

**University of Alberta**

Structural basis of TraD and *sbmA* recognition by TraM of F-like  
plasmids

by

Joyce Jia Wen Wong

A thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

©Joyce Jia Wen Wong

Fall 2012

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

## Abstract

Bacterial conjugation is the process of plasmid DNA transfer from a donor cell to a recipient cell. This process is mediated in F-like plasmids by proteins expressed by the *tra* operon. The relaxosome forms at *oriT* (origin of transfer), where the nicking and unwinding of a single-stranded copy of the plasmid begins, and the transferosome forms a transmembrane pore through which the DNA is transferred.

TraM is a tetrameric relaxosomal protein which binds to 3 sites at *oriT* – *sbmA*, *sbmB*, and *sbmC*. TraD is an inner membrane protein of the transferosome that is homologous to FtsK/SpoIIIE hexameric ATPases. The interaction between the C-terminal tail of TraM and TraD is essential for high conjugation efficiency. The structural basis of this interaction is revealed by the crystal structure of F TraM<sup>58-127</sup> in complex with TraD<sup>711-717</sup>. Electrostatic complementarity is a key feature of TraM-TraD interaction, which includes the TraM K99-TraD D715 and TraM R110-TraD F717 C-terminal carboxylate interactions. An additional feature is the fit of the phenyl side chain of F717 into a hydrophobic pocket. The importance of the TraD C-terminal tail for binding to TraM was tested with a pulldown assay comparing TraD constructs with and without a C-terminal truncation of 8 residues. *In vivo* assays confirmed the role of the C-terminal tail and its individual residues in conjugation.

TraM interacts with *sbmA* in a highly plasmid-specific manner. The basis of this is revealed by the crystal structure of pED208 TraM in complex with *sbmA*. The structure shows that the N-terminal domain of TraM is a dimeric ribbon-helix-helix fold which recognizes the DNA bases which make up the binding motif. Two

tetramers are bound to *sbmA* on opposite faces of the DNA without protein-protein contact, confirming the TraM-*sbmA* binding stoichiometry obtained from various biophysical methods. The cooperative nature of TraM binding to *sbmA* is therefore entirely through DNA distortions observed in the crystal structure, which include underwinding and kinking. Efforts to determine the structural basis of F TraM-*sbmA* interaction were undertaken but no diffraction-quality crystals were obtained.

## **Acknowledgements**

I would like to thank my supervisor, Mark Glover, for his guidance over the course of my graduate degree and giving me the chance to gain experience in a wide variety of scientific skills. I thank Laura Frost for her collaboration and insights in the area of plasmid conjugation, and for allowing me to work in her lab to learn additional techniques. Thanks also to Joanne Lemieux for serving on my supervisory committee.

I would like to thank Jun Lu for his help in getting me started in the Glover lab. I much appreciated his extensive knowledge of plasmid biology over the years and his intellectual contributions to the lab's conjugation projects. I also thank Ross Edwards –his expertise in crystallography data processing and Linux was invaluable for the work in this thesis and I am grateful for the patience in which he taught me those skills. Thanks also to Sheraz Khan for his assistance with the home x-ray generator and shipping of crystals for synchrotron data collection. Jan Manchak introduced me to cloning and other molecular biology techniques in the Frost lab –I was glad to learn from her. I much appreciated and would like to acknowledge Sue Smith's help in ensuring that administrative matters went smoothly.

Thanks to all the Glover lab members for their friendship and interesting conversations on all things scientific or not: the upstairs crew (Jun Lu, Nina Bernstein, Nico Coquelle, Zahra Havalishari, Gina Thede, Yun Peng, Ruth Green, and Sheraz Khan) and the downstairs crew (Stephen Campbell, Charles Leung, Ross Edwards, Danny Aceytuno, and Curtis Hodge). Also, thanks to the Frost lab members who made my "second home" on campus for a while, a much

better place: Isabella Lau-Wong, Jan Manchak, Barbara Klinger, Manuel Rodriguez-Maillard, Denis Arutyunov, and Bernadette Beadle.

## Table of Contents

	Page
<b>Chapter 1 – Introduction</b>	
Conjugation and bacterial genetic diversity, antibiotic resistance, and virulence	1
Mechanism and regulation of conjugation in F-like plasmids	3
Overview of relaxosome function	6
Proteins of the relaxosome and coupling protein complex	7
a) TraD	7
b) Tral	8
c) TraM	13
d) TraY	15
e) IHF	16
Relaxosome assembly, activity, and DNA topology	16
Prokaryotic DNA-binding motifs	18
Cooperativity in protein DNA binding	20
Thesis overview	23
Figures	24
References	31
<b>Chapter 2 – The crystal structure of the TraD C-terminus in complex with the TraM C-terminal domain</b>	
Overview	39
Introduction	40
Results	41

Discussion	47
Materials and Methods	48
Tables	51
Figures	53
References	58
<b>Chapter 3 - Functional Studies of TraD C-terminus -TraM interaction</b>	
Overview	60
Introduction	61
Results	62
Discussion	66
Materials and Methods	69
Tables	73
Figures	76
References	80
<b>Chapter 4 - Cooperative binding crystal structure of pED208 TraM</b>	
Overview	82
Introduction	83
Results	85
Discussion	97
Materials and Methods	100
Tables	109
Figures	113
References	127
<b>Chapter 5 - Specificity in pED208 TraM-DNA and TraM-TraD interactions</b>	
Overview	130

Introduction	131
Results	133
Discussion	138
Materials and Methods	142
Figures	145
Tables	147
References	156

## **Chapter 6 – Concluding Remarks**

Summary of Findings	157
Future Directions	160
References	164

## **Appendix A – Investigation of TraM binding to longer *oriT* fragments**

Introduction	165
Results	166
Discussion	170
Materials and Methods	172
Figures	176
References	182

## **Appendix B - Co-crystallization of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] with F *sbmA***

Introduction	183
Results	186
Discussion	189
Materials and Methods	191
Tables	193

Figures	194
References	197

## List of Tables

		Page
<b>Chapter 2</b>		
Table 2-1	Data collection and refinement of TraM <sup>58-127</sup> -TraD tail crystals	51
Table 2-2	Plasmids and oligonucleotides	52
<b>Chapter 3</b>		
Table 3-1	Effect of TraD mutations on F conjugation	73
Table 3-2	Effect of TraM mutations on F conjugation	73
Table 3-3	Compensatory mutations in TraD rescue TraM <sup>K99E</sup>	73
Table 3-4	Plasmids and oligonucleotides	74
<b>Chapter 4</b>		
Table 4-1	Data Collection and Refinement for pED208 TraM crystals	109
Table 4-2	Data Collection and Refinement for pED208 TraM crystals	110
Table 4-3	Data Collection for TraM-dIU4 <i>sbmA</i> complex	111
Table 4-4	Plasmids and oligonucleotides	112
<b>Chapter 5</b>		
Table 5-1	In vivo TraM-TraD tail interaction specificity between F and pED208 plasmids	145
Table 5-2	Plasmids and oligonucleotides	146
<b>Appendix B</b>		
Table B-1	Oligonucleotides for <i>oriT</i> and <i>sbmA</i> length variants	195
Table B-2	Variant <i>sbmA</i> oligos for crystallization optimization	195

## List of Figures

		Page
<b>Chapter 1</b>		
Figure 1-1	F plasmid <i>tra</i> operon overview	24
Figure 1-2	Coupling proteins	26
Figure 1-3	Relaxase structure and DNA-binding specificity	27
Figure 1-4	TraM C-terminal domain structure	28
Figure 1-5	Crystal structure of the IHF-DNA complex	29
Figure 1-6	Transcription factor structures	30
<b>Chapter 2</b>		
Figure 2-1	Crystallization and mass spectrometry of TraM <sup>58-127</sup> -TraD <sup>645-717</sup> crystals	53
Figure 2-2	Crystal structure and SigmaA-weighted 2mF <sub>o</sub> -DF <sub>c</sub> electron density of apo-TraM <sup>58-127</sup> determined from crystals grown from a TraM <sup>58-127</sup> -TraD <sup>707-717</sup> mixture.	54
Figure 2-3	Crystal structure of TraM <sup>58-127</sup> -TraD <sup>711-717</sup>	55
Figure 2-4	Sequence alignment of TraM and TraD homologues from F, pKPN3, And pED208	56
Figure 2-5	Comparison of the structure of the TraD-TraM complex with the unliganded TraM structure	57
<b>Chapter 3</b>		
Figure 3-1	Detailed view of TraM <sup>58-127</sup> -TraD <sup>711-717</sup> interactions	76
Figure 3-2	Affinity chromatography analysis of TraM interaction with TraD or its C-terminal truncation	77
Figure 3-3	Detailed view of potential TraD E712 and D715 interactions with TraM K99	78
Figure 3-4	Model of TraM avidity effect in binding to TraD	79
<b>Chapter 4</b>		
Figure 4-1	TraM binding stoichiometry to <i>sbmA</i>	113
Figure 4-2	Minimal length of <i>sbmA</i> required for binding of pED208 TraM	115
Figure 4-3	Evidence for the N-terminal domain of pED208 TraM being a ribbon-helix-helix fold	116
Figure 4-4	Crystals of pED208 TraM- <i>sbmA</i> complexes	117
Figure 4-5	Mass spectrometry confirmation of protein fragment in apo-N-terminal domain #2-52 crystals	118

Figure 4-6	Iodine peak from difference fourier map calculated after	119
Figure 4-7	Structure Solution of pED208 TraM- <i>sbmA</i> complex and TraM N-terminal domain	120
Figure 4-8	Crystal structure of pED208 TraM bound to <i>sbmA</i>	121
Figure 4-9	Crystal Structure of pED208 apo-N-terminal domain #2-52	123
Figure 4-10	Comparison of full-length and N-terminal domain pED208 TraM binding to 24bp <i>sbmA</i>	124
Figure 4-11	Cooperative recognition of <i>sbmA</i> DNA by TraM is mediated by DNA unwinding and kinking	125
Figure 4-12	Model for cooperative recognition of DNA by TraM	126
 <b>Chapter 5</b>		
Figure 5-1	Specificity of F-like plasmid TraM in binding to <i>sbmA</i>	147
Figure 5-2	DNA contacts of the pED208 TraM N-terminal domain	148
Figure 5-3	Role of residue-specific interactions in TraM- <i>sbmA</i> interaction	149
Figure 5-4	Cooperative binding of F TraM and pED208 TraM to a hybrid <i>sbmA</i>	150
Figure 5-5	Minimum length of F <i>sbmA</i> required for F TraM binding with full affinity	151
Figure 5-6	Possible specificity in TraM- <i>sbmA</i> interaction in from the TraM $\alpha$ 1- $\alpha$ 2 loop	152
Figure 5-7	Conserved mechanisms of allelic specificity of TraM- <i>sbmA</i> and TraM-TraD interaction	153
Figure 5-8	Top 9 conformations from docking of the F TraD tail to the F TraM tetramer in Autodock Vina	154
Figure 5-9	Top 9 conformations from docking of the pED208 TraD tail to the pED208 TraM tetramer in Autodock Vina	155
 <b>Appendix A</b>		
Figure A-1	EMSA of pED208 TraM binding to longer <i>oriT</i> fragments	179
Figure A-2	EMSA of F TraM binding to longer <i>oriT</i> fragments	180
Figure A-3	EMSA of pED208 TraM and IHF binding to <i>ihfC-sbmA</i> DNA	181
Figure A-4	Variation in binding affinity of pED208 TraM to <i>sbmA</i> with length of flanking DNA	183
 <b>Appendix B</b>		
Figure B-1	pED208 TraM- <i>sbmA</i> crystal contacts	196
Figure B-2	Crystallization hits of TraM [ $F^{1-55}$ :pED $^{56-127}$ ] – F <i>sbmA</i> complexes	197
Figure B-3	$\alpha$ 1- $\alpha$ 2 loop in F-like TraM	198

## List of Abbreviations

ADA	N-(2-Acetamido)iminodiacetic Acid
ATP	adenosine triphosphate
bp	base pair
BCA	bicinchoninic acid
Bis-Tris	Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane
BSA	bovine serum albumin
dIU	5'-iodo-2'-deoxyuridine
DNA	deoxyribonucleic acid
DSP	dithiobis(succinimidylpropionate)
DTBP	dimethyl 3,3'-dithiobispropionimidate
DTT	dithiothreitol
EMSA	electrophoretic mobility shift assay
GEC	genetic exchange community
HGT	horizontal gene transfer
H-NS	histone-like nucleoid structuring protein
HTH	helix-turn-helix
ICE	integrative conjugative element
IPTG	Isopropyl- $\beta$ -D-thio-galactoside
IHF	integration host factor
LB	Luria-Bertani
$K_d$	dissociation constant
kDa	kilodalton
MAD	multiple anomalous dispersion
MALDI-TOF	matrix-assisted laser desorption/ionization-time of flight
MALLS	multi-angle laser light scattering
MES	2-(N-morpholino)ethanesulfonic acid
MPD	2-methyl-2,4-pentanediol
NCS	non-crystallographic symmetry
Ni-NTA	nickel- nitrilotriacetic
NTP	nucleotide triphosphate
<i>oriT</i>	origin of transfer
PEG	polyethylene glycol
RHH	ribbon-helix-helix

RMSD	root-mean-square-deviation
RNA	ribonucleic acid
T4SS	Type IV secretion system
TBE	Tris-borate EDTA
TLS	<i>Translation Libration Screw-motion</i>
Tris	<i>tris</i> (hydroxymethyl)aminomethane
SAXS	small angle x-ray scattering
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
ssDNA	single-stranded DNA

## Chapter 1

### Introduction

#### *Conjugation and bacterial genetic diversity, antibiotic resistance, and virulence*

Conjugation, a form of horizontal gene transfer between bacterial cells, is an important contributor to bacterial genetic diversity. During the process of conjugation, a single-stranded copy of a plasmid is transferred from a donor bacterial cell to a recipient cell via the conjugative protein machinery expressed from the transfer operon on the plasmid (Llosa *et al.*, 2002, Juhas *et al.*, 2008). 17-25% of the *Escherichia coli* genome is thought to originate from horizontal gene transfer (HGT) (Narra & Ochman, 2006). Laterally-acquired genes are thought to be responsible for protein family expansion in 88-98% of genes across 8 genetically-distant bacterial clades (Treangen & Rocha, 2011). In fact, the concept of species boundaries among prokaryotes becomes a nebulous one, as more is discovered about the modes of horizontal gene transfer and their variations, and the extent of their contribution to bacterial genomes. Modes of HGT include transformation of DNA from the environment, phage transduction, conjugation, and integration into host genomes via homologous or non-homologous recombination and insertion sequences (Burrus & Waldor, 2004, Bennett, 2008, Skippington & Ragan, 2011, Boerlin & Reid-Smith, 2008). HGT allows for the creation of inter and intra-species genetic exchange communities (GECs), where the genetic material of all members of the GEC can be potentially shared among all others (Skippington & Ragan, 2011). The unfortunate consequence of HGT from a human health perspective is the wide and rapid

spread of antibiotic resistance genes among pathogenic bacteria (Boerlin & Reid-Smith, 2008, Bennett, 2008).

Conjugative plasmids are prevalent among both Gram-positive and Gram-negative bacteria, and have an important role in the rapid dissemination of antibiotic resistance and virulence factors in both. In one study, IncF1 replicons were detected in 28% of the *E. coli* strains tested, and in 35% of other *Enterobacteriaceae* (Mulec *et al.*, 2002). Conjugation is the primary mode for spread of extended-spectrum  $\beta$ -lactamases and AmpC  $\beta$ -lactamases among *Salmonella* and other *Enterobacteriaceae* (Su *et al.*, 2008). Also, plasmids are found in the vast majority of *Enterococcus* isolates and are the primary mode for the spread of antibiotic resistance among *E. faecium* and *E. faecalis* (Palmer *et al.*, 2010).

Upon integration into the bacterial genome via insertion sequence-mediated homologous recombination, chromosomal DNA can be transferred to bacteria of the same or different species after imperfect excision and Hfr-like transfer. This has been observed in F-like plasmids (Gubbins *et al.*, 2005), enterococcal pheromone-responsive plasmids (Palmer *et al.*, 2010), and integrative conjugative elements (ICEs) (Wozniak & Waldor, 2010). The F-derived plasmid pOX38 is capable of transfer to *Salmonella*, *Klebsiella*, and *Shigella* species (Mulec *et al.*, 2002), and evidence of horizontal propagation of transfer (*tra*) genes of the *E. coli* F plasmid have been found in a number of *Salmonella* strains (Boyd & Hartl, 1997). A pheromone-responsive antibiotic resistance plasmid in gram-positive *Streptococcus galactiae*, pIP501, was found to also be capable of transfer to gram-negative *E. coli* (Palmer *et al.*, 2010).

Plasmids of the IncF incompatibility groups are narrow host-range plasmids typically found in the *Enterobacteriaceae* family (Mulec et al., 2002). Examples include the prototypical F plasmid, and the R1, R100, and pED208 plasmids. IncF plasmids are relatively large. For example, the F plasmid is 99159 bases in length (Genbank accession number NC002483). Members of the F plasmid family are responsible for some of the earliest instances of antibiotic resistance, such as the emergence of multidrug resistant *Shigella* in Japan in the mid-1950's (Watanabe, 1963). F-like plasmids, many of them conjugative, continue to mediate a wide range of antibiotic resistance mechanisms in recent times (Conly, 2002, Strahilevitz *et al.*, 2009, Potron *et al.*, 2011). The conjugative ability of F-like plasmids has been directly linked to biofilm formation in *E. coli*, a virulence trait that allows the plasmids to be propagated throughout populations and facilitates the colonization of new environments (Ghigo, 2001). F-like replicons and portions of F-like *tra* systems are found in the majority of large virulence plasmids documented in *E. coli* and *Salmonella*, indicating a prominent role for F-like plasmids in their evolution. These large plasmids typically encode a combination of virulence factors like colicins, fimbriae, siderophores, *spv* genes, and hemolysins, as well as antibiotic resistance factors (Chu & Chiu, 2006, Johnson & Nolan, 2009, Porwollik & McClelland, 2003). The majority are non-transmissible because their transfer genes are incomplete but some of them are, like pSLT of *Salmonella typhimurium* (Ahmer *et al.*, 1999).

#### *Mechanism and regulation of conjugation in F-like plasmids*

Over the last decade, macromolecular structures have become available that provide insights into machinery of conjugation at the atomic level. The structural biology of conjugative Type IV secretion systems (T4SS) has been

reviewed extensively (Schroder & Lanka, 2005, Juhas et al., 2008, Alvarez-Martinez & Christie, 2009, Terradot & Waksman, 2011, Llosa *et al.*, 2009). The machinery of conjugation in F-like plasmids includes a DNA-processing complex (the relaxosome) that assembles on the plasmid's origin of transfer (*oriT*) and a Type IV secretion system (the transferosome) through which the DNA is transferred (Lawley *et al.*, 2003) with a coupling protein acting as the link between the two complexes (de la Cruz *et al.*, 2009). The proteins which carry out F conjugation are encoded on the *tra* operon, which consists of 3 promoters –  $P_M$ , from which TraM is transcribed,  $P_J$  from which TraJ is transcribed, and the  $P_Y$  promoter, from which the majority of *tra* genes are transcribed  $P_Y$  (Will *et al.*, 2004) (Figure 1-1A). TraM is an essential protein for conjugation that also autoregulates its own expression (Frost *et al.*, 1994). TraJ acts as an activator of transcription at  $P_Y$  by counteracting silencing by the bacterial global regulatory factor, H-NS (Will et al., 2004). TraY also activates  $P_Y$  expression and is part of the relaxosome (Silverman & Sholl, 1996) (Figure 1-1A).

Cell-cell contact is mediated via the pilus, which is a protein polymer consisting of the pilin subunit, TraA. After cell-cell contact, one strand of the plasmid DNA (the transfer strand) is unwound and is actively transported through the conjugative pore into the recipient cell. Being energetically expensive, conjugation is tightly regulated and highly responsive to physiological and environmental stimuli. F-plasmid transfer begins to decline in mid-exponential phase to undetectable levels in stationary phase, but is able to quickly become transfer positive when small amounts of glucose are added (Frost & Manchak, 1998). H-NS silencing of *tra* promoters is needed for the decrease in transfer efficiency in stationary phase (Will et al., 2004). F plasmid transfer is regulated by

a number of host factors responsive to environmental cues such as the Cpx system, which regulates response to extracytoplasmic stress, and the cyclic AMP receptor protein (Crp) and leucine-responsive regulatory protein (Lrp) proteins that regulate gene expression contingent upon nutritional availability (Figure 1-1A). CpxAR mediates degradation of TraJ via the HslVU and GroEL proteases in response to stress (Frost & Koraimann, 2010, Lau-Wong *et al.*, 2008). Crp and Lrp modulate TraJ transcription by binding to the DNA upstream of the TraJ gene (Gubbins *et al.*, 2005).

An additional mode of regulation in F-like plasmids is the FinO-FinP fertility inhibition system that represses transfer (Frost *et al.*, 1994, Gubbins *et al.*, 2005, Finnegan & Willetts, 1971, Frost *et al.*, 1989). The FinO/FinP system provides a mechanism for transfer repression which acts by reducing the level of TraJ protein, the main transcriptional activator of P<sub>Y</sub>. Translation of *traJ* mRNA is blocked by the 79-nt antisense RNA FinP, which is complementary to the 5'-UTR of *traJ* mRNA, blocking the ribosome binding site of P<sub>J</sub> (Frost, *et al.*, 1994, Gubbins, *et al.*, 2005). Regulation of *traJ* mRNA by FinP critically depends on a plasmid encoded protein, FinO. FinO is an RNA chaperone that increases the lifetime of FinP by protecting it from degradation by RNaseE (Jerome, *et al.*, 1999), while enhancing duplex formation of FinP and *traJ* mRNA (van Biesen & Frost, 1994) (Figure 1-1A). The crystal structure of a proteolytically stable fragment of FinO<sup>26-186</sup> showed that it forms a novel, largely  $\alpha$ -helical fold that is elongated due to an extended N-terminal  $\alpha$ -helix. (Ghetu, *et al.*, 2000). RNase protection experiments reveal that the lower half of the stem-loop and the 3'-tail single-stranded tail, are contacted by FinO in a manner that is dependent on the presence of a free 3'-hydroxyl (Arthur *et al.*, 2011). These protected regions in

FinP SLII, in addition to the FinO crosslinks determined in Ghetu et. al., 2002, and structural data from small-angle X-ray scattering (SAXS), were used as restraints in generating models for FinP-FinO interactions. (Arthur et al., 2011).

### *Overview of Relaxosome Function*

The relaxosome is a large multiprotein-DNA complex that forms at the plasmid origin of transfer. The protein components in F are TraI, TraY, TraM, and IHF. TraI is a bifunctional relaxase /helicase that recognizes the *nic* sequence within *oriT* and introduces a nick on the transfer strand that results in the covalent attachment of TraI to the 5' end of the nick. TraI then unwinds the DNA in a 5' →3' direction and is transported into the recipient cell along with the transfer strand (Dostal et al., 2011, Lang et al., 2010). TraY is an accessory protein that binds to two regions at *oriT* and to the P<sub>Y</sub> promoter (Nelson et al., 1993, Howard et al., 1995, Luo et al., 1994). TraM is thought to be the signaling molecule for the initiation of conjugative transfer, and binds to 3 sites at *oriT*. IHF is a small host-encoded protein that binds to several sites at *oriT*, and is necessary for relaxosome formation and nicking in F (Howard et al., 1995, Nelson et al., 1995) and F-like plasmids (Kupelwieser et al., 1998, Karl et al., 2001, Inamoto et al., 1994) (Figure 1-1C and 1-1D). Single-stranded DNA is transferred through the transferosome, a large transmembrane complex consisting of multiple Tra proteins that make up a Type IV secretion system (Lawley et al., 2003) (Figure 1-1B). The relaxosome is brought in close proximity to the transferosome through a key interaction between the coupling protein, TraD, and TraM (Disque-Kochem & Dreiseikelmann, 1997, Beranek et al., 2004, Lu et al., 2008). In general, interactions between relaxosomal *tra* components selectively occur between proteins of the same plasmid; heterotypic interactions are much less stable. This

has been demonstrated in TraD-TraM interaction, TraD-Tral interaction, and binding of Tral, TraY, and TraM to *oriT* DNA. In the following sections these will be described further.

### *Proteins of the relaxosome and coupling protein complex*

#### *a) TraD*

The coupling protein, TraD, is an integral membrane protein localized to the inner membrane of the bacterial cell envelope that interacts with TraM (Disque-Kochem & Dreiseikelmann, 1997). TraD of the F plasmid has a molecular weight of 82 kDa and contains 717 amino acids. It is a hexameric ATPase of the FtsK/SpoIII family (Gomis-Ruth *et al.*, 2001), consisting of an N-terminal membrane-spanning region and a C-terminal cytoplasmic domain that makes up the bulk of the protein (Frost *et al.*, 1994). Within the cytoplasmic domain are Walker A (G/AxxxxGKS/T) and Walker B (hhhhD, h=hydrophobic) nucleotide-binding boxes which are characteristic of F<sub>1</sub>-ATPase family of hexameric NTPases (Schroder *et al.*, 2002) (Figure 1-2A). TraD is able to bind to both single- and double-stranded DNA, with a preference for single-stranded DNA (Schroder *et al.*, 2002). Both the N- and C-terminal domains make contributions to oligomerization in TraD, and dimerize through interaction between a face on one side of the monomer and another face on the other side of the second monomer *in vivo*. These dimers then form hexamers in the presence of the F plasmid *in vivo* (Haft *et al.*, 2007). ATPase activity has been shown for TrwB (Figure 1-2B), the TraD homologue in the R388 plasmid transfer system (Tato *et al.*, 2007). Interestingly, TrwB ATPase activity is stimulated in the presence of the relaxosome accessory protein it interacts with, TrwA (Tato *et al.*, 2007). TraD contains a C-terminal extension (residues 577-717 of the F plasmid

TraD) beyond the ATPase domain (Figure 1-2A) which is not present in TrwB and its homologues. The last 38 amino acids (residues 680-717) are the minimal fragment shown to bind to TraM (Beranek et al., 2004).

*b) Tral*

Tral activity is modulated by several proteins and negative cooperativity between two domains for DNA binding. F plasmid Tral is a 182 kDa protein of 1756 amino acids consisting of a relaxase domain (~1-306) (Byrd *et al.*, 2002), two putative RecD-like helicase folds (~303-844 and ~830-1473) (Dostal & Schildbach, 2010), and a C-terminal domain of unknown function (~1476-1756) that also appears to be required for F conjugation (Guogas *et al.*, 2009). The relaxase domain is responsible for cleavage at *nic* through a conserved tyrosine, Tyr16. This tyrosine is part of a 4-tyrosine motif, YY-X<sub>5-6</sub>-YY (Tyr16, Tyr17, Tyr23, and Tyr24 in F Tral), which is largely conserved in the Mob<sub>F</sub> family of conjugative relaxases. The Tyr16 hydroxyl performs a nucleophilic attack on the phosphate backbone, becoming covalently attached to the 5'-phosphate of the transferred strand and leaving a free 3'-hydroxyl on the other strand (Byrd & Matson, 1997). Although other tyrosines exist near the active site in Tral and its homologues, only Tyr16 is essential for the nicking reaction to occur in F Tral (Dostal et al., 2011). Similar results have been found for TrwC, where the crucial tyrosine is also Tyr16 (Grandoso *et al.*, 2000). Binding and nicking activity of the relaxase at *nic* is highly sequence-specific, and therefore plasmid-specific (Stern & Schildbach, 2001, Harley & Schildbach, 2003, Gonzalez-Perez *et al.*, 2009, Fekete & Frost, 2000).

One classification scheme based on genetic lineage for conjugative relaxases divides them into the Mob<sub>F</sub>, Mob<sub>Q</sub>, Mob<sub>P</sub>, Mob<sub>H</sub>, Mob<sub>C</sub>, Mob<sub>V</sub> families (Garcillan-Barcia *et al.*, 2009). Crystal structures have been solved for Mob<sub>F</sub> class relaxases from three plasmids (F, pCU1, R388), and one Mob<sub>Q</sub> class relaxase from the plasmid R1162 (Garcillan-Barcia *et al.*, 2009). These are of the relaxase domain of R388 TrwC (Guasch *et al.*, 2003), F Tral (Datta *et al.*, 2003), pCU1 relaxase (Nash *et al.*, 2010), and R1162 MobA (Monzingo *et al.*, 2007), representing the IncW, IncF, IncN, and IncQ incompatibility groups, respectively. All structures share a conserved fold, consisting of a 5-stranded  $\beta$ -sheet “palm”, with a pair of long  $\alpha$ -helices on one face and 2 largely  $\alpha$ -helical domains on the DNA binding face (Figure 1-3A). The  $\alpha$ -helical flap that closes over the bound DNA are the “fingers” that become ordered upon binding (Larkin *et al.*, 2005). The 3 Mob<sub>F</sub> structures are very similar to each other while showing some differences with the Mob<sub>Q</sub> structure (R1162). The Mob<sub>F</sub> structures have a protruding 2-stranded  $\beta$ -sheet (“ $\beta$ -protrusion”) that contacts the DNA hairpin which is missing in the Mob<sub>Q</sub> (MobA) structure. Instead MobA has an  $\alpha$ -helical protrusion nearby (Monzingo *et al.*, 2007). Structures of relaxase-*nic* DNA complexes have revealed that the relaxase binds to a single-stranded DNA U-turn stabilized by intramolecular contacts between the DNA bases (Guasch *et al.*, 2003, Larkin *et al.*, 2005). The TrwC-DNA structure shows that there are contacts between the protein and the hairpin on the 3' side of the *nic* site, however the significance of this contact has yet to be shown (Guasch *et al.*, 2003). There is strong structural conservation of a triple-histidine divalent cation coordination site in close proximity to the active site tyrosines. An enzymatic mechanism where the active site metal catalyzes nucleophilic attack by the tyrosine hydroxyl is likely (Larkin *et al.*, 2005, Boer *et al.*, 2006).

The relaxases of F-like plasmids show a high level of binding specificity to the *nic* site of its cognate plasmids. Harley and Schildbach (2003) have shown that Tral of F and R100 plasmids bind to their cognate *nic* sites (Figure 1-3B) three orders of magnitude more tightly than to the *nic* site of the non-cognate plasmid. This selectivity is largely due to the interactions of a non-conserved pair of amino acid residues, Gln 193 and Arg 201 in F Tral, and a pair of single-stranded bases at 145' and 147' (according to the base-numbering scheme of the *nic* site in (Frost et al., 1994) (Figure 1-3C). The specificity of binding can be swapped to some extent between R100 and F by switching residues only at these positions (Harley & Schildbach, 2003). The crystal structure of Tral bound to *nic* DNA bases 144'-153' provides an explanation for the role of Gln193, Arg201, G145', and G147' in binding specificity. In addition to revealing hydrogen bonds between the DNA bases and the side chains, Arg201 forms part of a pocket entered by G147' (Larkin et al., 2005) (Figure 1-3D). Comparison between the structures of F Tral and R388 TrwC (Boer et al., 2006) reveal further plasmid specificity. None of the above-mentioned specificity determinants are conserved. Residues corresponding to that of F Tral Gln193 and Arg 201, Thr189 and Asn197 of TrwC, are not appropriately positioned for interaction with bases in the R388 *nic* site corresponding to F 145' and 147'. Instead, a hydrogen bond is formed between His4 and A19, and Arg 190 forms a cation-pi stacking interaction with T21 (Figure 1E). A further site of specific binding is at the position immediately 5' to the *nic* site, which is T in TrwC but is G in the *nic* sites of other F-like plasmids (Figure 1C). It was predicted that Lys 262, which interacts with T, would be precluded from interaction with guanine in other F-like relaxases by the bulkier G side chain, and is supported by site-directed mutagenesis (Gonzalez-Perez et al., 2009).

The N-terminal domain is followed by the helicase domain and the C-terminal domain that may interact with TraM (Lang et al., 2010, Ragonese *et al.*, 2007). Of the 2 putative helicase folds, the C-terminal fold is the functional helicase, while the N-terminal fold functions as a binding domain for ssDNA, possibly as a negative autoregulatory domain for Tral activity (Haft *et al.*, 2006). Supporting this, the C-terminal fold contains helical motifs and a  $\beta$ -hairpin required for helicase activity in *E. coli* RecD, but the N-terminal fold does not. In addition, ssDNA binding ability has been localized to the N-terminal fold, possibly residues 309-331, while the C-terminal fold shows no ssDNA DNA binding ability (Dostal & Schildbach, 2010). Both helicase domains have been shown to contain translocation signals essential for transfer into the recipient cell (Lang et al., 2010). The crystal structure of the region C-terminal to the helicase domain consisting of residues 1476-1629 of F Tral has been solved. It contains separate  $\alpha$  and  $\alpha$ - $\beta$  domains joined by a proline-containing loop that may supply rigidity to their relative orientations. Although truncations in this region are very detrimental to conjugation, the precise function of this region is yet to be determined (Guogas et al., 2009). Since models were found for all the major domains of Tral, a model of full-length Tral was constructed, using a SAXS envelope derived from full-length Tral. The SAXS data shows that Tral has an elongated, linear conformation in solution (Cheng *et al.*, 2011).

Several findings indicate that there is negative cooperativity in single-stranded DNA binding between the relaxase and N-terminal helicase domains of Tral. Truncation of the relaxase domain leads to greater unwinding activity compared to the full-length protein (Sut *et al.*, 2009). Twice as much DNA as expected was required to reach binding saturation with the full-length protein,

indicating that binding of the relaxase site interferes with binding to the helicase site (Dostal & Schildbach, 2010). High affinity binding of the relaxase domain to the DNA hairpin formed by an inverted repeat 3' to *nic* is hypothesized to act as a “switch” between an inactive state to a helicase active state (Sut et al., 2009, Dostal & Schildbach, 2010, Mihajlovic *et al.*, 2009).

The nature of Tral interaction with transferosome components still needs to be clarified. Tral nicking activity is stimulated by the presence of TraD (Mihajlovic et al., 2009). The region responsible for this is Tral<sup>1-992</sup>, which is transported to the recipient cell while it is attached to the transferred plasmid DNA (Dostal et al., 2011, Lang et al., 2010), therefore interaction with the conjugative pore is necessary. Evidence suggests that this interaction occurs in a sequence-specific manner through its translocation sequences. Residue Leu626 in the first translocation sequence of F Tral is essential for transfer. Only TraD from the cognate plasmid can counteract the dominant negative effect of excess Tral when both are expressed in trans in an F system (Lang et al., 2010). Tral remains in the soluble cytoplasmic fraction when expressed, but colocalizes with TraD in the membrane fraction when TraD is co-expressed (Dash *et al.*, 1992). These results suggest that the interaction of Tral with the transferosome involves TraD in some manner, whether direct or indirect. Direct interaction of coupling proteins with the relaxase has been reported in R388, RP4, and the RP4-mobilizable plasmids pBHR1 and pLV22a (Llosa *et al.*, 2003, Szpirer *et al.*, 2000, Schroder et al., 2002, Thomas & Hecht, 2007). It has been hypothesized that there is a signaling conduit from TraD through the Tral<sup>1-992</sup> for export or import of substrates through the T4SS (Lang et al., 2010, Lang *et al.*, 2011).

### c) TraM

DNA nicking by the Tral relaxase is enhanced by the relaxosome protein, TraM (Ragonese *et al.*, 2007, Kupelwieser *et al.*, 1998, Karl *et al.*, 2001), and evidence suggests that the mechanism of enhancement is indirect through DNA rather than through direct protein-protein interaction between Tral and TraM (Mihajlovic *et al.*, 2009). Tral helicase activity is also enhanced (Sut, *et al.*, 2009). F plasmid TraM binds to three sites, *sbmA*, *sbmB*, and *sbmC*, at *oriT* that contain plasmid-specific DNA binding motifs. These sites overlap with the two TraM promoters, and as a result, TraM negatively regulates its own expression (Penfold *et al.*, 1996) (Figure 1-1A). Binding of TraM to these sites is cooperative, with the highest affinity binding site being *sbmA* (Fekete & Frost, 2002). In addition, TraM interacts with the coupling protein TraD of the transferosome, forming a physical tether between the transferosome and relaxosome (Disque-Kochem & Dreiseikelmann, 1997, Beranek *et al.*, 2004).

TraM is a tetrameric protein consisting of a C-terminal tetramerization domain (Miller & Schildbach, 2003, Verdino *et al.*, 1999) and an N-terminal dimerization and DNA binding domain (Kupelwieser *et al.*, 1998, Miller & Schildbach, 2003, Lu *et al.*, 2004, Schwab *et al.*, 1993). Oligomerization of TraM is essential for TraM binding to DNA and for *in vivo* function (Lu *et al.*, 2004). (Figure 1-4). An NMR structure of the N-terminal domain of R1 TraM at pH 4.0 showed a monomeric structure consisting of 3  $\alpha$ -helices (Stockner *et al.*, 2001). The crystal structure of the C-terminal domain of F TraM showed that it forms an 4-fold symmetrical  $\alpha$ -helical bundle (Lu *et al.*, 2006). Random mutagenesis of TraM revealed 3 functional categories of mutants (DNA-binding, dimerization, and tetramerization), which were consistent with previous functional assignments

of TraM regions (Lu et al., 2004) (Figure 1-4B). A fourth class of mutants that was defective in conjugation but did not show defect in autoregulation or tetramerization was discovered, and their role was shown to be that of interaction with TraD. The most deleterious of these was K99E (Lu & Frost, 2005).

The crystal structure of the TraM tetramerization domain reveals the mechanism of how it regulates conjugation in response to increased pH or temperature (Lu et al., 2006). Glu88 is buried within the hydrophobic core of the helical bundle (Figure 1-4). Basic pH and increased temperature result in its protonation, leading to decreased tetramer stability, DNA binding ability, and conjugation. In addition, tetramerization is essential for interaction of TraM with TraD as mutation of Glu88 to a non-ionizable leucine led to persistence of TraM-TraD interaction at increased temperature (Lu et al., 2006). Thus, the deprotonation of Glu88 appears to be a direct mechanism by which conjugation can be repressed in non-optimal pH and temperature. This residue is conserved among the IncFI and FII plasmids F, R1, and R100, but is not in others like the IncFV plasmid pED208. It remains to be seen if TraM from pED208 or other plasmids exhibit the same pH and temperature-dependent stability.

A protein with an analogous function to TraM in the R388 plasmid is TrwA, a relaxosome component with a putative RHH (ribbon-helix-helix) fold and a C-terminal tetramerization domain (Moncalian & de la Cruz, 2004). The N-terminal domain is the DNA-binding domain, and the C-terminal domain is a tetramerization domain that interacts with TrwB, the coupling protein of the R388 system (Llosa et al., 2003). It also functions as a negative transcriptional regulator of the *trw* operon and enhances activity of TrwC, the relaxase

(Moncalian *et al.*, 1997). The TrwA-TrwB interaction is more than simply a bridge between the relaxosome and transferosome, as TrwA affects the ATPase activity and oligomerization state of TrwB. In the absence of TrwA and DNA, TrwB is a monomer with weak ATPase activity. Both TrwA and DNA stimulate TrwB's ATPase activity and formation of TrwB hexamers (Tato, *et al.*, 2007). Whether this also occurs in the F-plasmid has yet to be shown. However, evidence exists that TraD occurs in dimers *in vivo* in the absence of the F plasmid, but forms higher order oligomers when F is present (Haft *et al.*, 2007). This suggests that F plasmid proteins, possibly TraM, are required for hexamer formation.

#### d) *TraY*

TraY of F-like plasmids regulates P<sub>Y</sub> promoter activity (Taki *et al.*, 1998, Silverman & Sholl, 1996) and stimulates the activity of TraI (Howard *et al.*, 1995, Karl *et al.*, 2001) when bound to its DNA sites (*sby*) (Figure 1-1C). As is not uncommon with transcription factors which act at *tra* operon promoters, TraY has opposite effects in different plasmids. In the F plasmid, TraY stimulates P<sub>Y</sub> transcription (Maneewannakul *et al.*, 1996, Silverman & Sholl, 1996) but in the R100 plasmid, it represses it (Taki *et al.*, 1998) .

Ribbon-helix-helix (RHH) domains are a commonly-used DNA-binding module in prokaryotes which consists of a homodimer where the N-terminal residues form a short N-terminal  $\beta$ -sheet followed by two  $\alpha$ -helices (Schreiter & Drennan, 2007). Three alternating residues of the  $\beta$ -sheet on the protein's surface are specificity determinants for contacting DNA bases in the major groove of the binding site. The relaxosome accessory proteins TrwA of the R388 plasmid and TraY of F-like plasmids are putative RHH-folds (Lum & Schildbach,

1999, Moncalian & de la Cruz, 2004, Moncalian et al., 1997). TraY is encoded by a single polypeptide consisting of two predicted RHH domains in tandem (Bowie & Sauer, 1990). Mutagenesis of the predicted DNA-contacting  $\beta$ -sheet residues resulted in DNA binding defect, supporting the predicted RHH fold (Lum & Schildbach, 1999). Upon binding to DNA, it induces a bend of  $\sim 50^\circ$  (Lum & Schildbach, 1999).

e) *IHF*

IHF binding to *oriT* is necessary for relaxosome formation and nicking in F (Howard, et al., 1995, Nelson, et al., 1995) R1 (Kupelwieser et al., 1998, Karl et al., 2001), and R100 (Inamoto et al., 1994). A crystal structure was solved for IHF bound to  $\lambda$  phage DNA (Rice et al., 1996). The IHF heterodimer induces a  $160^\circ$  bend when bound to the minor groove of DNA (Rice et al., 1996) (Figure 1-5). The ability of IHF to stimulate plasmid transfer is highly dependent on the spacing and rotational orientation of the binding site on the DNA helix relative to those of other relaxosome components (Williams & Schildbach, 2007).

*Relaxosome assembly, activity, and DNA topology.*

The order of assembly of the relaxosome components, and their effects on relaxosome and *tra* operon function varies from plasmid to plasmid. In the F plasmid, Tral nicking activity is much weakened unless Tral and TraY are added first (Howard et al., 1995). TraY may not be needed for relaxosome function in pED208 because there is an insertion element, IS2, in the pED208 TraY gene (Finlay et al., 1986, Lu et al., 2002). In R1, TraM can substitute for TraY in stimulation of Tral *oriT* nicking activity (Kupelwieser et al., 1998, Karl et al., 2001), while TraM in the F plasmid cannot (Ragonese et al., 2007). Furthermore,

in R1, plasmid mobilization can occur at efficiencies comparable to wild-type in the absence of TraY (Karl et al., 2001), while TraY insertion mutants in the F plasmid reduced mobilization to low levels (Maneewannakul et al., 1996). However, the majority of the TraY sequence is located downstream of IS2 in pED208, with a putative promoter and a weak ribosome binding site located upstream of a start codon (Finlay et al., 1986), so the possibility that a functional TraY product is produced cannot be ruled out at this time.

Indirect evidence suggesting an interaction between TraM and the C-terminal domain of Tral has been reported by one group (Ragonese et al., 2007), but 2 other groups could not confirm the interaction (Guogas et al., 2009) (Wong, J.J.W., and Lu, J., unpublished observations). Plasmid-specific TraD-Tral interaction has been strongly suggested by mutagenesis studies (Lang et al., 2010).

Many factors suggest a complex DNA topology at *oriT*. The DNA bound by relaxosome components is subject to bends induced by TraY (~50°) (Lum & Schildbach, 1999) and IHF (~160°) (Rice et al., 1996). *oriT* sequences of all the commonly studied F-like plasmids contain a fair number of poly-AT tracts (Ostermann *et al.*, 1984, Di Laurenzio *et al.*, 1991, Frost et al., 1994, Abo & Ohtsubo, 1995), which are intrinsically bent and subject to many structural subtleties depending on the exact sequence of the tract and whether consecutive As are 5' or 3' to consecutive Ts. These poly-AT tracts may also affect the structure of the flanking non-AT regions (Haran & Mohanty, 2009). The presence of TraM induces underwinding of plasmid DNA (Mihajlovic et al., 2009). The orientation of relaxosome components relative to each other is important, as

several studies show that alteration of spacing between *oriT* binding sites decreases transfer efficiency. In R100, insertion of base pairs between *sbyA* and *sbmA* was detrimental to transfer (Abo & Ohtsubo, 1995). In F, transfer is phase variation-dependent in the presence of deletions or insertions between *sbyA* and *sbmC* (whole turns are tolerated whereas half turns are not) (Fu *et al.*, 1991), and transfer is greatly reduced in the presence of insertions between *ihfA* and *sbyA* while base substitutions are fairly well tolerated (Williams & Schildbach, 2007).

### *Prokaryotic DNA-binding motifs*

Transcription factors in prokaryotes very often contain one of 2 folds as their DNA binding module - the helix-turn-helix (HTH) fold or ribbon-helix-helix (RHH) fold. The core of the HTH fold consists of 3  $\alpha$ -helices which form an open helical bundle, where the  $\alpha_3$  is the “recognition helix” that forms specific contacts with DNA bases in the major groove. Many variations on the HTH fold exist, including 4-helical bundles, and winged-HTH folds with an additional 2-4 stranded  $\beta$ -sheet (Aravind *et al.*, 2005). The wings make contacts with the DNA in diverse ways –they may contact in various combinations, the phosphate backbone, bases in the minor groove, or bases in the major groove (Huffman & Brennan, 2002, Kenney, 2002) (Figure 1-6A).

The RHH fold is a homodimer where the N-terminal residues form a short antiparallel  $\beta$ -sheet and are followed by 2  $\alpha$ -helices. The  $\beta$ -sheet specifically recognizes bases in the DNA major groove while the N-terminal end of  $\alpha_2$  forms non-specific contacts with the DNA backbone (Schreiter & Drennan, 2007) (Figure 1-6B). The RHH motif is speculated to have evolved from the HTH motif,

due to the similarity of the HTH  $\alpha$ 2-loop- $\alpha$ 3 fold to the RHH  $\alpha$ 1-loop- $\alpha$ 2 fold (Aravind et al., 2005).

Both families of folds have been used as the DNA recognition modules of transcription factors with a wide variety of functions. HTH or RHH folds are very often part of a larger protein which contains a ligand-responsive and/or oligomerization domain. Therefore, bacterial transcription factors of these two families frequently regulate transcription in response to small molecule stimuli. Binding of a ligand causes allosteric changes in the protein structure that makes it more or less amenable to binding its operator site. Often, multiple copies of the protein bind cooperatively to operator DNA. The HTH family includes many of the “classical” nutrition-responsive transcription factors, like catabolite repressor protein (Crp) (Lawson *et al.*, 2004) (Figure 1-6A), the Lac repressor (Lewis, 2005), AraC (Schleif, 2010), and TrpR (Youderian & Arvidson, 1994). Other functions mediated by HTH transcription factors include antibiotic resistance (TetR family) (Ramos *et al.*, 2005), heavy metal resistance (ArsR-SmtB family, Fur family) (Osman & Cavet, 2010), response regulation in 2-component signal transduction systems (OmpR family) (Kenney, 2002) and plasmid partitioning (ParB, SopB, KorB) (Schumacher & Funnell, 2005, Schumacher *et al.*, 2010, Khare *et al.*, 2004). In addition to binding small molecules, HTH proteins may interact with other proteins or with DNA. In 2-component systems, the DNA-binding of the HTH-containing response regulator is affected by phosphorylation by the membrane-bound sensor kinase. ParB homologues form DNA bridges to position plasmids for segregation, and may form complex structures with multiple variations due to a high degree of linker flexibility between the HTH and

dimerization domains, well as the DNA-binding ability of the dimerization domain (Schumacher *et al.*, 2007, Khare *et al.*, 2004).

Like HTH domains, RHH-containing transcription factors mediate a wide variety of functions which also include plasmid partitioning (ParG and omega), plasmid copy number control (CopG), nutritional response (MetJ) (Figure 1-6B), metal homeostasis (NikR), and post-segregational killing of cells that do not receive a copy of the plasmid following cell division (FitA, ParD) (Schreiter & Drennan, 2007). HTH and RHH proteins are also found in bacteriophages, where the extensively characterized HTH-containing  $\lambda$  cI repressor (Stayrook *et al.*, 2008), and RHH-containing Arc repressor (Raumann *et al.*, 1994) regulate the transition between the lytic and lysogenic stages.

Ribbon-helix-helix (RHH) protein folds are widely distributed among the plasmid kingdom, as members have been found in the Mob<sub>F</sub>, Mob<sub>P</sub>, and Mob<sub>Q</sub> classes. In addition to TraY and TrwA, a family of relaxosome accessory proteins predicted to form RHH folds, represented by MbeC of the ColE1 plasmid has been discovered (Varsaki *et al.*, 2009). The structures of several RHH proteins with conjugative regulatory functions have recently been solved, including ArtA of the plasmid pSK41 in complex with its consensus DNA (Ni *et al.*, 2009), NikA, an auxiliary relaxosome component of plasmid R64 (Yoshida *et al.*, 2008), and an unusual RHH protein of the Ti plasmid, VirC2, where the RHH fold is formed by a single polypeptide chain (Lu *et al.*, 2009).

#### *Cooperativity in protein-DNA binding*

Cooperativity in protein-ligand interactions is defined as the increased affinity of ligand binding to other protein subunits following the initial binding of the ligand to one of the subunits. The degree of cooperativity can be determined experimentally by measuring the degree of saturation of the protein by the ligand at different ligand concentrations and plotting the data using a Hill plot. This uses the linear form of the Hill equation,  $\log(\theta/1-\theta) = n \cdot \log[L] - \log[K_D]$ . By measuring the slope of the plot of  $\log(\theta/1-\theta)$  vs.  $\log[L]$ , a Hill coefficient can be determined. When the Hill coefficient is greater than one, it indicates positive cooperativity between the protein subunits, and if less than one, negative cooperativity (Nelson & Cox, 2004). In the case of protein binding to DNA, the presence of cooperativity can also be detected by visual inspection of the gel bands resulting from an electrophoretic mobility shift assay (EMSA). Instead of sequential loading of proteins onto the binding sites, resulting in multiple shifted species corresponding to occupancy of each binding site of the protein, the presence of a single shifted species is an early indicator of cooperativity, as occupancy of one site appears to increase the affinity of the other protein subunit(s) for the others so that intermediate species are not present in significant quantity (Fekete & Frost, 2002, Phillips *et al.*, 1989). Another method for visualization of cooperativity is described in (Hochschild & Ptashne, 1986). Binding of the lambda repressor to a single operator site results in a visible DNA footprint at a higher concentration than when two operator sites are present (Hochschild & Ptashne, 1986).

Cooperative recognition of DNA is very common among transcription factors. Cooperativity is typically mediated through protein-protein interactions between the DNA-binding proteins, often through a second oligomerization

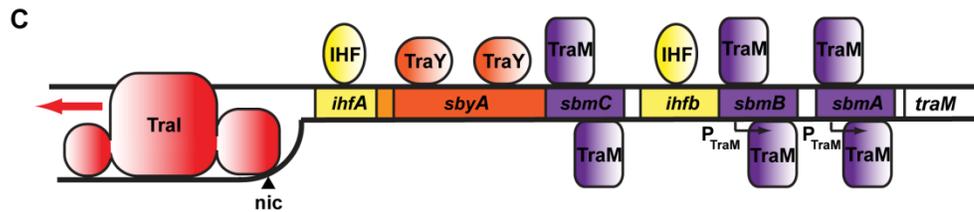
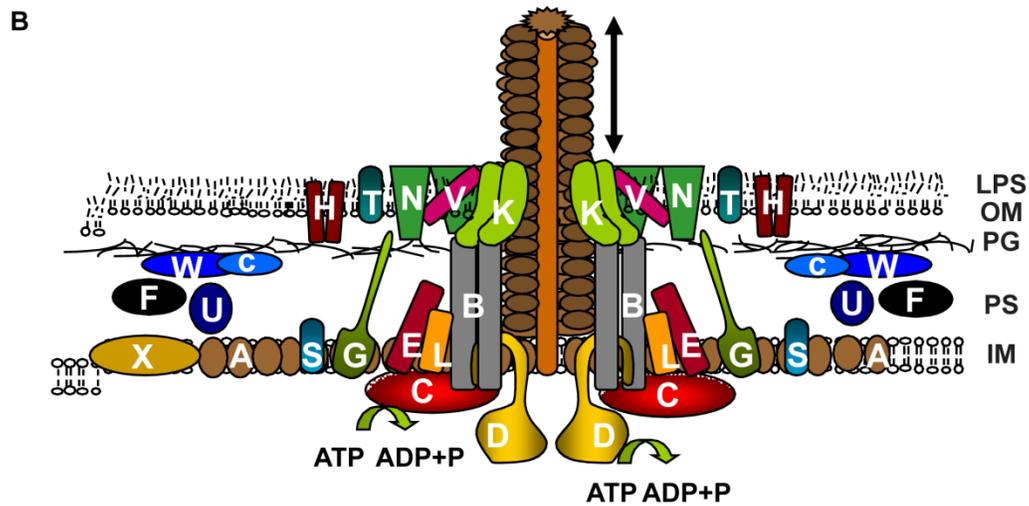
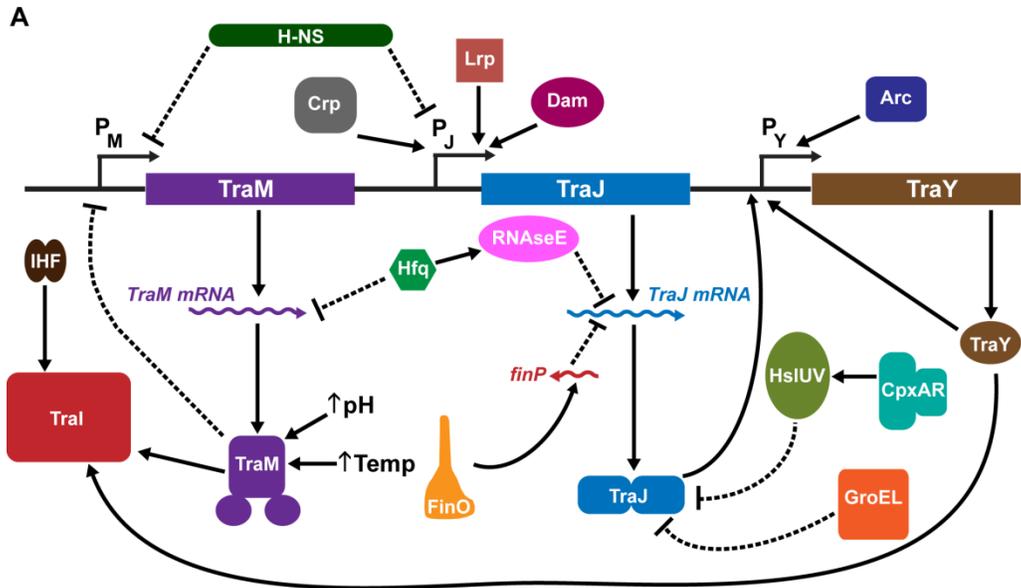
domain that may also function as an allosteric sensor to intracellular molecules. MetJ, is one such example – it represses methionine biosynthesis with maximal affinity when bound to a methionine catabolite, S-adenosylmethionine (Weissbach & Brot, 1991). The dimeric RHH domains interact via the C-terminal sensor domain with another MetJ dimer bound to a neighboring operator site (Somers & Phillips, 1992) (Figure 1-6B), which is how it appears to bind to these sites in a cooperative manner with high affinity ( $K_d=45$  nM) (Phillips et al., 1989). Another example is the Lac repressor, where two dimers, each with HTH domains bound to the major groove of one operator, interact with another dimer bound to a distant operator site via its ligand-binding domain (Senear & Brenowitz, 1991, Lewis *et al.*, 1996). Numerous other examples where the crystal structure have revealed the mechanism of cooperative DNA binding are the lambda repressor (Stayrook et al., 2008), p53 (Malecka *et al.*, 2009), Arc (Smith & Sauer, 1995), and NikR (Schreiter *et al.*, 2006). However, there exists a relatively small number of transcription factors that bind cooperatively to DNA without any protein-protein interaction. These include the TetR (Ramos et al., 2005) (Figure 1-6C), DtxR/MntR (Guedon & Helmann, 2003), and SMAD (Baburajendran *et al.*, 2011) (Figure 1-6D) families of transcription factors.

Cooperativity also occurs between proteins on operator sites separated by substantial numbers of DNA bases. This has been shown to occur between TraM binding sites at *oriT* in the F plasmid (Fekete & Frost, 2002). Classical examples of this are the Lac and the lambda repressors, which are thought to form DNA loops which enhance repression. The lac repressor dimerization domain is able to homodimerize, while both are bound to different operator DNA sites (Bell & Lewis, 2001). One operator site is ~80 bp from the

central site, and the other is ~410 bp away. Even so, Lac is able to enhance repression via operator sites up to 1000 bp apart (Lewis, 2005). The lambda repressor tetramer-DNA complex is able to bind to another lambda tetramer to form an octamer, bridging together operator sites that are ~2400 bp apart (Stayrook et al., 2008).

### *Thesis overview*

In the following thesis the structural characterization of F-like plasmid TraM-TraD interaction and TraM-*sbmA* interaction, as well as the probing of those interactions by *in vitro* and *in vivo* methods will be presented and discussed. The crystallization and analysis of the crystal structure of the C-terminal domain of F TraM bound to the C-terminal 7 residues of its cognate TraD is described in Chapter 2. Chapter 3 investigates the effect of the TraD tail on TraM binding by pulldown assays, and the effect of the tail and individual residues of the TraD tail and TraD binding pocket in TraM on conjugation *in vivo*. In Chapter 4, the crystal structure of pED208 TraM bound to *sbmA* and its apo-N-terminal domain is described. The nature of the cooperative binding of TraM tetramers to *sbmA* is probed by various biophysical methods. In Chapter 5, the specific interaction of TraM with *sbmA* DNA is discussed in more detail and the role of these interactions in binding is tested *in vitro*. Specificity in TraM-DNA binding in the pED208 system is investigated. The thesis work is summarized and future directions for research are proposed in Chapter 6. Appendix A covers the binding of pED208 TraM and IHF to *oriT*. Appendix B documents the attempts to obtain diffraction-quality crystals of a chimeric TraM in complex with F *sbmA* in order to gain information on F TraM binding to its cognate *sbmA*.



D

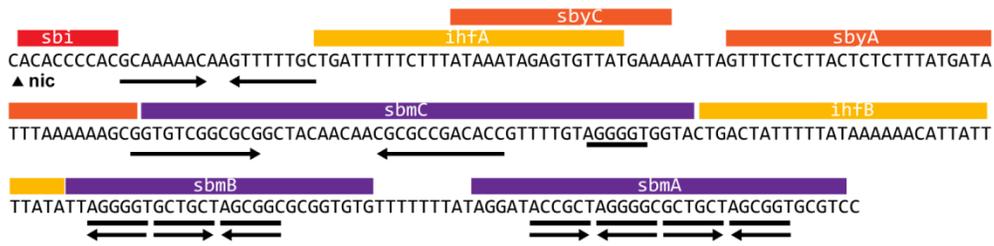


Figure 1-1 F plasmid *tra* operon overview

- tra* operon regulation by plasmid and host factors
- Schematic of the transferosome and its Tra protein components. LPS=lipopolysaccharide. OM=outer membrane. PG=peptidoglycan. PS=periplasmic space. IM=inner membrane. The ability of the pilin subunits (brown ovals) to polymerize and depolymerize to extend and retract the pilus is indicated by the double-ended arrow (Lawley et al., 2003)
- Relaxosome proteins and their binding sites at the F *oriT*
- F plasmid *oriT* sequence. TraM binding motifs are underlined. Inverted repeats are indicated by arrows.

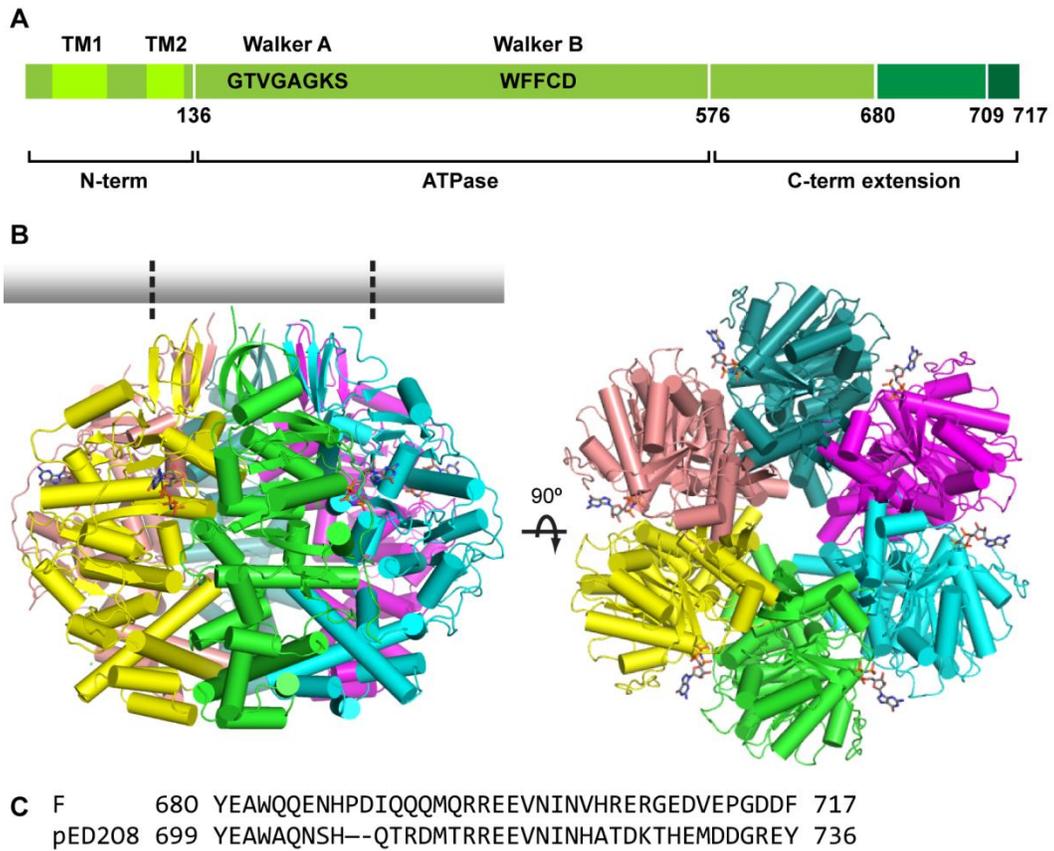


Figure 1-2 Coupling proteins

- a) TraD. Residues shown to bind TraM are shown in dark green.  
 b) TrwB cytoplasmic domain crystal structure (PDB: 1GL6) (Gomis-Ruth et al., 2001). Inner membrane is shown by a shaded grey box. GDPnP is shown in cyan sticks.  
 c) Sequence alignment of the last 38 amino acids of F and pED208 TraM

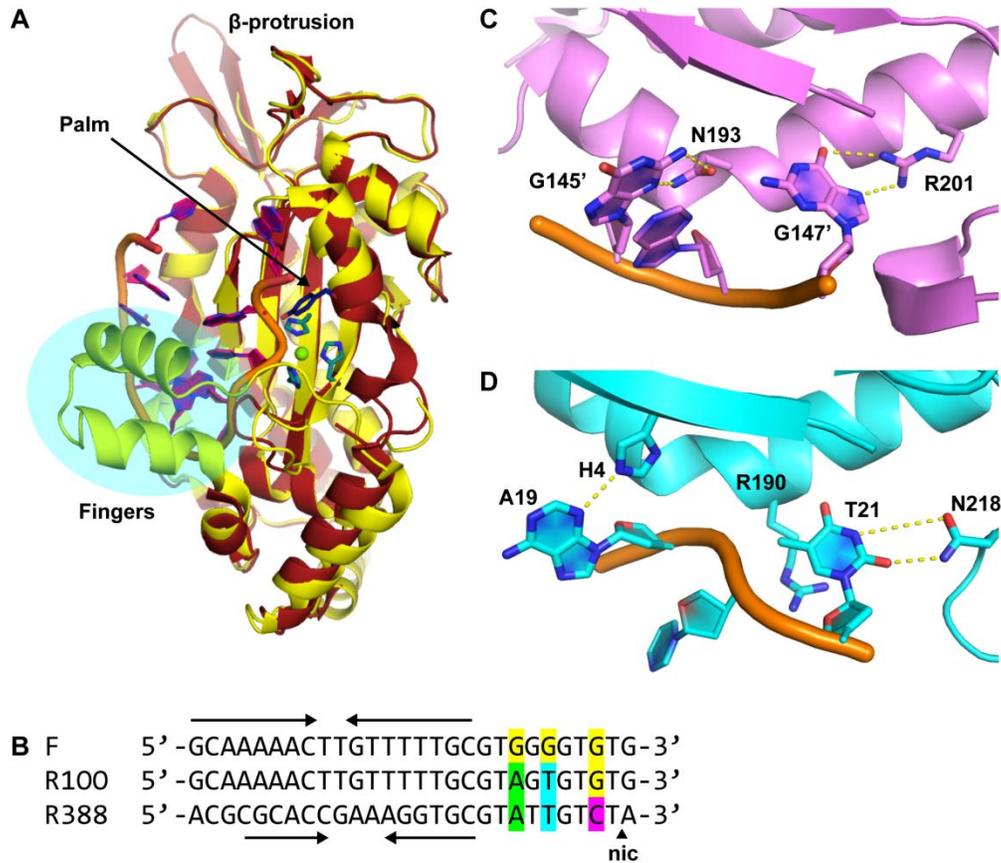


Figure 1-3 Relaxase structure and DNA-binding specificity

- a) Alignment of crystal structures from F Tral in its apo-form (red) (PDB: 1P4D) (Datta et al., 2003) and DNA-bound form (yellow) (PDB: 2A0I) (Larkin et al., 2005). *oriT* DNA from the Tral-DNA complex is shown in fuchsia. F Tral H146, H157, H159 are shown in blue, and Y16 in dark blue.  $Mg^{2+}$  is shown as a green sphere
- b) Alignment of relaxase binding sequences adjacent to *nic* at *oriT*. The inverted repeat is indicated by arrows.
- c) F Tral residue-specific interactions with DNA bases adjacent to *nic*
- d) R388 TrwC residue-specific interactions with DNA bases adjacent to *nic*

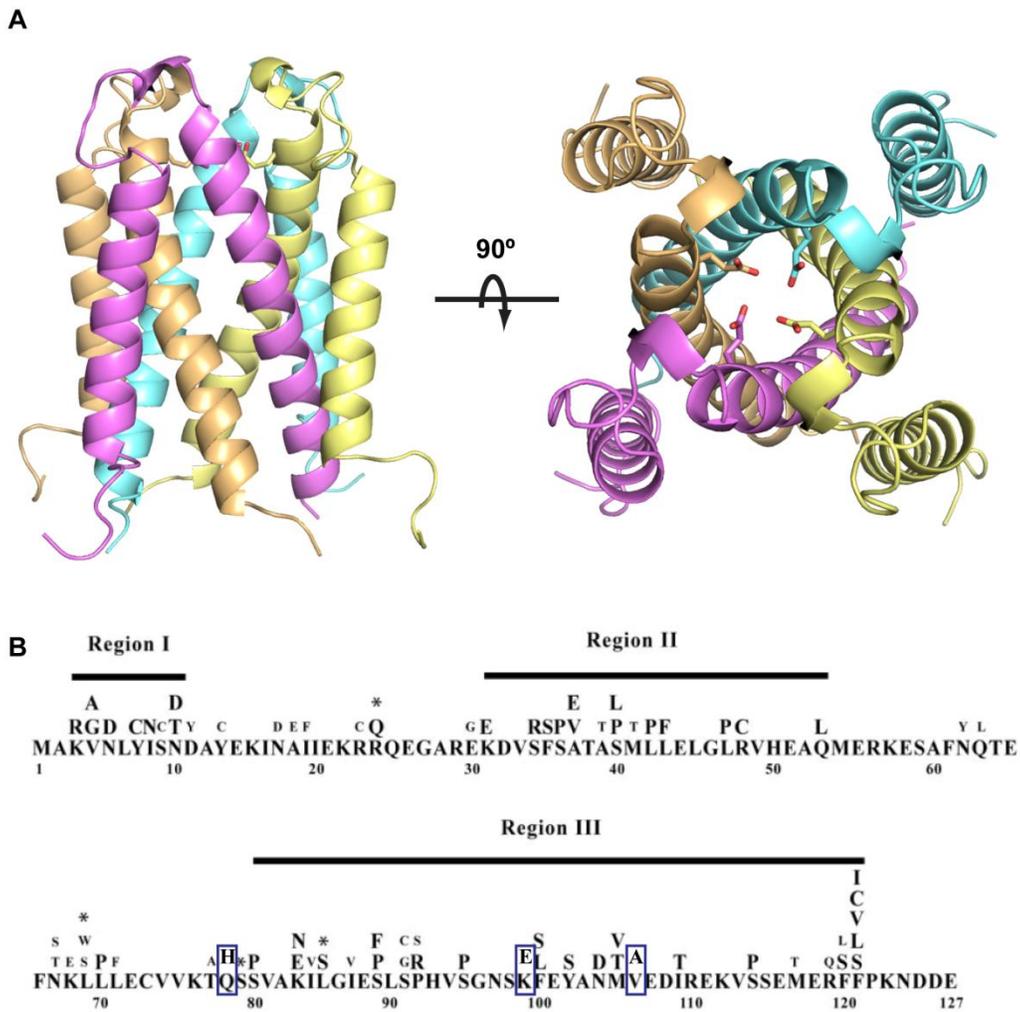


Figure 1-4 TraM

- a) Crystal structure of the TraM C-terminal domain. Glu88 is shown as sticks. (PDB: 2G7O) (Lu et al., 2004)
- b) TraM functional regions as determined by the effect of random mutagenesis. Region I: DNA binding. Region II: dimerization. Region III: tetramerization. (Lu & Frost, 2005) TraD-binding mutants are boxed in blue (Lu & Frost, 2005). Highly defective mutants in conjugation are shown above the TraM sequence in large font, while those that are not are shown in smaller font.



Figure 1-5 Crystal structure of the IHF-DNA complex (PDB: 1IHF) (Rice et al., 1996)

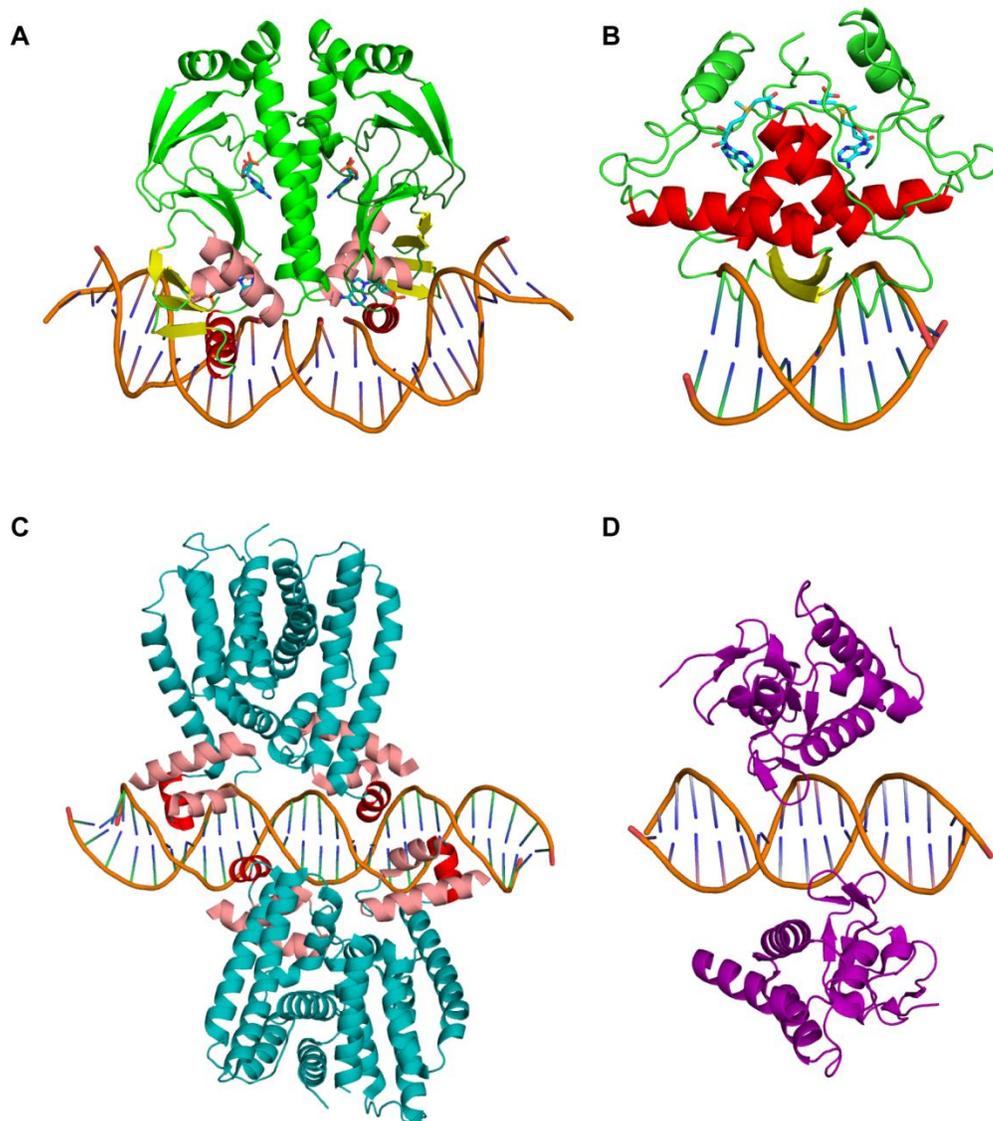


Figure 1-6 Transcription factor structures

- a) Crp crystal structure. (PDB: 2CGP) (Passner & Steitz, 1997)  
Winged-HTH fold is colored in salmon ( $\alpha$ -helices), red (recognition helix), and yellow ( $\beta$ -sheet that makes up the wings). Ligand is shown in cyan sticks.
- b) MetJ-operator crystal structure. (PDB: 1CMB) (Phillips *et al.*, 1989)  
RHH fold is colored in red ( $\alpha$ -helices), and yellow ( $\beta$ -sheet). Ligand is shown in cyan sticks.
- c) Crystal structure of QacR (TetR family) in complex with operator DNA (PDB: 1JT0) (Schumacher *et al.*, 2002)
- d) Crystal structure of Smad4-MH1 (SMAD family) in complex with operator DNA (PDB: 3QSV) (Baburajendran *et al.*, 2011)

## References

- Abo, T. & E. Ohtsubo, (1995) Characterization of the functional sites in the oriT region involved in DNA transfer promoted by sex factor plasmid R100. *J Bacteriol* **177**: 4350-4355.
- Ahmer, B. M., M. Tran & F. Heffron, (1999) The virulence plasmid of *Salmonella typhimurium* is self-transmissible. *J Bacteriol* **181**: 1364-1368.
- Alvarez-Martinez, C. E. & P. J. Christie, (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**: 775-808.
- Aravind, L., V. Anantharaman, S. Balaji, M. M. Babu & L. M. Iyer, (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev* **29**: 231-262.
- Arthur, D. C., R. A. Edwards, S. Tsutakawa, J. A. Tainer, L. S. Frost & J. N. Glover, (2011) Mapping interactions between the RNA chaperone FinO and its RNA targets. *Nucleic Acids Res* **39**: 4450-4463.
- Baburajendran, N., R. Jauch, C. Y. Tan, K. Narasimhan & P. R. Kolatkar, (2011) Structural basis for the cooperative DNA recognition by Smad4 MH1 dimers. *Nucleic Acids Res* **39**: 8213-8222.
- Bell, C. E. & M. Lewis, (2001) Crystallographic analysis of Lac repressor bound to natural operator O1. *J Mol Biol* **312**: 921-926.
- Bennett, P. M., (2008) Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br J Pharmacol* **153 Suppl 1**: S347-357.
- Beranek, A., M. Zettl, K. Lorenzoni, A. Schauer, M. Manhart & G. Koraimann, (2004) Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. *J Bacteriol* **186**: 6999-7006.
- Boer, R., S. Russi, A. Guasch, M. Lucas, A. G. Blanco, R. Perez-Luque, M. Coll & F. de la Cruz, (2006) Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J Mol Biol* **358**: 857-869.
- Boerlin, P. & R. J. Reid-Smith, (2008) Antimicrobial resistance: its emergence and transmission. *Anim Health Res Rev* **9**: 115-126.
- Bowie, J. U. & R. T. Sauer, (1990) TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *J Mol Biol* **211**: 5-6.
- Boyd, E. F. & D. L. Hartl, (1997) Recent horizontal transmission of plasmids between natural populations of *Escherichia coli* and *Salmonella enterica*. *J Bacteriol* **179**: 1622-1627.
- Burrus, V. & M. K. Waldor, (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* **155**: 376-386.
- Byrd, D. R. & S. W. Matson, (1997) Nicking by transesterification: the reaction catalysed by a relaxase. *Mol Microbiol* **25**: 1011-1022.
- Byrd, D. R., J. K. Sampson, H. M. Ragonese & S. W. Matson, (2002) Structure-function analysis of *Escherichia coli* DNA helicase I reveals non-overlapping transesterase and helicase domains. *J Biol Chem* **277**: 42645-42653.
- Cheng, Y., D. E. McNamara, M. J. Miley, R. P. Nash & M. R. Redinbo, (2011) Functional characterization of the multidomain F plasmid Tral relaxase-helicase. *J Biol Chem* **286**: 12670-12682.
- Chu, C. & C. H. Chiu, (2006) Evolution of the virulence plasmids of non-typhoid *Salmonella* and its association with antimicrobial resistance. *Microbes Infect* **8**: 1931-1936.

- Conly, J., (2002) Antimicrobial resistance in Canada. *CMAJ* **167**: 885-891.
- Dash, P. K., B. A. Traxler, M. M. Panicker, D. D. Hackney & E. G. Minkley, Jr., (1992) Biochemical characterization of Escherichia coli DNA helicase I. *Mol Microbiol* **6**: 1163-1172.
- Datta, S., C. Larkin & J. F. Schildbach, (2003) Structural insights into single-stranded DNA binding and cleavage by F factor Tral. *Structure* **11**: 1369-1379.
- de la Cruz, F., L. S. Frost, R. J. Meyer & E. L. Zechner, (2009) Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**: 18-40.
- Di Laurenzio, L., L. S. Frost, B. B. Finlay & W. Paranchych, (1991) Characterization of the oriT region of the IncFV plasmid pED208. *Mol Microbiol* **5**: 1779-1790.
- Disque-Kochem, C. & B. Dreiseikelmann, (1997) The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD in vitro. *J Bacteriol* **179**: 6133-6137.
- Dostal, L. & J. F. Schildbach, (2010) Single-stranded DNA binding by F Tral relaxase and helicase domains is coordinately regulated. *J Bacteriol* **192**: 3620-3628.
- Dostal, L., S. Shao & J. F. Schildbach, (2011) Tracking F plasmid Tral relaxase processing reactions provides insight into F plasmid transfer. *Nucleic Acids Res* **39**: 2658-2670.
- Fekete, R. A. & L. S. Frost, (2000) Mobilization of chimeric oriT plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. *J Bacteriol* **182**: 4022-4027.
- Fekete, R. A. & L. S. Frost, (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.
- Finlay, B. B., L. S. Frost & W. Paranchych, (1986) Nucleotide sequence of the tra YALE region from IncFV plasmid pED208. *J Bacteriol* **168**: 990-998.
- Finnegan, D. J. & N. S. Willetts, (1971) Two classes of Flac mutants insensitive to transfer inhibition by an F-like R factor. *Mol Gen Genet* **111**: 256-264.
- Frost, L., S. Lee, N. Yanchar & W. Paranchych, (1989) finP and fisO mutations in FinP anti-sense RNA suggest a model for FinOP action in the repression of bacterial conjugation by the Flac plasmid JCFL0. *Mol Gen Genet* **218**: 152-160.
- Frost, L. S., K. Ippen-Ihler & R. A. Skurray, (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* **58**: 162-210.
- Frost, L. S. & G. Koraimann, (2010) Regulation of bacterial conjugation: balancing opportunity with adversity. *Future Microbiol* **5**: 1057-1071.
- Frost, L. S. & J. Manchak, (1998) F- phenocopies: characterization of expression of the F transfer region in stationary phase. *Microbiology* **144 ( Pt 9)**: 2579-2587.
- Fu, Y. H., M. M. Tsai, Y. N. Luo & R. C. Deonier, (1991) Deletion analysis of the F plasmid oriT locus. *J Bacteriol* **173**: 1012-1020.
- Garcillan-Barcia, M. P., M. V. Francia & F. de la Cruz, (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* **33**: 657-687.
- Ghigo, J. M., (2001) Natural conjugative plasmids induce bacterial biofilm development. *Nature* **412**: 442-445.
- Gomis-Ruth, F. X., G. Moncalian, R. Perez-Luque, A. Gonzalez, E. Cabezon, F. de la Cruz & M. Coll, (2001) The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* **409**: 637-641.

- Gonzalez-Perez, B., J. D. Carballeira, G. Moncalian & F. de la Cruz, (2009) Changing the recognition site of a conjugative relaxase by rational design. *Biotechnol J* **4**: 554-557.
- Grandoso, G., P. Avila, A. Cayon, M. A. Hernando, M. Llosa & F. de la Cruz, (2000) Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J Mol Biol* **295**: 1163-1172.
- Guasch, A., M. Lucas, G. Moncalian, M. Cabezas, R. Perez-Luque, F. X. Gomis-Ruth, F. de la Cruz & M. Coll, (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Biol* **10**: 1002-1010.
- Gubbins, M. J., W. R. Will & L. S. Frost, (2005) The F-plasmid, a paradigm for bacterial conjugation  
In: *The Dynamic Bacterial Genome*. P. Mullany (ed). Cambridge University Press, pp. 151-206.
- Guedon, E. & J. D. Helmann, (2003) Origins of metal ion selectivity in the DtxR/MntR family of metalloregulators. *Mol Microbiol* **48**: 495-506.
- Guogas, L. M., S. A. Kennedy, J. H. Lee & M. R. Redinbo, (2009) A novel fold in the Tral relaxase-helicase c-terminal domain is essential for conjugative DNA transfer. *J Mol Biol* **386**: 554-568.
- Haft, R. J., E. G. Gachelet, T. Nguyen, L. Toussaint, D. Chivian & B. Traxler, (2007) In vivo oligomerization of the F conjugative coupling protein TraD. *J Bacteriol* **189**: 6626-6634.
- Haft, R. J., G. Palacios, T. Nguyen, M. Mally, E. G. Gachelet, E. L. Zechner & B. Traxler, (2006) General mutagenesis of F plasmid Tral reveals its role in conjugative regulation. *J Bacteriol* **188**: 6346-6353.
- Haran, T. E. & U. Mohanty, (2009) The unique structure of A-tracts and intrinsic DNA bending. *Q Rev Biophys* **42**: 41-81.
- Harley, M. J. & J. F. Schildbach, (2003) Swapping single-stranded DNA sequence specificities of relaxases from conjugative plasmids F and R100. *Proc Natl Acad Sci U S A* **100**: 11243-11248.
- Hochschild, A. & M. Ptashne, (1986) Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. *Cell* **44**: 681-687.
- Howard, M. T., W. C. Nelson & S. W. Matson, (1995) Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem* **270**: 28381-28386.
- Huffman, J. L. & R. G. Brennan, (2002) Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr Opin Struct Biol* **12**: 98-106.
- Inamoto, S., H. Fukuda, T. Abo & E. Ohtsubo, (1994) Site- and strand-specific nicking at oriT of plasmid R100 in a purified system: enhancement of the nicking activity of Tral (helicase I) with TraY and IHF. *J Biochem* **116**: 838-844.
- Johnson, T. J. & L. K. Nolan, (2009) Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol Mol Biol Rev* **73**: 750-774.
- Juhas, M., D. W. Crook & D. W. Hood, (2008) Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol* **10**: 2377-2386.
- Karl, W., M. Bamberger & E. L. Zechner, (2001) Transfer protein TraY of plasmid R1 stimulates Tral-catalyzed oriT cleavage in vivo. *J Bacteriol* **183**: 909-914.
- Kenney, L. J., (2002) Structure/function relationships in OmpR and other winged-helix transcription factors. *Curr Opin Microbiol* **5**: 135-141.
- Khare, D., G. Ziegelin, E. Lanka & U. Heinemann, (2004) Sequence-specific DNA binding determined by contacts outside the helix-turn-helix motif of the ParB homolog KorB. *Nat Struct Mol Biol* **11**: 656-663.

- Kupelwieser, G., M. Schwab, G. Hogenauer, G. Koraimann & E. L. Zechner, (1998) Transfer protein TraM stimulates Tral-catalyzed cleavage of the transfer origin of plasmid R1 in vivo. *J Mol Biol* **275**: 81-94.
- Lang, S., K. Gruber, S. Mihajlovic, R. Arnold, C. J. Gruber, S. Steinlechner, M. A. Jehl, T. Rattei, K. U. Frohlich & E. L. Zechner, (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol Microbiol* **78**: 1539-1555.
- Lang, S., P. C. Kirchberger, C. J. Gruber, A. Redzej, S. Raffl, G. Zellnig, K. Zangger & E. L. Zechner, (2011) An activation domain of plasmid R1 Tral protein delineates stages of gene transfer initiation. *Mol Microbiol* **82**: 1071-1085.
- Larkin, C., S. Datta, M. J. Harley, B. J. Anderson, A. Ebie, V. Hargreaves & J. F. Schildbach, (2005) Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure* **13**: 1533-1544.
- Lau-Wong, I. C., T. Locke, M. J. Ellison, T. L. Raivio & L. S. Frost, (2008) Activation of the Cpx regulon destabilizes the F plasmid transfer activator, TraJ, via the HslVU protease in Escherichia coli. *Mol Microbiol* **67**: 516-527.
- Lawley, T. D., W. A. Klimke, M. J. Gubbins & L. S. Frost, (2003) F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* **224**: 1-15.
- Lawson, C. L., D. Swigon, K. S. Murakami, S. A. Darst, H. M. Berman & R. H. Ebright, (2004) Catabolite activator protein: DNA binding and transcription activation. *Curr Opin Struct Biol* **14**: 10-20.
- Lewis, M., (2005) The lac repressor. *C R Biol* **328**: 521-548.
- Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan & P. Lu, (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**: 1247-1254.
- Llosa, M., F. X. Gomis-Ruth, M. Coll & F. de la Cruz Fd, (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* **45**: 1-8.
- Llosa, M., C. Roy & C. Dehio, (2009) Bacterial type IV secretion systems in human disease. *Mol Microbiol* **73**: 141-151.
- Llosa, M., S. Zunzunegui & F. de la Cruz, (2003) Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc Natl Acad Sci U S A* **100**: 10465-10470.
- Lu, J., A. den Dulk-Ras, P. J. Hooykaas & J. N. Glover, (2009) Agrobacterium tumefaciens VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. *Proc Natl Acad Sci U S A* **106**: 9643-9648.
- Lu, J., R. A. Edwards, J. J. Wong, J. Manchak, P. G. Scott, L. S. Frost & J. N. Glover, (2006) Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. *EMBO J* **25**: 2930-2939.
- Lu, J. & L. S. Frost, (2005) Mutations in the C-terminal region of TraM provide evidence for in vivo TraM-TraD interactions during F-plasmid conjugation. *J Bacteriol* **187**: 4767-4773.
- Lu, J., J. Manchak, W. Klimke, C. Davidson, N. Firth, R. A. Skurray & L. S. Frost, (2002) Analysis and characterization of the IncFV plasmid pED208 transfer region. *Plasmid* **48**: 24-37.

- Lu, J., J. J. Wong, R. A. Edwards, J. Manchak, L. S. Frost & J. N. Glover, (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.
- Lu, J., W. Zhao & L. S. Frost, (2004) Mutational analysis of TraM correlates oligomerization and DNA binding with autoregulation and conjugative DNA transfer. *J Biol Chem* **279**: 55324-55333.
- Lum, P. L. & J. F. Schildbach, (1999) Specific DNA recognition by F Factor TraY involves beta-sheet residues. *J Biol Chem* **274**: 19644-19648.
- Luo, Y., Q. Gao & R. C. Deonier, (1994) Mutational and physical analysis of F plasmid traY protein binding to oriT. *Mol Microbiol* **11**: 459-469.
- Malecka, K. A., W. C. Ho & R. Marmorstein, (2009) Crystal structure of a p53 core tetramer bound to DNA. *Oncogene* **28**: 325-333.
- Maneewannakul, K., P. Kathir, S. Endley, D. Moore, J. Manchak, L. Frost & K. Ippen-Ihler, (1996) Construction of derivatives of the F plasmid pOX-tra715: characterization of traY and traD mutants that can be complemented in trans. *Mol Microbiol* **22**: 197-205.
- Mihajlovic, S., S. Lang, M. V. Sut, H. Strohmaier, C. J. Gruber, G. Koraimann, E. Cabezón, G. Moncalian, F. de la Cruz & E. L. Zechner, (2009) Plasmid r1 conjugative DNA processing is regulated at the coupling protein interface. *J Bacteriol* **191**: 6877-6887.
- Miller, D. L. & J. F. Schildbach, (2003) Evidence for a monomeric intermediate in the reversible unfolding of F factor TraM. *J Biol Chem* **278**: 10400-10407.
- Moncalian, G. & F. de la Cruz, (2004) DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* **1701**: 15-23.
- Moncalian, G., G. Grandoso, M. Llosa & F. de la Cruz, (1997) oriT-processing and regulatory roles of TrwA protein in plasmid R388 conjugation. *J Mol Biol* **270**: 188-200.
- Monzinger, A. F., A. Ozburn, S. Xia, R. J. Meyer & J. D. Robertus, (2007) The structure of the minimal relaxase domain of MobA at 2.1 Å resolution. *J Mol Biol* **366**: 165-178.
- Mulec, J., M. Starcic & D. Zgur-Bertok, (2002) F-like plasmid sequences in enteric bacteria of diverse origin, with implication of horizontal transfer and plasmid host range. *Curr Microbiol* **44**: 231-235.
- Narra, H. P. & H. Ochman, (2006) Of what use is sex to bacteria? *Curr Biol* **16**: R705-710.
- Nash, R. P., S. Habibi, Y. Cheng, S. A. Lujan & M. R. Redinbo, (2010) The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res* **38**: 5929-5943.
- Nelson, B. L. & D. M. Cox, (2004) *Lehninger Principles of Biochemistry*. W. H. Freeman.
- Nelson, W. C., M. T. Howard, J. A. Sherman & S. W. Matson, (1995) The traY gene product and integration host factor stimulate Escherichia coli DNA helicase I-catalyzed nicking at the F plasmid oriT. *J Biol Chem* **270**: 28374-28380.
- Nelson, W. C., B. S. Morton, E. E. Lahue & S. W. Matson, (1993) Characterization of the Escherichia coli F factor traY gene product and its binding sites. *J Bacteriol* **175**: 2221-2228.
- Ni, L., S. O. Jensen, N. Ky Tonthat, T. Berg, S. M. Kwong, F. H. Guan, M. H. Brown, R. A. Skurray, N. Firth & M. A. Schumacher, (2009) The Staphylococcus aureus pSK41 plasmid-encoded ArtA protein is a master

- regulator of plasmid transmission genes and contains a RHH motif used in alternate DNA-binding modes. *Nucleic Acids Res* **37**: 6970-6983.
- Osman, D. & J. S. Cavet, (2010) Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Nat Prod Rep* **27**: 668-680.
- Ostermann, E., F. Krichek & G. Hogenauer, (1984) Cloning the origin of transfer region of the resistance plasmid R1. *EMBO J* **3**: 1731-1735.
- Palmer, K. L., V. N. Kos & M. S. Gilmore, (2010) Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Curr Opin Microbiol* **13**: 632-639.
- Passner, J. M. & T. A. Steitz, (1997) The structure of a CAP-DNA complex having two cAMP molecules bound to each monomer. *Proc Natl Acad Sci U S A* **94**: 2843-2847.
- Penfold, S. S., J. Simon & L. S. Frost, (1996) Regulation of the expression of the traM gene of the F sex factor of Escherichia coli. *Mol Microbiol* **20**: 549-558.
- Phillips, S. E., I. Manfield, I. Parsons, B. E. Davidson, J. B. Rafferty, W. S. Somers, D. Margarita, G. N. Cohen, I. Saint-Girons & P. G. Stockley, (1989) Cooperative tandem binding of met repressor of Escherichia coli. *Nature* **341**: 711-715.
- Porwollik, S. & M. McClelland, (2003) Lateral gene transfer in Salmonella. *Microbes Infect* **5**: 977-989.
- Potron, A., L. Poirel & P. Nordmann, (2011) Plasmid-mediated transfer of the bla(NDM-1) gene in Gram-negative rods. *FEMS Microbiol Lett* **324**: 111-116.
- Ragonese, H., D. Haisch, E. Villareal, J. H. Choi & S. W. Matson, (2007) The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at oriT through an interaction with Tral. *Mol Microbiol* **63**: 1173-1184.
- Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan & R. Tobes, (2005) The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* **69**: 326-356.
- Raumann, B. E., M. A. Rould, C. O. Pabo & R. T. Sauer, (1994) DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature* **367**: 754-757.
- Rice, P. A., S. Yang, K. Mizuuchi & H. A. Nash, (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**: 1295-1306.
- Schleif, R., (2010) AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action. *FEMS Microbiol Rev* **34**: 779-796.
- Schreiter, E. R. & C. L. Drennan, (2007) Ribbon-helix-helix transcription factors: variations on a theme. *Nat Rev Microbiol* **5**: 710-720.
- Schreiter, E. R., S. C. Wang, D. B. Zamble & C. L. Drennan, (2006) NikR-operator complex structure and the mechanism of repressor activation by metal ions. *Proc Natl Acad Sci U S A* **103**: 13676-13681.
- Schroder, G., S. Krause, E. L. Zechner, B. Traxler, H. J. Yeo, R. Lurz, G. Waksman & E. Lanka, (2002) TraG-like proteins of DNA transfer systems and of the Helicobacter pylori type IV secretion system: inner membrane gate for exported substrates? *J Bacteriol* **184**: 2767-2779.
- Schroder, G. & E. Lanka, (2005) The mating pair formation system of conjugative plasmids-A versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54**: 1-25.
- Schumacher, M. A. & B. E. Funnell, (2005) Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* **438**: 516-519.

- Schumacher, M. A., A. Mansoor & B. E. Funnell, (2007) Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *J Biol Chem* **282**: 10456-10464.
- Schumacher, M. A., M. C. Miller, S. Grkovic, M. H. Brown, R. A. Skurray & R. G. Brennan, (2002) Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J* **21**: 1210-1218.
- Schumacher, M. A., K. M. Piro & W. Xu, (2010) Insight into F plasmid DNA segregation revealed by structures of SopB and SopB-DNA complexes. *Nucleic Acids Res* **38**: 4514-4526.
- Schwab, M., H. Reizenzein & G. Hogenauer, (1993) TraM of plasmid R1 regulates its own expression. *Mol Microbiol* **7**: 795-803.
- Senear, D. F. & M. Brenowitz, (1991) Determination of binding constants for cooperative site-specific protein-DNA interactions using the gel mobility-shift assay. *J Biol Chem* **266**: 13661-13671.
- Silverman, P. M. & A. Sholl, (1996) Effect of traY amber mutations on F-plasmid traY promoter activity in vivo. *J Bacteriol* **178**: 5787-5789.
- Skippington, E. & M. A. Ragan, (2011) Lateral genetic transfer and the construction of genetic exchange communities. *FEMS Microbiol Rev* **35**: 707-735.
- Smith, T. L. & R. T. Sauer, (1995) P22 Arc repressor: role of cooperativity in repression and binding to operators with altered half-site spacing. *J Mol Biol* **249**: 729-742.
- Somers, W. S. & S. E. Phillips, (1992) Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by beta-strands. *Nature* **359**: 387-393.
- Stayrook, S., P. Jaru-Ampornpan, J. Ni, A. Hochschild & M. Lewis, (2008) Crystal structure of the lambda repressor and a model for pairwise cooperative operator binding. *Nature* **452**: 1022-1025.
- Stern, J. C. & J. F. Schildbach, (2001) DNA recognition by F factor Tral36: highly sequence-specific binding of single-stranded DNA. *Biochemistry* **40**: 11586-11595.
- Stockner, T., C. Plugariu, G. Koraimann, G. Hogenauer, W. Bermel, S. Prytulla & H. Sterk, (2001) Solution structure of the DNA-binding domain of TraM. *Biochemistry* **40**: 3370-3377.
- Strahilevitz, J., G. A. Jacoby, D. C. Hooper & A. Robicsek, (2009) Plasmid-mediated quinolone resistance: a multifaceted threat. *Clin Microbiol Rev* **22**: 664-689.
- Su, L. H., C. Chu, A. Cloeckert & C. H. Chiu, (2008) An epidemic of plasmids? Dissemination of extended-spectrum cephalosporinases among *Salmonella* and other Enterobacteriaceae. *FEMS Immunol Med Microbiol* **52**: 155-168.
- Sut, M. V., S. Mihajlovic, S. Lang, C. J. Gruber & E. L. Zechner, (2009) Protein and DNA effectors control the Tral conjugative helicase of plasmid R1. *J Bacteriol* **191**: 6888-6899.
- Szipirer, C. Y., M. Faelen & M. Couturier, (2000) Interaction between the RP4 coupling protein TraG and the pBHR1 mobilization protein Mob. *Mol Microbiol* **37**: 1283-1292.
- Taki, K., T. Abo & E. Ohtsubo, (1998) Regulatory mechanisms in expression of the traY-I operon of sex factor plasmid R100: involvement of traJ and traY gene products. *Genes Cells* **3**: 331-345.
- Tato, I., I. Matilla, I. Arechaga, S. Zunzunegui, F. de la Cruz & E. Cabezon, (2007) The ATPase activity of the DNA transporter TrwB is modulated by

- protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J Biol Chem* **282**: 25569-25576.
- Terradot, L. & G. Waksman, (2011) Architecture of the Helicobacter pylori Cag-type IV secretion system. *FEBS J* **278**: 1213-1222.
- Thomas, J. & D. W. Hecht, (2007) Interaction of Bacteroides fragilis pLV22a relaxase and transfer DNA with Escherichia coli RP4-TraG coupling protein. *Mol Microbiol* **66**: 948-960.
- Treangen, T. J. & E. P. Rocha, (2011) Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet* **7**: e1001284.
- van Biesen, T. & L. S. Frost, (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, traJ mRNA, and promotes duplex formation. *Mol Microbiol* **14**: 427-436.
- Varsaki, A., G. Moncalian, P. Garcillan-Barcia Mdel, C. Drainas & F. de la Cruz, (2009) Analysis of ColE1 MbeC unveils an extended ribbon-helix-helix family of nicking accessory proteins. *J Bacteriol* **191**: 1446-1455.
- Verdino, P., W. Keller, H. Strohmaier, K. Bischof, H. Lindner & G. Koraimann, (1999) The essential transfer protein TraM binds to DNA as a tetramer. *J Biol Chem* **274**: 37421-37428.
- Watanabe, T., (1963) Infective heredity of multiple drug resistance in bacteria. *Bacteriol Rev* **27**: 87-115.
- Weissbach, H. & N. Brot, (1991) Regulation of methionine synthesis in Escherichia coli. *Mol Microbiol* **5**: 1593-1597.
- Will, W. R., J. Lu & L. S. Frost, (2004) The role of H-NS in silencing F transfer gene expression during entry into stationary phase. *Mol Microbiol* **54**: 769-782.
- Williams, S. L. & J. F. Schildbach, (2007) TraY and integration host factor oriT binding sites and F conjugal transfer: sequence variations, but not altered spacing, are tolerated. *J Bacteriol* **189**: 3813-3823.
- Wozniak, R. A. & M. K. Waldor, (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat Rev Microbiol* **8**: 552-563.
- Yoshida, H., N. Furuya, Y. J. Lin, P. Guntert, T. Komano & M. Kainosho, (2008) Structural basis of the role of the NikA ribbon-helix-helix domain in initiating bacterial conjugation. *J Mol Biol* **384**: 690-701.
- Youderian, P. & D. N. Arvidson, (1994) Direct recognition of the trp operator by the trp holorepressor--a review. *Gene* **150**: 1-8.

## Chapter 2

### The crystal structure of the TraD C-terminus in complex with the TraM tetramerization domain<sup>1</sup>

#### Overview

The interaction of TraD with TraM is thought to be a key step in the initiation of conjugative DNA transfer. The protein residues involved had been narrowed down to the last 38 amino acids of TraD and residues in the C-terminal domain of TraM. Therefore, crystallization of the TraM C-terminal domain with a proteolytic fragment of TraD consisting of the last 73 amino acids of TraD was attempted. The resulting structure showed the last 7 residues of TraD forming a  $\beta$ -hairpin which was bound to each of the 4 faces of TraM. The presence of only a short peptide fragment of TraD was confirmed with mass spectrometry. With the crystal structure, TraM residues previously identified to be important in TraD interaction were shown to be part of the TraD binding pocket. The nature of the TraM-TraD interaction is largely electrostatic, with an acidic tail binding to a basic pocket through mostly long-range electrostatic interactions. Lys99 makes a long-range electrostatic interaction with the side chain of TraD Asp715 and the main chain of TraD. Other key components of this interaction are the electrostatic interaction between the C-terminal carboxylate and basic TraM residues, Arg 110 and Lys76, and binding of the phenyl side chain of TraD Phe717 to a hydrophobic binding pocket in TraM. Sequence alignment of TraD and TraM homologues from pKPN3 and pED208 shows that the residues involved in TraD-TraM interaction are not entirely conserved, meaning that there is likely specificity in TraM-TraD interactions between the different plasmids.

---

<sup>1</sup> Part of this work was previously published: Lu J, Wong JJ, Edwards RA, Manchak J, Frost LS & Glover JN (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.

## Introduction

Recruitment of the plasmid to the coupling protein remains one of the least well understood aspects of conjugation. This process is thought to involve the constituent proteins of the relaxosome, which, in addition to Tral, contains TraM, TraY and IHF that occupy binding sites within *oriT* (Kupelwieser, *et al.*, 1998, Ragonese, *et al.*, 2007). In the F plasmid, there appears to be two stages in the recruitment process. The first involves a highly specific interaction between TraD and TraM (Disque-Kochem & Dreiseikelmann, 1997, Beranek, *et al.*, 2004) that is responsible for mediating specificity in relaxosome-transferosome interactions between plasmids. The second stage involves further interactions between more extensive regions of TraM and TraD (Sastre, *et al.*, 1998), presumably in preparation for the melting and unwinding of the DNA (Csitkovits, *et al.*, 2004). TraD has been shown to be critical for plasmid transfer, but not for pilus expression and assembly, consistent with the hypothesis that TraD-TraM interaction is responsible for the recruitment of the plasmid to the conjugative pore (Beranek, *et al.*, 2004). TraD interaction with Tral via the first helicase domain is also a crucial part of the transfer process (Lang, *et al.*, 2010, Lang, *et al.*, 2011), though it cannot substitute for the TraD-TraM interaction.

TraD of F-like plasmids contains a C-terminal extension (residues 576-717 in F) not found in TrwB, its homologue in the R388 plasmid, which is required for the efficient transfer of F (Sastre, *et al.*, 1998). This C-terminal region is responsible for increased specificity and efficiency of F TraD interaction with its cognate relaxosome at the expense of limiting the range of plasmids it can mobilize. Genetic and biochemical studies have shown that the C-terminal extension of TraD is required for its interaction with its cognate TraM, and that

the minimal fragment required is the last 38 amino acids of TraD (Sastre, *et al.*, 1998, Beranek, *et al.*, 2004). However, other regions of TraD are likely involved in TraM interaction, as full binding affinity of TraD to TraM is only attained with the full length cytoplasmic domain and not its truncation mutant without the last 38 residues (Beranek, *et al.*, 2004). TraM was shown to interact with TraD via its C-terminal tetramerization domain, as shown by the discovery of a single mutation, K99E, that abrogates TraM-TraD interaction without affecting autoregulation or tetramerization (Lu & Frost, 2005). Two additional mutations in TraM, V106A and Q78H, were discovered that had similar effects (Lu & Frost, 2005). However, it was not clear how TraD recognizes TraM in a plasmid-specific manner to recruit the relaxosome complex to the conjugation machinery.

The mechanism of this interaction was revealed at the atomic level by the crystal structure of the TraM C-terminal domain with the last 7 amino acids of TraD. The structure shows specific interactions between side chains of TraD and TraM that involve hydrophobic stacking, as well as electrostatic complementarity that hinges on the availability of the free carboxylate at the C-terminus of the TraD polypeptide chain. The TraD fragment observed in the crystal structure was much smaller than the original fragment, TraD<sup>645-717</sup>, due to proteolysis during crystallization.

## Results

### *Crystallization Screening and Optimization of TraM<sup>58-127</sup> - TraD<sup>645-717</sup> complexes*

Previous studies suggested that the C-terminal 38 residue region of TraD is involved with binding to TraM (Sastre, *et al.*, 1998, Beranek, *et al.*, 2004). Previous mutagenesis studies found three residues in the C-terminal domain of

TraM likely to be involved in TraD binding, which was confirmed with the most important residue, K99 (Lu & Frost, 2005). Therefore, the same C-terminal domain construct of TraM as that used for solving the C-terminal domain crystal structure, TraM<sup>58-127</sup> (Lu, *et al.*, 2006) was chosen for co-crystallization with a C-terminal TraD fragment. The TraD construct used was TraD<sup>645-717</sup>, a relatively stable fragment from trypsin proteolysis (Lu, J., unpublished findings).

TraM<sup>58-127</sup> and TraD<sup>645-717</sup> mixtures were used in manual crystallization screening setups using commercial and homemade screens. One hit was obtained in Nextal Classics Condition #91: 30% PEG 4000, 100 mM Tris pH 8.5, and 200 mM sodium acetate at 25°C. This condition was optimized to 30% PEG 2000, while keeping all other variables the same. Crystals grew after 2 weeks. Other conditions varied in attempts to improve the size and quality of the crystals were temperature (4°, 15°), pH of Tris (7.5, 8.0, 8.5, 9.0), different cations (lithium, potassium, magnesium, calcium, or ammonium acetate) at various concentrations 50 mM - 200 mM), different anions (sodium formate, nitrate, or chloride) at various concentrations (50 mM – 400 mM). In addition, 1 mM DTT was tested as an additive, microseeding, varying protein (15-25 mg/ml), buffer concentration (0-100 mM) TraD:TraM ratio (1.75-1.25:1), and drop ratio (2:1 to 1:2 protein:precipitant) was tested. However, the only variable that resulted in fewer and larger crystals per drop was changing the average length of PEG to 2000, which resulted in crystals of approximately 0.2-0.3 µm in size (Fig 2-1A). A dataset was collected from one of these crystals which diffracted to 2.55 Å and indexed to the space group R3 with cell dimensions of a=b=142.25, c=70.95, and  $\alpha=\beta=90^\circ$ ,  $\gamma=120^\circ$  (Table 2-1).

### *Mass Spectrometry of TraM<sup>58-127</sup>-TraD<sup>645-717</sup> Crystals*

MALDI-TOF mass spectrometry of the TraM<sup>58-127</sup>-TraD<sup>645-717</sup> crystal content indicated that the crystal contained peptides with molecular weights of 1080.1 and 1365.2, corresponding to the tryptic fragments that contain C-terminal 10 and 12 residues of TraD, respectively. (Figure 2-1C and D), The proteolysis of TraD<sup>645-717</sup> could be due to trace amount of trypsin present in purified TraM<sup>58-127</sup> because trypsin was used to cleave off the His<sub>6</sub>-tag from the trypsin-resistant TraM<sup>58-127</sup> during purification (Lu, *et al.*, 2006). Being relatively unstructured as determined by circular dichroism (data not shown), and susceptible to trypsin proteolysis even in the presence of TraM<sup>58-127</sup> (data not shown), even trace amounts of trypsin would result in the degradation of TraD<sup>645-717</sup>.

### *Crystallization of TraM<sup>58-127</sup>-TraD<sup>707-717</sup> or TraM<sup>58-127</sup>-TraD<sup>705-717</sup>*

Upon confirmation that the crystals resulting from the TraM<sup>58-127</sup>-TraD<sup>645-717</sup> mixture contained only the last 10 or 12 amino acids of TraD, synthetic peptides of TraD<sup>707-717</sup> and TraD<sup>705-717</sup> used in co-crystallization screens with TraM<sup>58-127</sup>. Crystals were obtained in the same conditions as the optimized TraM<sup>58-127</sup>-TraD<sup>645-717</sup> crystals, but none of them diffracted to as high resolution. In some drops of TraD<sup>707-717</sup> or TraD<sup>705-717</sup> and TraM<sup>58-127</sup>, bar-shaped, thick crystals ~0.3-0.4 μm long appeared (Fig 2-1B). These crystals often diffracted significantly better than the R3 crystals even on a home x-ray source. A dataset was collected to 1.33 Å at the Advanced Light Source (Table 2-1). Not encouragingly, the crystal indexed to the space group I4 with cell dimensions of a=b=51.45 and c=49.75, and α=β=γ=90°, which is effectively the same as that of the apo-TraM<sup>58-127</sup> crystals (I4, a=b=51.73, c=49.58, and α=β=γ=90°). Inspection of the electron

density map after molecular replacement and one round of rigid body refinement showed that no TraD peptide was present in the crystal. The TraD binding pocket of the TraM tetramer is instead occupied by water molecules and the C-terminal tail of TraM (Figure 2-2).

*Structure Solution of TraM -TraD complexes resulting from crystallization of TraM<sup>58-127</sup> -TraD<sup>645-717</sup>*

The complex structure was determined by molecular replacement using the previously determined TraM<sup>58-127</sup> tetramer as a search model, followed by manual building of TraD peptide residues into the remaining protein electron density. Following solution of the two TraM tetramers in the asymmetric unit, there was no space in the asymmetric unit for the full-length TraD<sup>645-717</sup> fragment. Mass spectrometry confirmed that only the C-terminal peptides were present, corresponding to the last 10 or 12 amino acids. Each TraM tetramer was bound by four TraD peptides, arranged in a 4-fold symmetric fashion around the outside of the tetramer (Figure 2-3A). One TraD chain contained the last 8 amino acids of TraD, and each of the other 3 chains contained the last 7 residues of TraD.

The electron density of the TraD peptides was validated by prime-and-switch phasing using the final refined structure as a starting model (Terwilliger, 2004). The resulting phases have a figure of merit of 0.56 and bias ratio of 1.01, indicating the elimination of the model bias from the prime-and-switch map (Figure 2-3B). The TraD C-terminal peptide binds to a pocket on each of the 4 symmetry-related faces of the TraM tetramer (Figure 2-3A). The peptide adopts a type II  $\beta$ -turn with Pro and Gly residues found at i+1 and i+2 positions. The  $\beta$ -turn structure is further stabilized by hydrogen bond interaction between the side

chain of Asp715 and the mainchain of Glu712. The conformation of each of the 8 independently determined TraD peptides is nearly identical, with an RMSD for C $\alpha$  atoms of ~0.3 Å (Figure 2-3C).

#### *Detailed TraD-TraM interactions*

The TraD peptide binds into a pocket composed of residues from three of the four TraM protomers (Figure 2-3D). The centre of the pocket is hydrophobic (including TraM residues Leu85, Val106, and Ile109) and binds the side chain of the C-terminal residue of TraD, Phe717. Further hydrophobic interactions are made by Val711 and Pro713, which largely pack against Tyr102 of TraM. Surrounding the hydrophobic pocket are a number of positively charged residues in TraM – Lys76, Lys83, Lys99, Lys112, and Arg110, which interact with an equal number of carboxylate groups in TraD. Most striking is the direct recognition of the C-terminal carboxylate of TraD by both the guanidinium group of Arg110 and the amino group of Lys76, both of which are within hydrogen bonding distance to the carboxylate. Acidic TraD residues Glu712, Asp715 and Asp716 also make electrostatic interactions with basic TraM residues, although none of these interactions are close enough (< 3.5 Å) to be considered true hydrogen bond/salt bridging interactions (Figure 2-3E).

#### *Conservation of TraM and TraD binding determinants*

Comparisons of TraM and TraD from other F-like plasmids suggest the TraD-TraM interactions observed in F are likely conserved in other members of the F plasmid family. We analyzed the sequences of TraM and TraD from two plasmids that have diverged significantly from F: pKPN3 isolated from *K. pneumoniae*, and pED208 from *S. typhimurium* (Figure 2-4). The residues

important for TraD-TraM interactions in the C-terminal tail of TraD are largely conserved. The glycine residue which is important for the  $\beta$ -turn is conserved in all three sequences, as is an aromatic residue (Phe or Tyr) at the end of the chain. In addition, each of the tails is overall highly acidic. In TraM, the positively charged residues that surround the TraD binding site are less well conserved, except at position 76, which is either a lysine (F and pKPN3) or arginine (pED208). This residue makes the closest approach to the TraD C-terminal carboxylate in the crystal structure. Residue 102 is a conserved planar aromatic (Tyr or Trp). In the TraD-TraM structure, this residue forms extensive hydrophobic contacts with the TraD  $\beta$ -turn. Thus, while the similarities of TraD and TraM in these plasmids suggest a similar mode of interaction, the differences are significant and may provide a degree of plasmid specificity.

#### *Comparison with the structure of unliganded TraM*

Overall, the structure of the TraM tetramer is nearly identical to that observed in free TraM determined in (Lu, *et al.*, 2006), with a RMSD based on all  $C_{\alpha}$  atoms of 0.65 Å. In particular, the TraD binding pocket is almost unchanged between the free and bound forms, with the exception of the side chains of Lys 76 and Arg110, which rotate in the complex with TraD to interact with the TraD C-terminal carboxylate (Figure 2-5). The free TraM tetramerization domain was previously shown to be metastable (Miller & Schildbach, 2003), largely due to the packing of four symmetry-related, protonated glutamic acid side chains (Glu 88) at the centre of the TraM helical bundle (Lu, *et al.*, 2006). These residues are packed in a similar manner within the TraD-TraM complex, and are within hydrogen bonding distance of one another, likely in a protonated state. As Glu88

and the TraD binding pocket are at the same end of the TraM tetramer, deprotonation-induced destabilization of the TraM tetramer will certainly affect the neighboring TraD-binding pocket as previously suggested (Lu & Frost, 2005, Lu, *et al.*, 2006)

## Discussion

Relaxosome-transferosome interaction, through which DNA is brought to the membrane-bound type IV secretion system for conjugative transfer, is a key step during bacterial conjugation. This interaction provides a potential pathway for the elusive mating signal, originating from donor-recipient contact, to reach the cytosolic relaxosome. In the F-plasmid-mediated conjugation system, a relaxosome component, TraM, specifically interacts with the coupling protein TraD of the type IV secretion machinery (Disque-Kochem & Dreiseikelmann, 1997). The K99E mutation that specifically reduces the binding affinity of TraM for TraD also decreases the efficiency of F conjugation (Lu & Frost, 2005), suggesting that the TraD-TraM interaction is important for F conjugation. Previous results have shown that the C-terminal 38 residues of TraD are necessary and sufficient to bind TraM (Beranek, *et al.*, 2004) although not as efficiently as intact TraD. In this work, we presented the crystal structure of the C-terminal 7-8 residues of TraD bound to the C-terminal tetramerization domain of TraM (TraM<sup>58-127</sup>), which is the first structural evidence of specific interactions between a relaxosome and its cognate transferosome in a bacterial conjugation system.

Many T4SSs in gram-negative bacteria have a membrane-bound SpoIIIE/FtsK-like ATPase that recognizes an unstructured C-terminal signal

sequence of a type IV secretion substrate to direct substrate translocation. For example, *A. tumefaciens* VirF has a C-terminal unstructured tail containing several arginine residues required for translocation (Vergunst, *et al.*, 2005). RalF of the *L. pneumophila* Dot/Icm system has an unstructured C-terminal signal sequence with a Leu residue at the –3 position critical for translocation (Nagai, *et al.*, 2005). A substrate of T4SS-like ESX-1/Snm secretion system in *Mycobacterium tuberculosis*, CFP-10, has an unstructured tail with a C-terminal phenylalanine residue essential for interacting with a SpoIIIE/FtsK-like ATPase Rv3871 during translocation (Renshaw, *et al.*, 2005, Champion, *et al.*, 2006). Although in the F system it is the TraD ATPase which bears the flexible tail that is recognized by the substrate TraM-DNA complex, it appears that a common feature in T4SS is the recognition of a flexible C-terminal peptide tail for substrate recruitment and translocation. The crystal structure of the TraM<sup>58-127</sup>-TraD<sup>711-717</sup> complex shows for the first time how this recognition occurs in a plasmid-specific manner, explaining why conjugative DNA transfer typically only occurs when transferosome and relaxosome components are from the same plasmid (Sastre, *et al.*, 1998).

## Materials and Methods

### *Cloning, expression, and purification of TraM<sup>58-127</sup> and TraD<sup>645-717</sup>*

Cloning, expression, and purification of TraM<sup>58-127</sup> was carried out as described in (Lu, *et al.*, 2006), while that of TraD<sup>645-717</sup> was carried out as described in (Lu, *et al.*, 2008). All the plasmids and oligonucleotides used in this work are listed in Supplementary Table 1. Plasmid pJLTraD<sup>645-717</sup>, which encodes the C-terminal 73 residues (645 to 717) of TraD, was constructed by ligating the

*EcoRI-BamHI* fragment of pT7-7 to the *EcoRI-BamHI* fragment of DNA amplified from pOX38-Km using JLU261 and JLU207 as primers.

#### *Crystallization of TraM<sup>58-127</sup>-TraD<sup>645-717</sup> complex and TraM<sup>58-127</sup> apo-protein*

TraD<sup>645-717</sup> and TraM<sup>58-127</sup> were concentrated to 15 mg/mL in 50 mM Tris, pH 7.5; and were mixed together in a 1.25 to 1.5:1 ratio. Crystals were obtained using the hanging drop vapor diffusion technique with 1  $\mu$ L of protein mixture mixed with 1  $\mu$ L well solution containing 30% PEG 2000, 100 mM Tris pH 8.5, and 200 mM sodium acetate at 25 °C for 2 weeks. Crystals were soaked in well solution plus 20% glycerol (v/v) for 20 min prior to flash freezing in liquid nitrogen. A native data set was collected to 2.55 Å from a TraM<sup>58-127</sup>-TraD<sup>645-717</sup> co-crystal at beamline 8.3.1 at Advanced Light Source, Lawrence Berkeley National Laboratory. 15 mg/mL TraM<sup>58-127</sup> in 50 mM Tris, pH 7.5 was mixed with TraD<sup>707-717</sup> or TraD<sup>705-717</sup> (synthesized by the Alberta Peptide Institute) in a 4 to 1.25:1 ratio in an attempt to obtain co-crystals. A native data set was collected to 1.33 Å from a crystal grown from a TraM<sup>58-127</sup>-TraD<sup>707-717</sup> mixture in the same conditions as the complex, which was later determined to be the TraM<sup>58-127</sup> apo-protein, also at Beamline 8.3.1. Data collection statistics are described in Table 1.

#### *Mass Spectrometry of TraM<sup>58-127</sup>-TraD<sup>645-717</sup> crystals*

Crystals grown from TraM<sup>58-127</sup>-TraD<sup>645-717</sup> mixtures were washed in mother liquor and dissolved in 50 mM ammonium acetate, pH 7.5 to prepare them for analysis. MALDI-TOF mass spectrometry was done on the sample, and peptide mass fragments were identified using Protein Prospector MS-Fit (Baker, P.R. and Clauser, K.R. <http://prospector.ucsf.edu>).

### *Structure determination of the TraM<sup>58-127</sup>–TraD<sup>645-717</sup> complex*

The TraM<sup>58-127</sup>–TraD<sup>645-717</sup> structure was solved by molecular replacement using the TraM<sup>58-127</sup> structure (Lu, *et al.*, 2006) as the search model in MOLREP (Vagin & Teplyakov, 2000). Amino acids corresponding to the C-terminal residues of TraD were manually built into electron density using Coot (Emsley & Cowtan, 2004). Manual model building and iterative cycles of refinement in REFMAC (Murshudov, *et al.*, 1997) with 8-fold NCS restraints were used to complete and refine the model (Table 1). 96.3% of residues fall within the most favored regions of the Ramachandran plot with the remaining in the additionally allowed regions. The molecular structure figures were prepared by using PyMOL (DeLano Scientific, <http://www.pymol.org>).

### *Structure determination of the TraM<sup>58-127</sup> apo-structure complex*

The TraM<sup>58-127</sup> apo-structure was solved by molecular replacement using the TraM<sup>58-127</sup> structure (Lu, *et al.*, 2006) as the search model in MOLREP (Vagin & Teplyakov, 2000). Waters were built using ARP/wARP Solvent (Langer, *et al.*, 2008) and the co-factors Na<sup>+</sup> and acetate were built manually. REFMAC (Murshudov, *et al.*, 1997) was used for iterative cycles of refinement.

<b>Table 2-1 Data collection and refinement of TraM<sup>58-127</sup>-TraD tail crystals</b>		
	<b>TraM<sup>58-127</sup>-TraD<sup>645-717</sup> complex structure</b>	<b>TraM<sup>58-127</sup> apo- structure</b>
<b>Data Collection</b>		
Space Group	R3	I4
<b>Cell Dimensions</b>		
a, b, c (Å)	142.25, 142.25, 70.95	51.45, 51.45, 49.75
α, β, γ (°)	90, 90, 120	90, 90, 90
Wavelength (Å)	1.1159	1.1159
Resolution (Å)	2.55	1.33
R <sub>sym</sub> <sup>b</sup>	0.061 (0.387)	0.055 (0.095)
I/σI	22.7 (4.0)	23.5 (7.2)
Redundancy	5.2 (4.2)	3.5 (1.3)
Completeness (%)	99.95 (100.00)	90.01 (41.10)
<b>Refinement</b>		
Resolution (Å)	46.5 - 2.55	36.39 - 1.33
Number of unique reflections	16609	12811
R <sub>work</sub> <sup>c</sup> / R <sub>free</sub> <sup>d</sup> (%)	21.8 / 25.7	15.5 / 20.9 <sup>e</sup>
Number of protein atoms in asymmetric unit	4456	547
B-factor (overall)	49.8	16.4
Bond angle r.m.s.d. (°)	1.06	1.35
Bond length r.m.s.d. (Å)	0.011	0.009
<sup>a</sup> Data of the highest resolution shell (2.55-2.59Å for TraM <sup>58-127</sup> -TraD TraD <sup>645-717</sup> complex, 1.33-1.35 Å for TraM <sup>58-127</sup> apo-structure) are shown in parentheses		
<sup>b</sup> R <sub>sym</sub> = $\sum_{hkl} \sum_i  I_i(hkl) - \langle I(hkl) \rangle  / \sum_{hkl} \sum_i I_i(hkl)$ , where I <sub>i</sub> (hkl) is the intensity for an observation of a reflection and <I(hkl)> is the average intensity of all symmetry-related observations of a reflection		
<sup>c</sup> R <sub>work</sub> = $\sum_{hkl}    F_{obs}  -  F_{calc}    / \sum_{hkl}  F_{obs} $		
<sup>d</sup> R <sub>free</sub> = R <sub>work</sub> calculated for 5% of reflections excluded from refinement		
<sup>e</sup> Refinement not completed, only to point where TraD binding pocket amino acid residues, waters, and co-factors were built and refined.		

---

**Table 2-2** Plasmids and oligonucleotides

---

Plasmid & oligos	Description & references
pJLTraD <sup>645-717</sup>	pT7-7 with a partial <i>traD</i> encoding residues 645 to 717; this work
pOX38-Km	Km <sup>r</sup> Cm <sup>r</sup> Tra <sup>+</sup> FinO <sup>-</sup> ; (Chandler & Galas, 1983)
pT7-7	Amp <sup>r</sup> ; (Tabor & Richardson, 1985)
JLU261	TAG AAT TCA CCA TCA CCA TCA CCA TGA GAA CCT GTA CTT CCA AGG GAT CGA GCA GGA GCT GAA AAT G
JLU207	TGG GGA TCC TGA GAA TTG AAG ACT GGA G

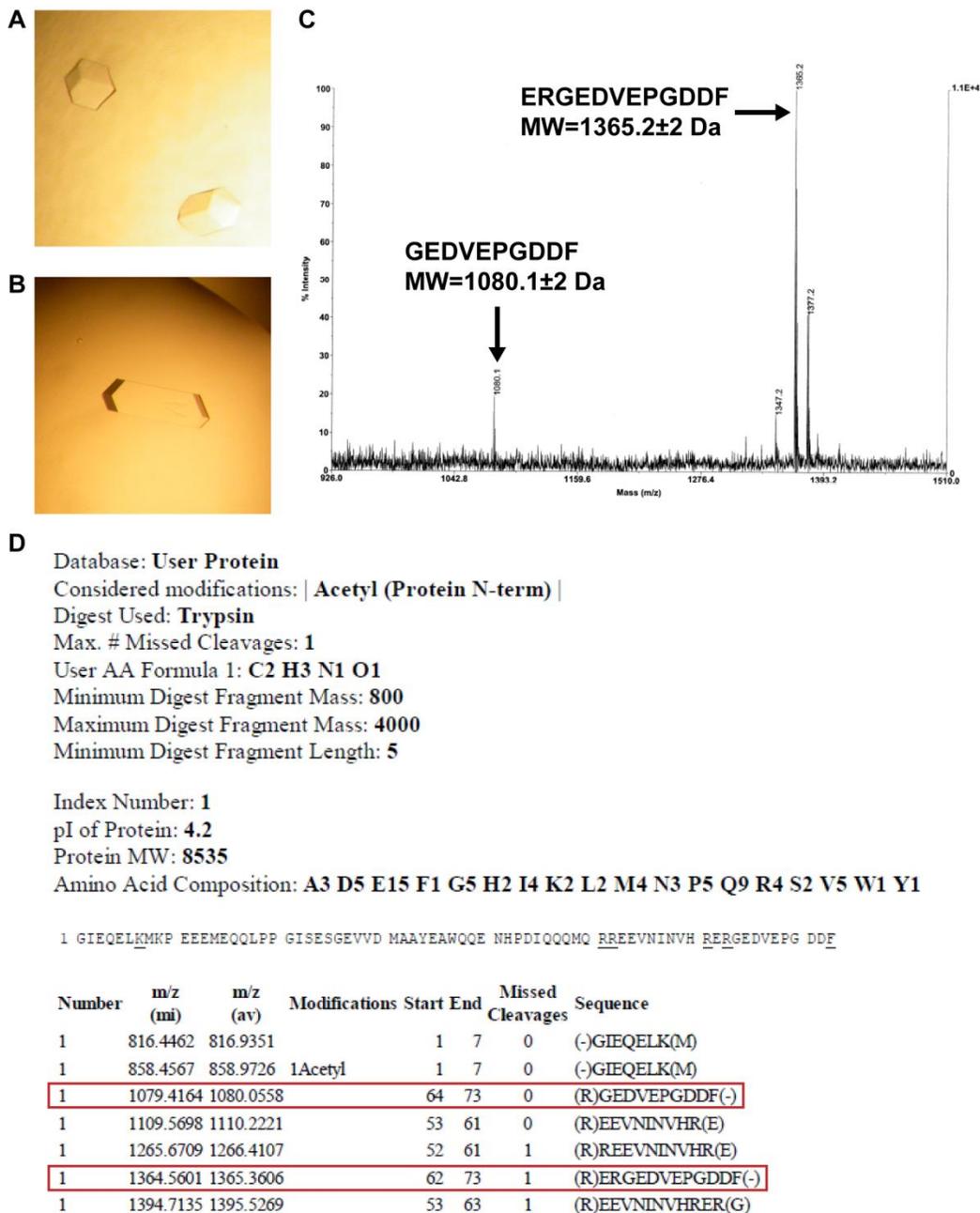


Figure 2-1 Crystallization and Mass Spectrometry of TraM<sup>58-127</sup>-TraD peptide crystals  
 a) TraM<sup>58-127</sup>-TraD<sup>711-717</sup> crystals from crystallization of TraM<sup>58-127</sup>-TraD<sup>645-717</sup> complexes  
 b) TraM<sup>58-127</sup> apoprotein crystals from crystallization of TraM<sup>58-127</sup>-TraD<sup>707-717</sup> complexes  
 c) Mass Spectrum of TraM<sup>58-127</sup>-TraD<sup>711-717</sup> crystals from crystallization of TraM<sup>58-127</sup>-TraD<sup>645-717</sup> complexes  
 d) Protein Prospector MS-Fit trypsin digest fragments of TraD<sup>645-717</sup>

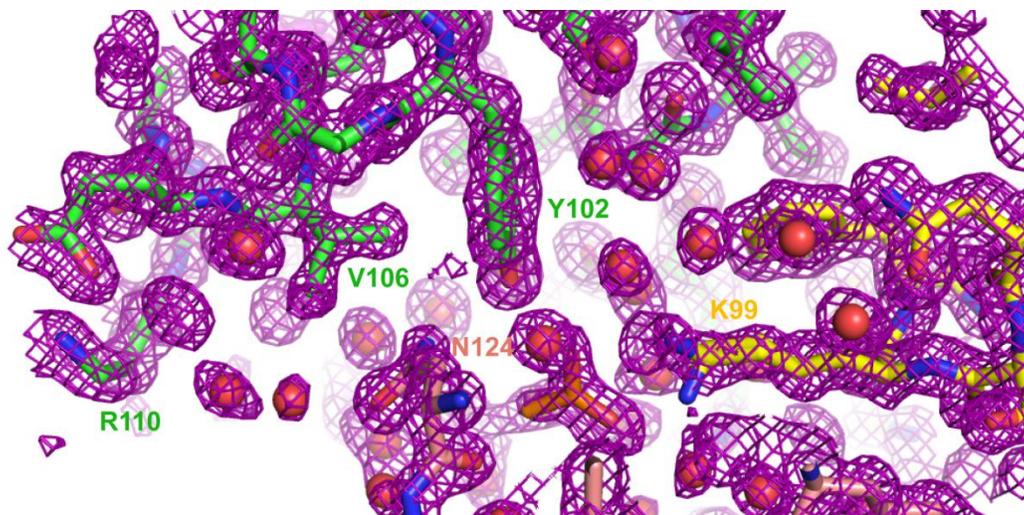


Figure 2-2 Crystal structure and SigmaA-weighted  $2mF_o-DF_c$  electron density of apo-TraM<sup>58-127</sup> determined from crystals grown from a TraM<sup>58-127</sup>-TraD<sup>707-717</sup> mixture. TraM chains are shown in green, yellow, and salmon. Waters are shown as red spheres. Acetate is colored orange.

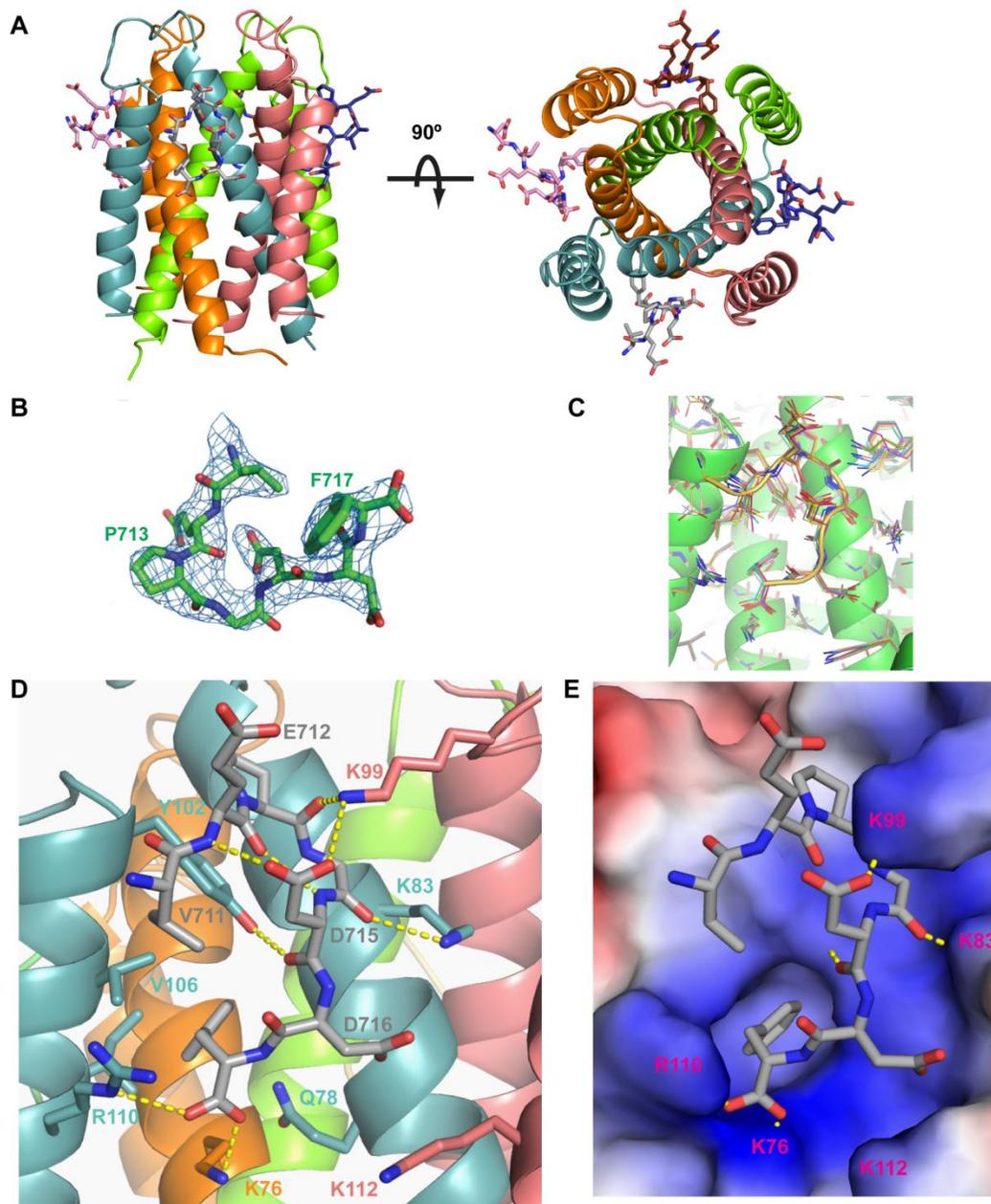


Figure 2-3 Crystal structure of TraM<sup>58-127</sup>-TraD<sup>711-717</sup>

- Orthogonal views of the TraD-TraM complex. The four TraM protomers that constitute the tetramerization domain are colored orange, green, teal, and salmon. The four TraD peptides are in pink, brown, grey, and blue.
- The electron density of the TraD peptide bound to TraM. The electron density is contoured at  $1.0 \sigma$  to  $2.55 \text{ \AA}$  resolution, phased by prime-and-switch method to remove model bias. The final refined model of the peptide is shown.
- Alignment of 8 non-crystallographic symmetry copies of the TraD C-terminus and binding pocket.
- Detailed view of TraM<sup>58-127</sup>-TraD<sup>711-717</sup> interactions. TraM and TraD are colored as in Figure 2-3A. Hydrogen bonding and electrostatic interactions  $<3.5 \text{ \AA}$  are indicated by yellow dashed lines.
- Same view as Figure 2-3D shown as an electrostatic surface. Electrostatic interactions  $<3.5 \text{ \AA}$  with TraM basic residues are indicated by yellow dashed lines.

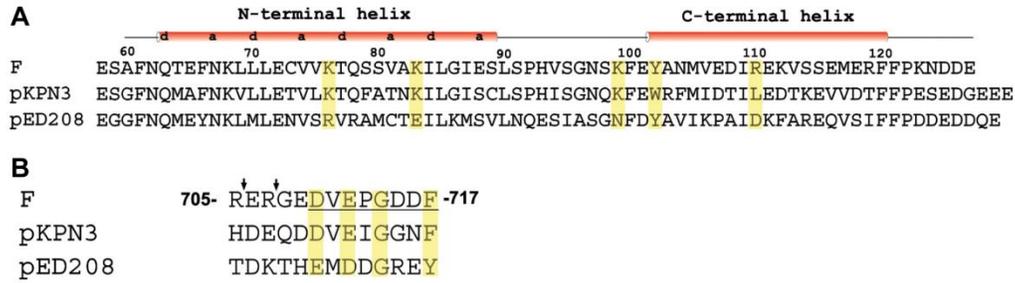


Figure 2-4 Sequence alignment of TraM and TraD homologues from F, pKPN3, and pED208

- Alignment of amino acid sequences of the TraM C-terminal tetramerization domain from F, pKPN3 and pED208 (Genbank accession No. BAA97941, YP\_001338610, and AAM90702, respectively); secondary structural elements were obtained from the X-ray structure of TraM<sup>58-127</sup>-TraD tail. The *a* and *d* positions of the heptad repeats in the N-terminal helices of TraM forming the central coiled coils are highlighted.
- Alignment of the amino acid sequences of the TraD C-terminal tails from F, pKPN3 and pED208 (Genbank accession No. BAA97972, YP\_001338644, and AAM90726, respectively). The underlined sequence is observed in the crystal structure of TraD-TraM complex, and arrows indicate sites of proteolysis.

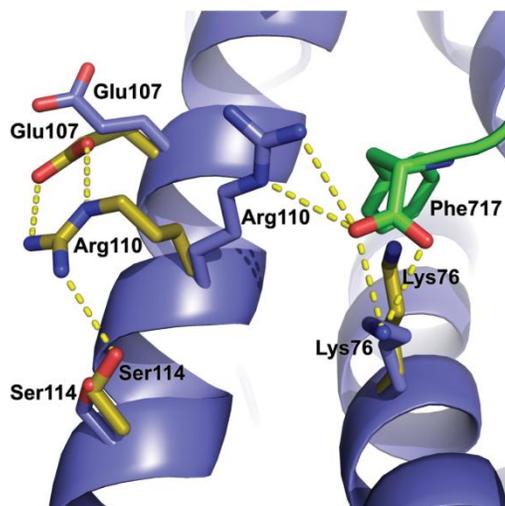


Figure 2-5 Comparison of the structure of TraD-TraM complex with the unliganded TraM structure  
Unliganded TraM is colored yellow. In the complex, TraM is colored blue, and TraD is colored green.

## References

- Beranek A, Zettl M, Lorenzoni K, Schauer A, Manhart M & Koraimann G (2004) Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. *J Bacteriol* **186**: 6999-7006.
- Champion PA, Stanley SA, Champion MM, Brown EJ & Cox JS (2006) C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*. *Science* **313**: 1632-1636.
- Chandler M & Galas DJ (1983) IS1-mediated tandem duplication of plasmid pBR322. Dependence on *recA* and on DNA polymerase I. *J Mol Biol* **165**: 183-190.
- Csitkovits VC, Dermic D & Zechner EL (2004) Concomitant reconstitution of Tral-catalyzed DNA transesterase and DNA helicase activity in vitro. *J Biol Chem* **279**: 45477-45484.
- Disque-Kochem C & Dreiseikelmann B (1997) The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD in vitro. *J Bacteriol* **179**: 6133-6137.
- Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132.
- Kupelwieser G, Schwab M, Hogenauer G, Koraimann G & Zechner EL (1998) Transfer protein TraM stimulates Tral-catalyzed cleavage of the transfer origin of plasmid R1 in vivo. *J Mol Biol* **275**: 81-94.
- Lang S, Kirchberger PC, Gruber CJ, *et al.* (2011) An activation domain of plasmid R1 Tral protein delineates stages of gene transfer initiation. *Mol Microbiol* **82**: 1071-1085.
- Lang S, Gruber K, Mihajlovic S, *et al.* (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol Microbiol* **78**: 1539-1555.
- Langer G, Cohen SX, Lamzin VS & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**: 1171-1179.
- Lu J & Frost LS (2005) Mutations in the C-terminal region of TraM provide evidence for in vivo TraM-TraD interactions during F-plasmid conjugation. *J Bacteriol* **187**: 4767-4773.
- Lu J, Wong JJ, Edwards RA, Manchak J, Frost LS & Glover JN (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.
- Lu J, Edwards RA, Wong JJ, Manchak J, Scott PG, Frost LS & Glover JN (2006) Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. *Embo J* **25**: 2930-2939.
- Miller DL & Schildbach JF (2003) Evidence for a monomeric intermediate in the reversible unfolding of F factor TraM. *J Biol Chem* **278**: 10400-10407.
- Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**: 240-255.
- Nagai H, Cambronne ED, Kagan JC, Amor JC, Kahn RA & Roy CR (2005) A C-terminal translocation signal required for Dot/Icm-dependent delivery of the *Legionella* RalF protein to host cells. *Proc Natl Acad Sci U S A* **102**: 826-831.

- Ragonese H, Haisch D, Villareal E, Choi JH & Matson SW (2007) The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at oriT through an interaction with Tral. *Mol Microbiol* **63**: 1173-1184.
- Renshaw PS, Lightbody KL, Veverka V, *et al.* (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *Embo J* **24**: 2491-2498.
- Sastre JI, Cabezon E & de la Cruz F (1998) The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J Bacteriol* **180**: 6039-6042.
- Tabor S & Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* **82**: 1074-1078.
- Terwilliger TC (2004) Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Crystallogr D Biol Crystallogr* **60**: 2144-2149.
- Vagin A & Teplyakov A (2000) An approach to multi-copy search in molecular replacement. *Acta Crystallogr D Biol Crystallogr* **56**: 1622-1624.
- Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouweland A & Hooykaas PJ (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proc Natl Acad Sci U S A* **102**: 832-837.

## Chapter 3

### Functional Studies of TraD C-terminus -TraM interaction<sup>1</sup>

#### Overview

The objective of the work in this chapter was to test the role of the TraD tail on TraM binding and conjugation. The crystal structure of the TraD tail showed a high degree of electrostatic and hydrophobic complementarity, suggesting that the binding of the TraD tail to TraM is highly specific. Using an affinity pulldown assay, the last 8 amino acids of TraD were confirmed to be important for binding, as constructs of the cytoplasmic domain of TraD or TraD<sup>645-717</sup> showed greater ability to pull down TraM than those without the tail. The need for efficient binding of the TraD tail to TraM for conjugation was shown *in vivo* in a mating assay. The specific interactions observed between amino acid side chains were tested for their importance *in vivo* by site-directed mutants in TraM and TraD, with most mutations having a significant effect on conjugation. This confirmed the functional significance of key features of the structure, including the C-terminal carboxylate of TraD and the aromatic side chain of TraD Phe717. The role of the electrostatic interaction between TraM Lys99 and Asp715 was tested in a second-site suppressor experiment where the charges on the TraD and TraM residues were reversed. Conjugation was restored in the TraD K99E / TraM D712K containing cells.

---

<sup>1</sup> Part of this work was previously published: Lu J, Wong JJ, Edwards RA, Manchak J, Frost LS & Glover JN (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.

## Introduction

The crystal structure of TraM<sup>58-127</sup> in complex with TraD<sup>711-717</sup> reveals that they interact in a sequence-specific manner. These interactions are due to a combination of electrostatic complementarity, shape complementarity, and hydrophobic interactions. The electrostatic complementarity comes from hydrogen bonds and long-range interactions between acidic residues and C-terminal carboxylate in the TraD C-terminal tail and basic residues lining the TraM pocket. The most notable example of hydrophobic interactions and shape complementarity occurs between an aromatic amino acid side chain in TraD and a TraM pocket lined with aliphatic residues. In order to validate the crystal structure as relevant to the biological function of the proteins, the role of the TraD tail in binding to TraM *in vitro* and in conjugation *in vivo* was tested.

Residue-specific electrostatic interactions between TraD and TraM include that between the carboxylic acid of TraD Asp715 and the amino group of TraM Lys99, and that between the functional groups of two TraM basic residues (the guanidinium group of TraM Arg110 and the amino group of Lys76) and the C-terminal carboxylate of TraD. The hydrophobic pocket into which the phenyl side chain of Phe717 binds is lined by TraM residues Val106, Leu85, and Ile109. Additional non-specific and/or long-range electrostatic interactions are provided by TraM Lys83, Lys99, and Tyr102 side chain interactions with main chain carbonyls of TraD (Figure 3-1). A previous study was highly suggestive of Lys99 and Val106 having a role in TraM-TraD binding, as mutations in these residues resulted in no autoregulation or oligomerization defect but were nevertheless defective in conjugation. One of the mutants, K99E, was confirmed to be less

efficient at binding to TraD than wild-type (Lu & Frost, 2005). The crystal structure provides validation for the results of that study.

TraM and TraD mutant proteins were created by site-directed mutagenesis for use in functional assays. The role of the entire TraD tail as observed in the crystal structure was tested with truncation mutants where the last 8 amino acids were deleted. In addition, the contribution of individual amino acids to TraD-TraM binding was tested with single-site mutants. Mutant proteins were tested in an *in vitro* pulldown assay to directly detect the effect of the residue on TraM-TraD binding, and in an *in vivo* conjugation assay to verify the biological relevance of the TraM-TraD interaction. Our *in vitro* results show that the last 8 amino acids are crucial for full TraM-binding affinity, which is correlated with conjugation frequency *in vivo*. Mutation of the majority of the residues tested had detrimental effects on conjugation *in vivo*. In summary, our results validate the observed interactions in the crystal structure.

## Results

### *The C-terminal tail of TraD is necessary for efficient interactions with TraM in solution*

We used Ni-NTA affinity chromatography to test the importance of the C-terminal 8 amino acids of TraD for interactions with TraM in solution (Lu & Frost, 2005, Lu, *et al.*, 2006). His<sub>6</sub>-tagged TraD cytoplasmic domain, either full length (TraD<sup>125-717</sup>) or lacking the C-terminal 8 residue tail (TraD<sup>125-709</sup>), was bound to Ni-NTA agarose and used to pull down purified full length TraM. The results showed that while TraD<sup>125-717</sup> interacts with TraM in this assay, interaction with TraD<sup>125-709</sup> is much weaker when equivalent amounts of protein are used (Figure 3-2).

Similar results were seen with TraD<sup>645-717</sup> compared to TraD<sup>645-709</sup> (Figure 3-2). These results indicate that the last 8 residues of TraD are required for efficient interaction with TraM.

The converse experiment was performed in an attempt to show that the last 8 amino acids of TraD are able to bind to TraM. The ability of a ubiquitin-TraD<sup>705-717</sup> fusion protein (hUb-D12) to pulldown TraM was tested in comparison to wild-type ubiquitin (hUb), but no binding for either protein was detected (data not shown).

*The C-terminal tail of TraD is necessary for efficient conjugation in vivo*

To understand the importance of the C-terminal tail of TraD for conjugation, we tested the ability of TraD mutants to complement conjugation of a plasmid derived from F, pOX38-D411, in which *traD* is knocked out (Table 3-1). Deletion of the C-terminal 141 amino acids of TraD (TraD 576\*) resulted in a 10<sup>4</sup>-fold decrease in conjugation efficiency, compared to wild type TraD, consistent with previous results (Sastre, *et al.*, 1998). This deletion removes the entire C-terminal extension but leaves intact the ATPase domain. Deletion of just the C-terminal 8 amino acids (TraD 709\*) resulted in a 10<sup>3</sup>-fold decrease in conjugation compared to the wild type control, indicating that this region is critical for the normal function of the C-terminal tail in initiating contact with TraM.

To identify the relative importance of different residues in the TraD tail to conjugation, a subset of the residues within TraD<sup>709-717</sup> were individually mutated and assessed in the mating assay (Table 3-1). Mutation of Phe717, which constitutes the hydrophobic core of the TraD portion of the interface, to alanine

yielded the largest effect with a  $10^4$ -fold reduction in mating efficiency. Several negatively charged residues in TraD make long-range electrostatic interactions with positively charged residues in TraM. We mutated two of these residues, Glu712 and Asp715, to lysine to test the importance of these interactions. Only TraD D715K showed a significant loss in conjugation (~10-fold), whereas TraD E712K yielded no significant change in conjugation efficiency. TraD Asp715 not only is in proximity with Lys99 in TraM, but also stabilizes the TraD  $\beta$ -turn through a hydrogen bond between the carboxylate of Asp715 and the main chain amide of Glu712. This hydrogen bonding interaction may explain the increased sensitivity of Asp715 to mutation compared to Glu712.

The TraD C-terminal carboxylate makes close contact with positively charged groups in TraM, Lys76 and Arg110. To test the importance of this group, we created a mutant with an additional glycine residue appended to the TraD C-terminus (TraD<sup>\*</sup>718G) (Table 3-1). This mutation reduced conjugation efficiency by  $10^3$ -fold, similar to the TraD F717A and TraD 709\* mutations. This result demonstrates the importance of the C-terminal carboxylate in TraM – TraD interactions.

TraM residues lining the TraD binding pocket were also tested for their role in conjugation. Consistent with previous results, K99E and V106A resulted in conjugative defect. R110E and K76E also resulted in conjugative defect, confirming their importance in forming electrostatic interactions with the free c-terminal carboxylate of the TraD peptide (Table 3-2). Their effects were confirmed to be not due to lack of TraM-binding ability, as all mutants were able to suppress LacZ reporter gene activity with close to wild-type levels (Table 3-2).

### *Suppression of the TraM K99E conjugation defect by a compensatory mutation in TraD*

TraM K99E was initially identified in a screen for random *traM* mutations that greatly decreased conjugation efficiency (Lu, *et al.*, 2003, Csitkovits, *et al.*, 2004). Unlike the other mutants isolated in this screen, the TraM K99E mutation did not affect either TraM tetramerization, or its ability to bind DNA and repress its own promoter, but abrogated interactions with TraD *in vitro* (Lu & Frost, 2005). The structure of the TraM – TraD complex explains these results. Lys99 lies over the TraD C-terminal tail, hydrogen bonding with the main chain carbonyl of Pro713 of TraD, and making long range electrostatic interactions with Glu712 and Asp715 (Figure 3-3). Thus the lysine to glutamate substitution at residue 99 would be expected to disrupt this hydrogen bonding interaction and result in electrostatic repulsions with Glu712 and Asp715. We reasoned that if this was the case, that TraM K99E could be rescued by compensatory mutations in TraD that introduce positively charged side chains at either residue 712 or 715. To test this hypothesis, we generated an F derivative plasmid, pOX38-DM, which is deficient for both TraM and TraD, and tested the ability of TraM and TraD mutants, either alone or in combination, to facilitate the conjugation of pOX38-DM when supplied *in trans* (Table 3-3).

As previously shown (Lu & Frost, 2005), TraM K99E results in a dramatic,  $10^5$ -fold reduction in conjugation efficiency compared to wild type TraM in cells expressing wild type TraD. This defect is significantly rescued ( $\sim 10^2$ -fold) by a compensatory mutation TraD E712K. Since TraD E712K does not significantly alter conjugation when co-expressed with wild type TraM, this result strongly suggests that electrostatic repulsions between TraM and TraD are responsible

for the conjugation defect in TraM K99E/TraD wt cells, and that electrostatic interactions between TraM Lys99 and TraD Glu712 stabilize the interaction of the wild type proteins. In contrast, TraD D715K alone did not rescue TraM K99E, perhaps because this side chain also stabilizes the  $\beta$ -turn in the TraD tail (Figure 3-3).

## Discussion

*In vivo* mating assays indicated that deletion of the C-terminal 8-residue tail of TraD resulted in a 500-fold decrease in the mating efficiency of the F plasmid (Table 3-1), whereas the naturally occurring fertility inhibition system (FinOP) decreases the mating efficiency of the F or F- like plasmids by approximately 100-fold (Lu, *et al.*, 2002). Point mutations potentially affecting interactions between TraD and TraM decreased the efficiency of F conjugation (Table 3-1 and 3-2). Structure-based second-site suppressor experiments showed that a mutation in TraD (E712K) could rescue the TraM K99E conjugation defect by over 30-fold (Table 3-3). These results demonstrate that the TraD-TraM interaction visualized in our crystal structure is involved in F conjugation.

A residual level of mating efficiency was observed in the F conjugation system with a deletion of the very C-terminal 8-residues of TraD (Table 3-1), suggesting that other regions of TraD might also interact with TraM or other components of the relaxosome to achieve a basal level of DNA transfer in the absence of the C-terminal tail. This is consistent with previously reported results from conjugation assays with TraD truncation mutants. TraD truncation at residue 576 allows it to mobilize the plasmids R388 and RSF1010 with a  $10^3$ -fold

increase in conjugative frequency, while decreasing transfer of the F plasmid by  $10^{-4}$  (Sastre, *et al.*, 1998). The TraD C-terminal extension appears to mediate specificity for its cognate TraM, while preventing it from mobilizing relaxosomes from other plasmids.

Several methods have been attempted to demonstrate *in vitro* a direct interaction between the last 8 amino acids of TraD and TraM. Since no pulldown of TraM was detected with His<sup>6</sup>-tagged hUb-D12, the converse experiment was attempted with GST-tagged TraM and *in vitro* transcribed S<sup>35</sup>-labelled hUb-D12 followed by detection with anti-Ub antibodies, but no pulldown was detected. Crosslinking of hUB-D12 to GST-TraM with DTBP and DSP was attempted during the pulldown experiment, but no significantly higher level of TraD pulldown was detected in hUb-D12 compared with the hUb control (data not shown). No interaction between synthesized TraD<sup>705-717</sup> and TraD<sup>707-717</sup> peptides and TraM<sup>58-127</sup> was detectable by isothermal titration calorimetry and fluorescence anisotropy (data not shown).

Because we were unable to detect interactions between the TraD C-terminal peptide and TraM interaction *in vitro*, we hypothesized that either the oligomerization of TraD or TraD sequences N-terminal to the 8 residue C-terminal tail may be an important factor in TraD-TraM interactions. As TraD exists as a hexameric, membrane anchored ring, and the TraM tetramer binds to multiple sites on plasmid DNA (Di Laurenzio, *et al.*, 1992, Fekete & Frost, 2002, Haft, *et al.*, 2007), multivalent contacts between the TraD hexamer and multiple TraM tetramers on DNA likely stabilize this complex *in vivo* (Figure 3-4). The structure of the TraM-TraD complex indicates that as many as 4 TraD C-terminal

tails can simultaneously bind a single TraM tetramer. However, constraints imposed by the other domains of these proteins, their oligomerization, interactions with DNA, and perhaps other components of the conjugation machinery will undoubtedly influence the stoichiometry and overall architecture of the complex. The lack of stable secondary structure in TraD<sup>645-717</sup> as determined by circular dichroism spectroscopy (Lu and Glover, unpublished observations) may provide a degree of flexibility that is important to initiate the formation of this oligomeric protein-DNA complex.

To test this idea, we assayed for interactions between TraD and TraM bound to DNA using electrophoretic mobility shift assays (EMSA). Titration of TraD<sup>137-717</sup> against TraM bound to *sbmA* did not show any evidence of TraD-TraM-DNA complex formation (data not shown). We then hypothesized that the hexameric state of TraD, which is known to be mediated by its transmembrane domain (Haft, *et al.*, 2007), may not be recapitulated with the cytoplasmic domain constructs, and therefore full-length TraD may be needed to observe interaction with DNA-bound TraM. Purification of full-length TraD, which required solubilization in a detergent fos-choline, was carried out. However, interference of the detergent with the ability of TraM to bind to DNA precluded the titration of TraD onto TraM-*sbmA* complexes (data not shown). As TraM [F<sup>1-55</sup> pED208<sup>56-127</sup>] does not show the same sensitivity to detergent as wild-type F TraM, a possible future experiment to test for TraD interaction with TraM-*sbmA* is to obtain purified full-length pED208 TraD for titration onto pED208 TraM-*sbmA* complexes or F<sup>1-55</sup> pED208<sup>56-127</sup>-F *sbmA* complexes.

## Materials and Methods

### *Bacterial growth media and strains*

Cells were grown in LB (Luria-Bertani) broth or on LB solid medium containing appropriate antibiotics. Antibiotics were used at the following final concentrations: ampicillin (Amp), 50 µg/mL; kanamycin (Km), 25 µg/mL; spectinomycin (Spc), 100 µg/ml; and nalidixic acid (Nal), 20 µg/mL. The following *Escherichia coli* strains were used: XK1200 [ $F^-$  Nal<sup>r</sup>  $\Delta$ lacU124  $\Delta$ (nadA aroG gal attLbio gyrA; (Moore, et al., 1987), MC4100 ( $F^-$ , Sm<sup>r</sup>  $\Delta$ lacU169, araD139, rpsL150, relA1, ptsF, rbs, flbB5301); (Sauter, et al., 1992), DH5  $\alpha$  [ $\Delta$ lacU169 ( $\Phi$ 80dlacZ $\Delta$ M15) supE44 hsdR17 recA1 endA1 gyrA96 (Nal<sup>r</sup>) thi-1 relA1] (Hanahan, 1983), and BL21(DE3) [ $F^-$  dcm ompT hsdS ( $r_B^-$  m $_B^-$ ) gal  $\lambda$ (DE3)] (Stratagene).

### *Cloning of His<sup>6</sup>-TraD constructs, TraD mutants, and TraM*

Cloning of TraD His<sup>6</sup>-TraD constructs and TraM mutant proteins was carried out as described in (Lu, et al., 2008).

pJMACTraM and pJMACTraM<sup>K99E</sup> were constructed by ligating the EcoRI-ScaI fragment from pACYC184 to the EcoRI-ScaI fragments from pRFM200 and pJLM105, respectively. Site-directed mutagenesis in *traM* by overlap extension was performed as described previously (Lu, et al., 2006). The primer pairs used for introducing point mutations in *traM* were JLU284 and JLU284rev for K76E, and JLU285 and JLU285rev for R110E (Table 3-4).

pJMTraD<sup>front</sup> was constructed by ligating the *EcoRI-PstI* fragment from pNLK5 to the *EcoRI-PstI* fragment of pBAD24. The 3' half of TraD was generated by PCR product using the primers LFR120 and JMA66, which was then blunt-end cloned into pTOPO (Invitrogen). The corresponding *PstI-HindIII* fragment in pTOPO was ligated to pJMTraD<sup>front</sup> digested with the same enzymes, to generate pJMTraD. Similar to the construction of pJMTraD C-terminal mutations were generated through PCR amplification of the 3' half of TraD using primer LFR120 and mutagenesis primers listed in Table 3-4.

An *NdeI-EcoRI* fragment of the PCR product amplified from pBAD24-TraD using the forward primer JWG01 and the reverse primer JWG02 was blunt-ended and cloned using a TOPO kit (Invitrogen). The corresponding *NdeI-EcoRI* fragment was ligated into the *NdeI-EcoRI* fragment of pT7-7, resulting in pJWTraD<sup>125-717</sup>. pJWTraD<sup>125-709</sup> was constructed similarly, except using the reverse primer JWG03. Plasmid pJLTraD<sup>645-717</sup>, which encodes the C-terminal 73 residues (645 to 717) of TraD, was constructed by ligating the *EcoRI-BamHI* fragment of pT7-7 to the *EcoRI-BamHI* fragment of DNA amplified from pOX38-Km using JLU261 and JLU207 as primers. Plasmid pJLTraD<sup>645-709</sup> was made similarly to pJLTraD<sup>645-717</sup>, except JWG03 was used instead of JLU207.

The *traM-traD* double mutant derivative of pOX38-Km was generated by inserting a chloramphenicol (Cm) cassette in the *traM* gene of pOX38-D411. Briefly, the *traM* gene amplified from pRS27 using primers LFR28 and RW168 was cloned into the TOPO vector. The *traM* gene was subsequently disrupted by insertion of a *SmaI* fragment containing a Cm cassette from pUC4CIXX at the *HincII* site. The disrupted *traM* gene was PCR-amplified using primers LFR28

and RW168 and was transformed into *E. coli* strain DY330 containing pOX38-D411 for recombination according to the previously described procedure (Yu, *et al.*, 2000). DY330 cells containing the double mutant pOX38-DM were selected on media containing chloramphenicol and kanamycin. DNA sequencing verified that pOX-DM contains a *tram* gene disrupted by a Cm cassette.

### *Mating assays*

Donor XK1200 cells and MC4100 recipient cells were grown to mid-exponential phase at 37°C, with arabinose added to 0.05% to donor cells for induction of 1 hour prior to mating. Mating assays were performed as previously described (Lu, *et al.*, 2002). Mating efficiency is calculated as the number of transconjugants divided by the number of donors. Each assay was repeated 3-4 times and the averaged values are reported. Standard deviations of all mating assays were within one log unit.

### *Affinity chromatography analysis of TraD-TraM interaction*

75 mL of cultured cells expressing His<sub>6</sub>-TraD constructs (His<sub>6</sub>-TraD<sup>125-717</sup>, His<sub>6</sub>-TraD<sup>125-709</sup>, His<sub>6</sub>-TraD<sup>645-717</sup>, or His<sub>6</sub>-TraD<sup>645-709</sup>) were induced at ~0.7 OD<sub>600</sub> with 0.5mM IPTG for 3 hours, then pelleted. Cells were lysed by sonication in 15 mL of Buffer A (20 mM Tris, 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.5, 8 M Urea, 300 mM NaCl, 10% glycerol, 1 mM DTT) with 1 Complete Mini EDTA-free protease inhibitor tablet. After centrifugation at 17,000 x g for 40 minutes, 500 μL of supernatant was added to 30 μL of Ni-NTA beads and incubated at room temperature on shaker for 2 hours. After washing with 300 μL of Buffer A once and 1 mL of Buffer B (same as Buffer A except without urea) 3 times, beads were incubated

in 250  $\mu\text{L}$  of Buffer B with 2  $\mu\text{g}$  of purified TraM and 5  $\mu\text{g}$  of BSA on a shaker at room temperature for 1 hour. Beads were washed with 1 mL of Buffer B 3 times. TraD was eluted from beads with 100  $\mu\text{L}$  of Buffer C (250 mM imidazole in Buffer B). 10  $\mu\text{L}$  of eluate was taken for SDS-PAGE gel, and proteins were visualized by Coomassie blue staining or Western blot as previously described (Penfold, *et al.*, 1996).

**Table 3-1 Effect of TraD mutations on F conjugation**

TraD <sup>a</sup>	-	Wt	D576*	E709*	E712K	D715K	F717A	*718G
Mating efficiency <sup>b</sup>	$<5 \times 10^{-7}$	1	$5 \times 10^{-5}$	$2 \times 10^{-3}$	4	$9 \times 10^{-2}$	$2 \times 10^{-4}$	$9 \times 10^{-4}$

<sup>a</sup>“-“ represents vector control. “\*” represents a stop codon. Mutants are named after their codon changes.

<sup>b</sup>Determined by assaying donor ability of cells containing pOX38-D411 and pJMTraD or one of its mutant derivatives. Mating efficiency is normalized to that of the wild type TraD.

**Table 3-2 Effect of TraM mutations on F conjugation**

TraM <sup>a</sup>	-	Wt	K76E	R110E	K99E	V106A	N5D <sup>c</sup>
Mating efficiency <sup>b</sup>	$<5 \times 10^{-7}$	1	$5 \times 10^{-5}$	$2 \times 10^{-3}$	$<1 \times 10^{-5}$	$9 \times 10^{-2}$	$<1 \times 10^{-5}$
LacZ activity (MU)	45.0±3.1	7.9±0.6	7.8±0.5	7.7±0.6	8.0±0.3	8.8±0.6	44.0±2.1

<sup>a</sup> Wild type TraM or its mutants were expressed by pRFM200 or its derivatives, respectively. “-“ represents vector control. “wt” represents wild type. Mutants are named after their codon changes.

<sup>b</sup> Determined by assaying donor ability of cells containing pOX38-MK3, and pRFM200 or one of its mutant derivatives. Mating efficiency is normalized to that of the cells expressing wild type TraM from pRFM200.

<sup>c</sup> LacZ Positive control -TraM mutant deficient in autoregulation (Lu, *et al.*, 2004)

**Table 3-3 Compensatory mutations in TraD rescue TraM<sup>K99E</sup>**

TraM <sup>1</sup>	-	-	wt	wt	wt	wt
TraD <sup>1</sup>	-	wt	-	wt	E712K	D715K
Mating Efficiency <sup>2</sup>	$<1 \times 10^{-7}$	$<1 \times 10^{-7}$	$<1 \times 10^{-7}$	1	$4 \times 10^{-1}$	$5 \times 10^{-3}$
TraM <sup>1</sup>			K99E	K99E	K99E	K99E
TraD <sup>1</sup>			wt	E712K	D715K	E712K D715K
Mating Efficiency <sup>2</sup>			$3 \times 10^{-5}$	$1 \times 10^{-3}$	$7 \times 10^{-5}$	$2 \times 10^{-4}$

<sup>1</sup>“-“ represents vector control. “wt” represents wild type. Mutants are named after their codon changes.

<sup>2</sup> Determined by assaying donor ability of cells containing pOX38-DM, pJMACTraM or pJMACTraM<sup>K99E</sup>, and pJMTraD or one of its mutant derivatives. Mating efficiency is normalized to that of the cells containing wild type TraD and TraM.

Mating assays were performed by Jan Manchak, Jun Lu, and Joyce Wong

**Table 3-4** Plasmids and oligonucleotides

Plasmid & oligos	Description & references
pACYC184	Tc <sup>r</sup> ; (Chang & Cohen, 1978)
pACPM24fs:: <i>lacZ</i>	pACYC184 with a <i>PtraM-lacZ</i> fusion for promoter repression assays (Lu, <i>et al.</i> , 2003)
pBAD24	Amp <sup>r</sup> ; (Guzman, <i>et al.</i> , 1995)
pJLM105	pT7-4 with an F <i>traM</i> mutant, <i>K99E</i> ; (Lu & Frost, 2005)
pJLTraD <sup>645-717</sup>	pT7-7 with a partial <i>traD</i> encoding residues 645 to 717; (Lu, <i>et al.</i> , 2008)
pJLTraD <sup>645-709</sup>	pT7-7 with a partial <i>traD</i> encoding residues 645 to 709; (Lu, <i>et al.</i> , 2008)
pJMACTraM	pACYC184 with <i>traM</i> from pRFM200; (Lu, <i>et al.</i> , 2008)
pJMACTraM <sup>K99E</sup>	pACYC184 with <i>traM</i> mutant <i>K99E</i> from pJLM105; (Lu, <i>et al.</i> , 2008)
pJMTraD	pBAD24- <i>traD</i> front with <i>traD</i> from pNLK5; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>D576*</sup>	pBAD24 with a <i>traD</i> mutant encoding first 576 residues; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>D715K</sup>	pBAD24 with a <i>traD</i> mutant <i>D715K</i> ; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>E709*</sup>	pBAD24 with a <i>traD</i> mutant missing the last 8 residues; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>E712K</sup>	pBAD24 with a <i>traD</i> mutant <i>E712K</i> ; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>E712K+D715K</sup>	pBAD24 with a <i>traD</i> double mutant <i>D715K</i> plus <i>D715K</i> ; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>F717A</sup>	pBAD24 with a <i>traD</i> mutant <i>F717A</i> ; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>front</sup>	pBAD24 with the front half of <i>traD</i> from pNLK5; (Lu, <i>et al.</i> , 2008)
pJMTraD <sup>*718G</sup>	pBAD24 with a <i>traD</i> mutant with an extra C-terminal glycine codon; (Lu, <i>et al.</i> , 2008)
pJWTraD <sup>125-717</sup>	pT7-7 with a partial <i>traD</i> encoding residues 125 to 717; (Lu, <i>et al.</i> , 2008)
pJWTraD <sup>125-709</sup>	pT7-7 with a partial <i>traD</i> encoding residues 125 to 709; (Lu, <i>et al.</i> , 2008)
pRS27	pSC101 with an F fragment from <i>oriT</i> to <i>traV</i> (Achtman, <i>et al.</i> , 1978)
pUC4CIXX	Derivative of pUC4KIXX (Gibco/BRL) carrying a Cm resistance cassette (Maneewannakul, <i>et al.</i> , 1992)
pNLK5	pBAD18 with F <i>traD</i> ; (Lee, <i>et al.</i> , 1999)
pOX38-D411	Km <sup>r</sup> Cm <sup>r</sup> Tra <sup>+</sup> FinO <sup>-</sup> TraD <sup>-</sup> ; (Maneewannakul, <i>et al.</i> , 1996)
pOX38-DM	Km <sup>r</sup> Cm <sup>r</sup> Tra <sup>+</sup> FinO <sup>-</sup> TraD <sup>-</sup> TraM <sup>-</sup> ; (Lu, <i>et al.</i> , 2008)
pOX38-Km	Km <sup>r</sup> Cm <sup>r</sup> Tra <sup>+</sup> FinO <sup>-</sup> ; (Chandler & Galas, 1983)
pRFM200	pT7-5 with an F <i>Bst</i> BI- <i>Bgl</i> II fragment from <i>traM</i> to <i>P<sub>finP</sub></i> ; (Lu, <i>et al.</i> , 2003)
pT7-7	Amp <sup>r</sup> ; (Tabor & Richardson, 1985)

JLU261 TAG AAT TCA CCA TCA CCA TCA CCA TGA GAA CCT GTA CTT  
CCA AGG GAT CGA GCA GGA GCT GAA AAT G

JLU207 TGG GGA TCC TGA GAA TTG AAG ACT GGA G

JLU284 GCTTCTTGAATGCGTTGTAGAAACACAATCATCAGTAGCGAAAA  
TTTTGGG

JLU284rev GTGTTTCTACAACGCATTCAAGAAGCAATTTATTAACACTCAG

JLU285 GGTTGAAGATATCGAGGAGAAGGTATCATCTGAGATGG

JLU285rev CCTTCTCCTCGATATCTTCAACCATATTGGCATATTCAAAC

LFR120 CGT ATC TGA TGC GTA ATG ACC

JMA66 AAG CTT CAG AAA TCA TCT CCC GGC

JMA67 CAA GCT TAC TCC CCG CGC TCC CGG; for TraD<sup>D576\*</sup>

JMA68 CAA GCT TAG TCA CGC GGA ATA AAC CTC; for TraD<sup>E709\*</sup>

JMA70 AAG CTT CAT CCGA AAT CAT CTC CCG GCT CAA C; for  
TraD<sup>\*718G</sup>

JMA71 AAG CTT CAG GCA TCA TCT CCC GGC TTC AAC; for TraD<sup>F717A</sup>

JMA72 AAG CTT CAG AAA TCT TTT CCC GGC TCA ACA TC; for  
TraD<sup>D715K</sup>

JMA73 AAG CTT CAG AAA TCA TCT CCC GGC TTA ACA TCC TC; for  
TraD<sup>E712K</sup>

JMA75 AAG CTT CAG AAA TCT TTT CCC GGC TTA ACA TCC TC; for  
TraD<sup>E712K+D715K</sup>

JWG01 CAT ATG CAT CAT CAT CAT CAT GAA AAT CTT TAT TTT  
CAA GGT GGA TCC ATT ACC TTC TTT GTT GTC TCC TGG

JWG02 CGG AAT TCT TAG AA ATC ATC TCC GGG CTC AAC

JWG03 GCG AAT TCT TAC TCC CCG CGC TCC CGG TG

LFR28 CGAATTCGTCCCTGTTTGCATTATGA

RW168 TTTCCAGCAGATCTATTTGACGAGCA

---

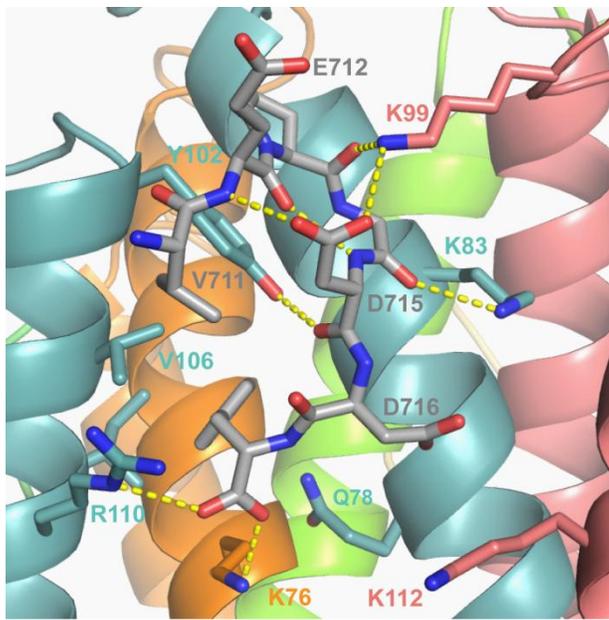


Figure 3-1 Detailed view of TraM<sup>58-127</sup>-TraD<sup>711-717</sup> interactions  
The four TraM protomers that constitute the tetramerization domain are colored orange, green, teal, and salmon. The TraD peptide are in grey. Hydrogen bonding and electrostatic interactions <3.5 Å are indicated by yellow dashed lines.

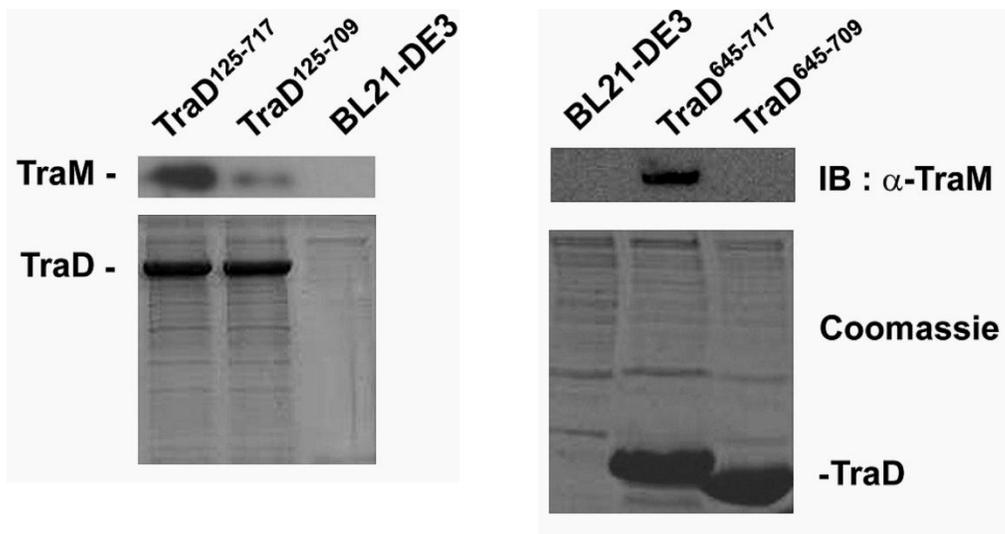


Figure 3-2 Affinity chromatography analysis of TraM interaction with TraD or its C-terminal truncation

Ni-NTA agarose beads pre-incubated with a TraD-free cell lysate or a cell lysate containing His<sub>6</sub>-tagged TraD<sup>137-717</sup> or His<sub>6</sub>-TraD<sup>137-709</sup> was used to pull down TraM. TraM was visualized by western blot using anti-TraM antibody (top panel); and TraD was visualized by Coomassie blue staining of the same immunoblot membrane (bottom panel).

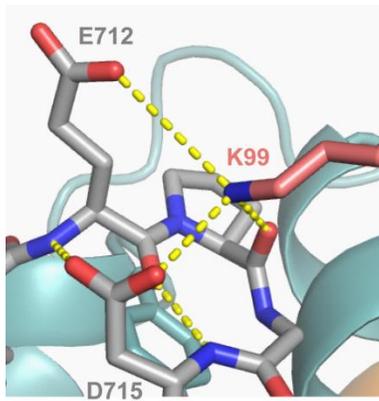


Figure 3-3 Detailed view of potential TraD E712 and D715 interactions with TraM K99. Potential electrostatic interactions between the E712 and D715 side chains and K99 are indicated with dashed yellow lines. The role of these interactions in TraD-TraM interaction were tested by second-site suppressor mutations *in vivo*.

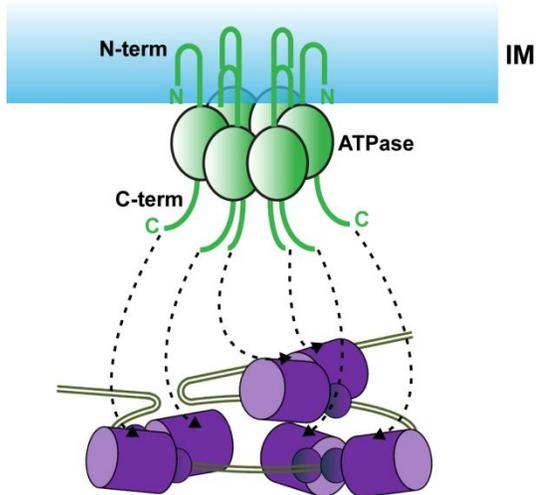


Figure 3-4 Model of TraM avidity effect in binding to TraD  
 IM=inner membrane. TraD is shown in green, and TraM in purple. TraM N-terminal domains are shown as ellipsoids, and TraM C-terminal domains are shown as cylinders. Multiple TraM tetramers are bound to 3 *sbmA* sites at *oriT* in a compact area due to nucleosome-like DNA wrapping. The localized concentration of TraM tetramers facilitates interaction between multiple TraM binding sites and multiple TraD C-termini.

## References

- Achtman M, Skurray RA, Thompson R, Helmuth R, Hall S, Beutin L & Clark AJ (1978) Assignment of tra cistrons to EcoRI fragments of F sex factor DNA. *J Bacteriol* **133**: 1383-1392.
- Chandler M & Galas DJ (1983) IS1-mediated tandem duplication of plasmid pBR322. Dependence on recA and on DNA polymerase I. *J Mol Biol* **165**: 183-190.
- Chang AC & Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**: 1141-1156.
- Csitkovits VC, Dermic D & Zechner EL (2004) Concomitant reconstitution of TraI-catalyzed DNA transesterase and DNA helicase activity in vitro. *J Biol Chem* **279**: 45477-45484.
- Di Laurenzio L, Frost LS & Paranchych W (1992) The TraM protein of the conjugative plasmid F binds to the origin of transfer of the F and ColE1 plasmids. *Mol Microbiol* **6**: 2951-2959.
- Fekete RA & Frost LS (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.
- Guzman LM, Belin D, Carson MJ & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Haft RJ, Gachelet EG, Nguyen T, Toussaint L, Chivian D & Traxler B (2007) In vivo oligomerization of the F conjugative coupling protein TraD. *J Bacteriol*.
- Haft RJ, Gachelet EG, Nguyen T, Toussaint L, Chivian D & Traxler B (2007) In vivo oligomerization of the F conjugative coupling protein TraD. *J Bacteriol* **189**: 6626-6634.
- Hanahan D (1983) Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**: 557-580.
- Lee MH, Kosuk N, Bailey J, Traxler B & Manoil C (1999) Analysis of F factor TraD membrane topology by use of gene fusions and trypsin-sensitive insertions. *J Bacteriol* **181**: 6108-6113.
- Lu J & Frost LS (2005) Mutations in the C-terminal region of TraM provide evidence for in vivo TraM-TraD interactions during F-plasmid conjugation. *J Bacteriol* **187**: 4767-4773.
- Lu J, Fekete RA & Frost LS (2003) A rapid screen for functional mutants of TraM, an autoregulatory protein required for F conjugation. *Mol Genet Genomics* **269**: 227-233.
- Lu J, Zhao W & Frost LS (2004) Mutational analysis of TraM correlates oligomerization and DNA binding with autoregulation and conjugative DNA transfer. *J Biol Chem* **279**: 55324-55333.
- Lu J, Wong JJ, Edwards RA, Manchak J, Frost LS & Glover JN (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.
- Lu J, Manchak J, Klimke W, Davidson C, Firth N, Skurray RA & Frost LS (2002) Analysis and characterization of the IncFV plasmid pED208 transfer region. *Plasmid* **48**: 24-37.
- Lu J, Edwards RA, Wong JJ, Manchak J, Scott PG, Frost LS & Glover JN (2006) Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. *Embo J* **25**: 2930-2939.

- Maneewannakul K, Kathir P, Endley S, Moore D, Manchak J, Frost L & Ippen-Ihler K (1996) Construction of derivatives of the F plasmid pOX-tra715: characterization of traY and traD mutants that can be complemented in trans. *Mol Microbiol* **22**: 197-205.
- Maneewannakul S, Kathir P & Ippen-Ihler K (1992) Characterization of the F plasmid mating aggregation gene traN and of a new F transfer region locus trbE. *J Mol Biol* **225**: 299-311.
- Moore D, Wu JH, Kathir P, Hamilton CM & Ippen-Ihler K (1987) Analysis of transfer genes and gene products within the traB-traC region of the Escherichia coli fertility factor, F. *J Bacteriol* **169**: 3994-4002.
- Penfold SS, Simon J & Frost LS (1996) Regulation of the expression of the traM gene of the F sex factor of Escherichia coli. *Mol Microbiol* **20**: 549-558.
- Sastre JI, Cabezon E & de la Cruz F (1998) The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J Bacteriol* **180**: 6039-6042.
- Sauter M, Bohm R & Bock A (1992) Mutational analysis of the operon (hyc) determining hydrogenase 3 formation in Escherichia coli. *Mol Microbiol* **6**: 1523-1532.
- Tabor S & Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* **82**: 1074-1078.
- Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG & Court DL (2000) An efficient recombination system for chromosome engineering in Escherichia coli. *Proc Natl Acad Sci U S A* **97**: 5978-5983.

## Chapter 4 Insights into cooperative binding of TraM to *oriT* DNA from the crystal structure of the pED208 TraM-*sbmA* complex<sup>1</sup>

### Overview

TraM of the F plasmid binds to its highest affinity binding site at *oriT*, *sbmA*, in a cooperative manner. EMSA and MALLS methods show that for both the F and pED208 plasmids, 2 tetramers of TraM bind to *sbmA*. pED208 TraM binds to its cognate *sbmA* site with high affinity and cooperativity. X-ray crystallography studies were carried out on the pED208 TraM-*sbmA* complex to elucidate the nature of TraM cooperative binding and the basis for plasmid-specific binding to *sbmA*. The structure of the complex was solved by a combination of molecular replacement and manual building. The structure of the apo-N-terminal domain of pED208 TraM was solved by molecular replacement. Two TraM tetramers bind to *sbmA* on opposite sides of the DNA without any protein-protein contact. The N-terminal domain was confirmed to be a ribbon-helix-helix (RHH) fold, which recognizes the bases of the pED208 GANTC binding motif in the major groove. The TraM-bound DNA is kinked due to electrostatic repulsion of the phosphate backbone from the acidic loop between the  $\alpha 1$  and  $\alpha 2$  helices of the RHH domain. In addition, the TraM-bound DNA is underwound, which places the GANTC sites bound by the RHH domains from the same tetramer in closer alignment than they would be if the *sbmA* DNA was in its B-form. EMSA assays show that the underwinding facilitates the binding of TraM tetramers and that the cooperative binding mechanism is conserved between pED208 and F.

---

<sup>1</sup> Part of this work was previously published: Wong JJ, Lu J, Edwards RA, Frost LS & Glover JNM (2011) Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**: 6775-6788.

## Introduction

F plasmid TraM binds to three sites (*sbmA*, *-B*, *-C*) within *oriT*. TraM represses its own gene by binding to *sbmA* and *sbmB* (Penfold, *et al.*, 1996), whereas *sbmC* is the most important site of the three for F conjugation (Fu, *et al.*, 1991). Cooperativity has been demonstrated for F TraM binding within *sbmA*, but not *sbmC*. Binding of F TraM *sbmA* is of high affinity ( $K_d = 5$  nM) and yields a single species of distinct migration in an electrophoretic mobility shift assay, whereas *sbmC* does not. Cooperativity also exists between the *sbmA* and *sbmB* sites, as well as between *sbmAB* and *sbmC*. At high F TraM concentrations, TraM binds to secondary sites that radiate out from these primary DNA binding sites, suggesting that it aggregates on the DNA in an orderly fashion (Fekete & Frost, 2002). Sequence-specific DNA recognition is mediated by the TraM N-terminal domain (residues 2-56 in F TraM), which forms dimers. The C-terminal domain of TraM (residues 58-127 in F TraM) is responsible for tetramerization (Verdino, *et al.*, 1999, Miller & Schildbach, 2003, Lu, *et al.*, 2004).

pED208 belongs to the IncFV incompatibility group and is a transfer-derepressed derivative of the  $F_{\text{lac}}$  plasmid originally isolated from *Salmonella typhi* (Falkow & Baron, 1962, Finlay, *et al.*, 1983). The derepression and multipiliation of pED208 is thought to be due to an insertion element in the *tray* gene (Finlay, *et al.*, 1986). Like F plasmid TraM, pED208 TraM binds to 3 sites in the *oriT* region, *sbmA*, *sbmB*, and *sbmC* (Di Laurenzio, *et al.*, 1991). pED208 TraM contains 127 amino acids and has a molecular weight of 14.6kDa. F TraM and pED208 TraM are 38% identical at the amino acid sequence level (Figure 4-1A and 4-8E) and bind to distinct *sbmA* DNAs. While both proteins bind *sbmA* sites containing sequence motifs spaced 12 bp apart with an inverted repeat

symmetry, the core sequence motif bound by pED208 TraM, GANTC (Di Laurenzio, *et al.*, 1991), which is also palindromic, is distinct from that bound by F TraM, A(G/C)CG(G/C)T (Figure 4-1B). In addition, the arrangement of relaxosome protein binding sites is in a different order in F (Figure 1-1C) compared to pED208 (Figure B-1A) (Frost, *et al.*, 1994, Di Laurenzio, *et al.*, 1995).

TraM proteins from F and various F-like plasmids function in a plasmid-specific manner are essential for conjugation, but only function in their cognate plasmid due to specificity in TraM-DNA interactions. However, the structural basis of plasmid-specific - DNA recognition by TraM is unknown. In addition, the mechanism for achieving cooperative DNA binding within and between the *sbm* sites, and how this relates to the proposed “signaling” function of TraM is not fully understood.

In this chapter I cover solution experiments on TraM-*sbmA* complexes that show that two tetramers of TraM bind to *sbmA* in a cooperative fashion in both F and F-like pED208 plasmid systems. The minimal length of *sbmA* required for pED208 TraM to bind with maximal affinity was also determined. The crystal structure of pED208 TraM bound to its *sbmA* site and that of the isolated N-terminal domain of pED208 TraM was determined. The structure of pED208 TraM bound to *sbmA* DNA reveals that alternating RHH modules from two different TraM tetramers contact staggered target GANTC sequence motifs within *sbmA*. In this way, each tetramer contacts a pair of GANTC motifs separated by 12 bp. Cooperative DNA recognition is achieved through the coordinated unwinding and kinking of the DNA to align alternating GANTC motifs on the same

side of the helix for recognition by TraM without protein-protein interaction. This unusual mechanism of cooperative binding allows TraM to simultaneously dictate plasmid specificity by binding to *oriT* and relaxosome-coupling protein interactions specificity by binding to TraD.

## Results

### *Stoichiometry of TraM binding to sbmA*

F TraM binds to three related DNA sequences, termed *sbmA*, *sbmB* and *sbmC*, all located within an ~200 bp region of DNA (*oriT*) located between the *traM* promoter and the plasmid *nic* site (Fekete & Frost, 2002). *sbmA* is bound with highest affinity and is distinguished by the highest level of sequence symmetry. Each *sbmA* contains 4 sequence repeats, A(G/C)CG(G/C)T, arranged around a centre of palindromic symmetry (Figure 4-1B). TraM itself is a tetramer (Verdino, *et al.*, 1999, Miller & Schildbach, 2003, Lu, *et al.*, 2004). To characterize the stoichiometry of binding of F TraM to its cognate *sbmA*, we used multi-angle laser light scattering (MALLS) to determine the molecular mass of TraM-DNA complexes purified by size exclusion chromatography (Figure 4-1C). This analysis showed that the TraM-*sbmA* complex forms a ~132 kDa complex. Given the mass of the *sbmA* DNA is ~19.8 kDa and the mass of a single TraM protomer is ~14.4 kDa, this result suggests that the stable complex formed contains two TraM tetramers bound to a single DNA.

Interactions between F TraM and a larger 75 bp DNA containing both the *sbmA* and *sbmB* sites was also characterized by MALLS (Figure 4-1C). This DNA formed a stable complex with TraM with molecular weight of ~260 kDa,

consistent with a binding stoichiometry of 4 tetramers per DNA (predicted M.W. = 274 kDa), indicating that a pair of F TraM tetramers also bind *sbmB*.

The TraM-*sbmA* result was confirmed using an electrophoretic mobility shift assay (EMSA) carried out at protein and DNA concentrations well in excess of the  $K_d$  of the complex. These results clearly demonstrate that 8 molar equivalents of F TraM are required to bind the *sbmA* duplex (Figure 4-1D). No evidence of higher mobility species corresponding to a single tetramer bound to *sbmA* was observed even at sub-stoichiometric amounts of TraM, indicating that pairs of TraM tetramers cooperatively recognize a single *sbmA*. In a similar experiment, 8 molar equivalents of pED208 TraM bound to its cognate DNA with no intermediate species, demonstrating that the requirement for 2 TraM tetramers to cooperatively bind *sbmA* is conserved between the two F-like plasmids (Figure 4-1E).

#### *Evidence for a ribbon-helix-helix fold in the N-terminal domain of TraM*

NMR analysis of the N-terminal domain of the R1 TraM protein suggested that this domain adopts a monomeric, helix-turn-helix structure, with a disordered N-terminal region at pH 4.0 (Stockner, *et al.*, 2001). However, sedimentation analyses of the DNA binding domains of F and R1 TraM clearly revealed that this domain adopts a dimeric structure at neutral pH (Verdino, *et al.*, 1999, Miller & Schildbach, 2003). Furthermore, analyses of the DNA binding properties of an extensive series of F TraM point mutants clearly indicated a critical role for the 10 N-terminal residues in specific DNA recognition (Lu, *et al.*, 2004). To reconcile these results, we hypothesized that the TraM DNA binding domain adopts a dimeric ribbon-helix-helix fold (Schreiter & Drennan, 2007) (Figure 4-2A), which

is commonly found in bacterial transcriptional repressors, as well as the R388 plasmid auxiliary DNA-processing protein TrwA (Moncalian & de la Cruz, 2004).

The  $\beta$ -ribbon of the RHH fold provides critical residues for DNA recognition that directly contact DNA base pairs. While members of the TraM family are quite similar throughout their N-terminal domains, many bind different DNA elements, providing a basis for allelic specificity (Schwab, *et al.*, 1993, Kupelwieser, *et al.*, 1998, Lu, *et al.*, 2002). For example, R100 TraM functions poorly in complementation assays for conjugative transfer of a TraM-deficient F plasmid derivative, pOX38-MK3 (Figure 4-2B). To understand the basis for this specificity, we generated a double mutant (R3K:I5N) of R100 TraM such that its predicted  $\beta$ -ribbon region is the same as that of the F plasmid TraM (Figure 4-2A). The side chains of residues 3 and 5 are predicted to be exposed on the surface of the RHH  $\beta$ -ribbon and in contact with DNA. When complementing pOX38-MK3, this R100 TraM mutant resulted in a 60,000-fold increase in conjugation efficiency compared to the wild-type R100 TraM. As a control, the F plasmid TraM N5D mutant, which is defective in DNA binding, was cloned in the same vector, resulting in undetectable levels of conjugation (Figure 4-2B). EMSA analysis indicated that the R100 double mutant has a dramatically increased binding affinity for F plasmid cognate DNA, *sbmA*, compared to the wild type (Figure 4-2C). Both F conjugation efficiency and *sbmA*-binding ability of the R100 TraM double mutant are slightly lower than those of F plasmid TraM, suggesting that some residues outside the N-terminal  $\beta$ -ribbon sequence might also contribute to cognate DNA binding.

*Minimum length of sbmA needed for maximal pED208 TraM binding*

Previous attempts to obtain the crystal structure of F TraM bound to *sbmA* were not successful. Large single crystals were obtained of this complex, but diffraction never exceeded 4.5 Å (Lu, J., unpublished findings). Therefore, we decided to try co-crystallization of the TraM homologue from the plasmid pED208 with its cognate *sbmA* in the hope that better diffracting crystals could be obtained. We hoped that given the relatively low level of sequence identity between TraM of F and pED208, the crystallization properties of pED208 would be significantly different from F, but that the overall mode of binding of TraM to *sbmA* for the F-like plasmid family could still be obtained from the crystal structure.

The minimal length of *sbmA* required for full binding affinity of pED208 TraM was determined by testing the ability of TraM to bind to *sbmA* sequences of varying lengths (Figure 4-3C) was determined by the same EMSA protocol used to describe the stoichiometry, with binding components at concentrations far exceeding the  $K_d$ . Since there was visible free DNA with a 22 bp *sbmA* at greater TraM:DNA ratios than 8:1, but not in oligos 24 bp in length or greater, the minimal length was found to be 24 base pairs. This sequence is centered on the palindromic centre (Figure 4-3A). The 24 bp length includes one more base pair beyond the outer GANTC motifs –deletion of the flanking base pairs to a 22 bp fragment disrupts the clean shift of unbound DNA at the 8:1 TraM monomer:*sbmA* ratio (Figure 4-3A). 23bp B shifts unbound DNA cleanly at 8:1 TraM monomers:*sbmA*, whereas 23bp A does not. Their different apparent binding affinities indicated that the flanking nucleotides on either side of the GANTC motifs have asymmetric contributions to TraM binding. The flanking base

pair which is most important for full TraM binding affinity is in 23bp B (Figure 4-3B).

#### *Screening of crystallization conditions for the pED208 TraM-sbmA complex*

Crystals of pED208 TraM in complex with 26 bp *sbmA* were obtained by vapour diffusion in hanging drops at room temperature. An initial hit was obtained in condition #17 of the JCSG+ (Qiagen) screen (40% 2-methyl-2,4-pentanediol (MPD), 5% PEG8000, 100 mM sodium cacodylate pH 6.5) with 1+1  $\mu$ l drops. Many crystallization variables were attempted to improve the quality of the crystals. These included varying the precipitant concentration, pH, buffer type, and PEG type. Additives (5 mM  $MgCl_2$ , 2 mM spermine, 100 mM KCl or NaCl), varying DNA length and overhangs (Table 4-1), temperatures (4°C, 15°C), TraM monomer:*sbmA* ratio 8:1-5:1, and ratio between the two precipitants were also varied. The variables that yielded significant improvement in crystal size and quality were changing the PEG length and decreasing the concentration of the precipitants, and increasing the drop size. The optimal ratio of protein:DNA was 6:1, and optimal length of DNA was 24 base pairs centered around the palindromic centre. The best crystals were obtained in 36% MPD, 4.5% PEG2000, 100 mM cacodylic acid pH 6.5, with 2+2  $\mu$ l drops grown at room temperature (22°C). The resulting crystals had varying morphologies and ranged in size from 0.3-0.6  $\mu$ m (Figure 4-4A). The mother liquor of the crystals is already an adequate cryoprotectant so these crystals were briefly washed in mother liquor from the harvested well prior to flash-freezing in liquid nitrogen.

These crystals were fragile and would easily crack or flake during handling for freezing. These crystals also typically did not diffract past 5 Å at the

home x-ray source, and frequently suffered from poor diffraction spot quality. The highest resolution data set of this crystal form was collected to 2.9 Å using remote data collection at Beamline 12.3.1 at the Advanced Light Source, Berkeley (Table 4-2), after screening ~140 crystals. The space group of this crystal form was C222<sub>1</sub>, confirmed by the presence of systematic absences along the [0 0 l] axis. Crystals were originally obtained using the wild-type *sbmA* sequence, but the final crystal used for structure solution was grown with the perfectly palindromic variant of the 24 base pair sequence, 24BF-24BR (Table 4-1). Eventually during the structure solution process, we realized that the *sbmA* fragment was centered along a crystallographic 2-fold axis and the palindromic variant eliminated hybrid electron density for bases which were not the same at the equivalent position around the 2-fold axis.

A second hit was obtained from crystallization screening for the pED208 TraM-*sbmA* complex, in condition #10 of the Matrix Screen (5% PEG4000, 5 mM MgSO<sub>4</sub>, 50 mM MES pH 6.0) with 1+1 µl drops at room temperature (22°C). These crystals did not require any additional optimization to yield crystals with high-resolution diffraction and very good diffraction spot quality, despite the very small size of the crystals (<0.05 µm) (Figure 4-4B). They were also quite resistant to damage from handling and transferring out of the drops they were grown in. These crystals were soaked in mother liquor plus 20% glycerol for 10 minutes for cryoprotection prior to flash-freezing in liquid nitrogen. The highest resolution data set of this crystal form was collected to 1.3 Å on-site at Beamline 8.3.1 at the Advanced Light Source (Table 4-2), which indexed to either P4<sub>1</sub>2<sub>1</sub>2 or P4<sub>3</sub>2<sub>1</sub>2 based on systematic absences on the [0 0 l] and [0 k 0] axes.

The second crystal form was shown to contain only the N-terminal domain of TraM residues 2-52 by MALDI mass spectrometry (Figure 4-5A and 4-5B), resulting from proteolysis of the full-length protein during crystallization.

*Structure solution of the pED208 TraM-sbmA complex and apo-N-terminal domain*

The only part of the TraM-*sbmA* structure which was straightforward to solve by molecular replacement is the C-terminal tetramerization domain. Molecular replacement with a model of the N-terminal domain based on the MetJ ribbon-helix-helix domain or with an ideal B-DNA model of the 24bp *sbmA* fragment failed to yield a solution. Selenomethionine-derivative TraM was expressed for obtaining additional phases by MAD methods. However, the Se-Met TraM failed to bind to *sbmA* (data not shown). Individual methionine residues were mutated to leucine and expressed as selenomethionine-derivative proteins in case only one of the Se-Mets were problematic. However, none of the Se-Met mutants were able to bind to *sbmA* (data not shown). Co-crystallization of native TraM with 5'-iodo-2'-deoxyuridine (dIU) substitutions for various thymines in the *sbmA* sequence was then attempted. Most of the iodinated bases resulted in loss of binding of TraM to the DNA. A few of the substitutions did not result in noticeable loss of binding, and were crystallized with TraM. dIU at base number 4 of the 24 base pair *sbmA* fragment gave the best crystals. A dataset was collected from one of these crystals to 3.1Å (Table 4-3). No solution was found using SOLVE and a difference Patterson map did not yield convincing peaks for the iodine (data not shown). This could be due to the mixed occupancy at the dIU4 position due to substitution in only one of the DNA strands and the presence of a two-fold axis down the center of the DNA. However, a peak for the iodine (height = 9.94

$\sigma$ ) was clearly visible in a difference fourier map created after the tetramer was solved, and its location matched up very well with where it would be expected in the substituted uridine (Figure 4-6A). Nevertheless, the rest of the TraM-*sbmA* complex was solved without any additional experimental phases.

An overview of the structure solution process is shown in (Figure 4-7A). The asymmetric unit of the TraM-*sbmA* complex contains 1 tetramer, with the biological unit formed by a 2-fold crystallographic axis. The  $\alpha$ -helices of the N-terminal domain were built into density visible after restrained refinement with NCS restraints between the four chains of the tetramer. The resulting arrangement of helices confirmed that the N-terminal domain was a ribbon-helix-helix fold, and was eventually replaced with a model of the RHH-domain based on MetJ.

The apo-N-terminal domain crystal form was initially not solvable by molecular replacement with models derived from CopG or MetJ. No heavy atom derivative crystals from heavy atom soaks with defined heavy atom sites were found. The partially refined N-terminal domain monomer containing residues #2-47 was successfully used as a model in molecular replacement when the data was scaled as P4<sub>1</sub>2<sub>1</sub>2. Two monomers of the TraM N-terminal domain were in the asymmetric unit, with crystallographic 2-fold axes creating the biological units. The residues and ordered waters were then built using an automated model building program Arp/wArp (Cohen, *et al.*, 2008) and the phases from molecular replacement. The apo-N-terminal domain structure was eventually refined with alternate conformations at partial occupancies for certain side

chains, as well as additional co-factors ( $Mg^{2+}$ , glycerol) to an  $R_{work}$  and  $R_{free}$  of 15.1% and 17.3%.

The partially refined apo-N-terminal domain structure was then used in place of the MetJ-based model in the TraM-*sbmA* structure. The availability of a high-resolution structure was particularly helpful in building the loop between the  $\alpha 1$  and  $\alpha 2$  helices, as it was not in the correct conformation in the MetJ-based model. Manual placement of the ideal B-DNA model of *sbmA* into density showed that it deviated quite significantly from experimental density –simulated annealing (DNA restraints) refinement as well as manual model building was required to refine the DNA conformation. The initial model was a 12 bp double-stranded model, as a 2-fold crystallographic axis runs perpendicular to the DNA. Following the first round of simulated annealing refinement, the DNA backbone skipped over density corresponding to a nucleotide, so the nucleotides were manually moved followed by more simulated annealing refinement. The 12 bp double-stranded model was converted in the final stages of DNA refinement to a 24 bp single-stranded model in order to join the DNA backbone. The loops joining the N- and C-terminal domains were then built manually. The final  $R_{work}$  and  $R_{free}$  of the structure was 25.0 and 27.8% respectively (Table 4-2). A prime-and-switch model bias-reduced electron density map for the protein-DNA interface is shown in (Figure 4-7B).

#### *Overall Structure of the pED208 TraM-sbmA complex*

The structure of the TraM-*sbmA* complex reveals that two TraM tetramers bind *sbmA* (Figure 4-8A), validating the MALLS and EMSA results (Figure 4-1B, 4-1C, 4-1D). As predicted by the specificity swap experiments, the N-terminal

regions adopt dimeric RHH folds that each contact one of the GANTC sites within *sbmA*. The two tetramers contact *sbmA* in a staggered arrangement in which the two N-terminal domains of one tetramer bind the 1<sup>st</sup> and 3<sup>rd</sup> GANTC motifs, while those of the second tetramer bind the 2<sup>nd</sup> and 4<sup>th</sup> GANTC repeats (Figure 4-8A). In this way the two tetramers are arranged on nearly opposite sides of the DNA. Each RHH domain is connected to the C-terminal tetramerization domain via flexible peptide linkers corresponding to residues 56-60. The end of  $\alpha 2$  preceding the linker is unwound to different degrees in different protomers (4-8B), providing additional conformational flexibility to the tetramer which may facilitate TraM-DNA binding. The RHH and tetramerization domains do not contact each other, suggesting that these domains are flexibly tethered to one another and otherwise do not interact.

#### *N-terminal domain structure and TraM-sbmA interactions*

The N-terminal strands of each RHH domain form a two-stranded antiparallel  $\beta$ -sheet that provides amino acid side chains, Lys3, Gln5, and Tyr7, that enter the DNA major groove and contact the pseudo-palindromic GANTC repeat in a sequence-specific and symmetric manner (Figure 4-8C and 4-8D). The conserved A-T base pair is recognized by Gln5, which makes a pair of hydrogen bonds with the face of the adenine base, an interaction commonly observed in protein-nucleic acid interactions (Luscombe, *et al.*, 2001). The orientation of the Gln5 side chain is stabilized in the correct orientation through an additional hydrogen bond to Lys3 in the other  $\beta$ -strand. The conserved G-C pair is recognized by Tyr7. Tyr7 also forms a hydrogen bond with Lys3, which likely helps to position this side chain and ensure that the terminal hydroxyl group is oriented to donate a hydrogen bond to the guanine N7 atom. Lys3 may also

directly contribute to recognition of the guanine through a long hydrogen bond between Lys3 Nε and thymine O4. The RHH domain also makes strong, symmetric interactions with the DNA backbone on either side of the GANTC motif. The interaction involves the N-terminus of α2, which is aligned so that its helix dipole, the main chain NH of Leu33, and the hydroxyl groups of Ser32 and Ser34, interact with the DNA phosphate backbone of the base preceding the G of the GANTC motif, Ade1 (Figure 4-8C and 4-8D). This explains why the *sbmA* construct 23bp B is able to bind to TraM with the same apparent affinity as 24 bp, whereas 23bp A cannot (Fig. 4-2)

A comparison of the structure of the DNA-bound form of TraM and the free TraM RHH domains reveals subtle conformational change induced by DNA binding (Figure 4-9A). The α1-α2 loop comes into closer contact with the DNA backbone (Figure 4-9B), and β1 DNA-contacting side chains are stabilized into a single conformation by a hydrogen bond network (Figure 4-9C and 4-9D) upon DNA binding. Overall, the structure is very similar to that of the bound form, with an r.m.s.d. of 0.721.

*TraM-DNA interactions are stabilized by cooperative DNA unwinding and distortion*

The crystal structure of the TraM-*sbmA* complex reveals that binding of *sbmA* is achieved without direct contact between the two TraM tetramers, suggesting that the mechanism of cooperative DNA binding must act through the DNA itself (Figure 4-11). Analysis of the DNA structure reveals that it is significantly distorted compared to standard B-DNA. Strikingly, the DNA is significantly underwound with an average helical twist of 32° between staggered GANTC sites, compared to 36° per base pair in standard B-DNA. This unwinding

results in a duplex structure with ~12 bp per turn, which aligns alternating GANTC sites on the same side of the DNA, and facilitates recognition of the alternating GANTC sites by RHH domains from a single tetramer. The binding of one tetramer to one face of the *sbmA* DNA would thus unwind and align the unbound GANTC sites on the opposite side of the DNA such that they would be positioned to interact with the second TraM tetramer, thereby facilitating cooperative recognition of *sbmA*.

pED208 TraM binds its cognate *sbmA* with high affinity ( $K_d = 4.8$  nM, Figure 4-10A), similar to the affinity of F TraM for its cognate *sbmA* site (Fekete & Frost, 2002). Tetramerization is essential for high-affinity binding, as the isolated N-terminal domain binds with ~500 times lower affinity (Figure 4-10B).

We reasoned that if DNA unwinding is important for cooperativity, then binding of a single TraM tetramer to DNA containing a pair of GANTC repeats could be achieved by reducing the spacing between the repeats to better match the helical pitch of B-DNA (Figure 4-11A and 4-11B). To test this idea, we compared the ability of pED208 TraM to bind a double-stranded DNA containing only a single pair of GANTC sites separated by 12 bp (as in *sbmA*), with DNAs in which the pair of GANTC sites are separated by either 11 or 10 bp using EMSA (Figure 4-11E). Interaction between TraM and DNA with the 12 bp spacing was weak. However, reduction of the spacing to 11 bp significantly enhanced TraM binding. Further reduction of the spacing to 10 bp essentially abrogated binding. Modeling suggests that docking of two RHH modules on GANTC sites separated by 10 bp on a B-form double helix would lead to significant clashes between the  $\alpha 1$  helices and  $\alpha 1$ - $\alpha 2$  loops, explaining the loss of binding with this DNA. While

these experiments support the importance of the positioning of the GANTC sites on the same side of the DNA double helix for TraM binding, the fact that the binding of the 11 bp spacer DNA is still much weaker than the intact *sbmA* (Figure 4-10A) indicates that other mechanisms must also facilitate cooperative TraM-DNA interactions.

Further comparison of the *sbmA* structure with that of ideal B-DNA reveals significant kinking of the DNA helix axis, mainly localized to the junctions between the GANTC repeats (Figure 4-11C and 4-11D). This kinking appears to be largely due to interactions between TraM and the DNA backbone. As noted above, the N-termini of the  $\alpha 2$  helices bind the phosphate backbone, anchoring the RHH domain to both sides of the DNA major groove, facilitating a conformational change that results in the unwinding of the DNA within the GANTC site, as well as widening and deepening of the major groove. At the same time, a pair of acidic residues, Glu29 and Glu30, is positioned in the loop between  $\alpha 1$  and  $\alpha 2$  to make unfavorable electrostatic interactions with the DNA helical backbone. In response, the DNA bends into the major groove to minimize these interactions (Figure 4-11C and 4-11D). This push of the backbone away from one RHH helps to wrap the next GANTC site around its RHH domain.

## **Discussion**

*TraM is a member of the RHH family of plasmid regulatory proteins*

Our work reveals that TraM is a member of the large class of bacterial repressor proteins that bind DNA through a ribbon-helix-helix (RHH) DNA recognition module (Schreiter & Drennan, 2007). RHH modules appear to be a common DNA recognition module among plasmid regulatory proteins. For

example, F-like plasmids encode at least one other RHH DNA binding protein, TraY, which is critical for the activation of plasmid nicking through interactions with the nickase/relaxase, Tral (Inamoto, *et al.*, 1994, Howard, *et al.*, 1995, Karl, *et al.*, 2001). The R388 plasmid encodes a single RHH protein, TrwA, which likely fulfills both TraY and TraM functions in that it appears to interact with and activate both the TraD-like coupling protein, TrwB (Llosa, *et al.*, 2003, Tato, *et al.*, 2007), as well as the nickase/relaxase, TrwC (Moncalian & de la Cruz, 2004). The smallest of the plasmid regulatory RHH proteins is CopG, which regulates copy number of the streptococcal pMV158 plasmid (Gomis-Ruth, *et al.*, 1998). Particularly intriguing is the *Agrobacterium tumefaciens* T DNA regulatory protein, VirC2, which binds DNA through a novel, tandem repeat RHH module (Lu, *et al.*, 2009). In this case, the RHH dimer is formed from a single polypeptide chain, where the loop connecting the two RHH motifs wraps around the structure to form a topological knot (Bolinger, *et al.*, 2010).

#### *Cooperative DNA binding is mediated through DNA distortion*

The regulation of bacterial transcription often involves the formation of hierarchical assemblies of oligomeric DNA binding proteins on large, complex DNA elements. In general, these higher order interactions are mediated by protein-protein interactions, such as the interactions between dimers of  $\lambda$  repressor bound to adjacent DNA elements (Stayrook, *et al.*, 2008), or the interactions of tetrameric TgtV repressor with DNA containing two adjacent recognition elements (Lu, *et al.*, 2010). Thus, it was highly surprising that pairs of TraM tetramers cooperatively bind the high affinity *sbmA* DNA without direct contact between the tetramers. Instead, cooperativity is mediated by a protein-induced distortion of the DNA that facilitates simultaneous binding of both

tetramers. TraM binding induces an unwinding of the DNA to a ~12 bp/turn form that aligns alternating GANTC motifs on the same side of the DNA double helix for recognition by two RHH domains presented from a single tetramer (Figure 4-8 A). In addition, the protein also kinks the DNA and deforms the groove widths, through a combination of attractive electrostatic and hydrogen-bonding interactions, as well as repulsive interactions between negatively charged residues in the  $\alpha$ 1- $\alpha$ 2 loop and the DNA phosphodiester backbone (Figure 14-11 C and D). The fact that the 12 bp spacing of sequence elements is conserved in the *sbm* sites of other plasmids suggests that similar mechanisms will mediate DNA binding cooperativity in other TraM proteins (Figure 4-1B). Our demonstration that a hybrid pED208-F *sbmA* site in which DNA binding motifs alternate between pED208 and F, is only bound with high affinity by a mixture of pED208 TraM and F TraM, and not by the individual proteins, strongly suggests that the DNA deformations that facilitate DNA binding cooperativity are conserved throughout the TraM family (Figure 4-11F). We suggest that cooperative DNA recognition by TraM tetramers proceeds through the model outlined in (Figure 4-12). Initially, a single tetramer binds DNA, inducing the underwound and kinked DNA conformation in an unstable, high energy intermediate state. The DNA is thus primed for the binding of the second tetramer to the opposite face of the DNA, thereby stabilizing the cooperative TraM-DNA complex.

Interestingly, similar mechanisms of protein-induced DNA distortions have been uncovered in other bacterial transcriptional repressors. For example, three repressor proteins in the TetR family, QacR, IcaR, and CgmR, bind DNA sites in a staggered arrangement in which the DNA is underwound and kinked

(Schumacher, *et al.*, 2002, Jeng, *et al.*, 2008, Ito, *et al.*, 2010). Similarly, members of the iron-dependent regulator family (IdeR and DtxR), which are activated by the binding of divalent metal ions, also bind DNA in a similar manner, utilizing DNA deformation and not protein-protein interactions to effect cooperativity (White, *et al.*, 1998, Pohl, *et al.*, 1999, Pohl, *et al.*, 1999, Wisedchaisri, *et al.*, 2004).

## **Materials and Methods**

### *Growth media and bacterial strains*

Media and antibiotics for bacterial growth and bacterial strains used are described in (Wong, *et al.*, 2011).

Cells were grown in LB (Luria-Bertani) broth or on LB solid medium containing appropriate antibiotics or other supplements. Antibiotics were used at the following final concentrations: ampicillin (Amp), 50 µg/mL; kanamycin (Km), 25 µg/mL; spectinomycin (Spc), 100 µg/mL; nalidixic acid (Nal), 40 µg/mL, and tetracycline (Tet), 10 µg/ml. IPTG (isopropylthio-β-D-galactoside) was used at a final concentration of 1 mM. The following *Escherichia coli* strains were used: XK1200 [F<sup>-</sup> Δ*lacU124* Δ(*nadA gal attλ bio*) *gyrA* (Nal<sup>r</sup>)] (Moore, *et al.*, 1987), ED24 (F<sup>-</sup> Lac<sup>-</sup>Spc<sup>r</sup>) (Willettts & Finnegan, 1970), DH5α [Δ*lacU169* (Φ80 *lacZ*Δ*M15*) *supE44 hsdR17 recA1 endA1 gyrA96* (Nal<sup>r</sup>) *thi-1 relA1*] (Hanahan, 1983) and BL21-DE3 [F<sup>-</sup> *ompT hsdS* (r<sub>B</sub><sup>-</sup>m<sub>B</sub>) *galλ*(DE3)] (Stratagene).

### *Primers, Plasmids, and cloning of TraM mutants*

Primers, plasmids, and cloning of all TraM constructs used are described in (Wong, *et al.*, 2011) and Table 4-4.

Plasmids pJLM400 (Lu, *et al.*, 2004), pJLM403 (Lu, *et al.*, 2004), pED208 (Falkow & Baron, 1962), pRFM200 (Lu, *et al.*, 2003), and pT7-5 (Tabor & Richardson, 1985) have been described previously. The 0.5-kb EcoRI-BamHI fragment of the PCR products amplified from pED208 using primer pair JLU91 was cloned into the EcoRI-BamHI sites of pT7-5 or pBluescript KS (+), resulting in pJLEM200 or pJLM404, respectively. The 0.5-kb EcoRI-BamHI fragment of the PCR products amplified from pRF105 using primer pair JLU601 and JLU602, or JLU603 and JLU602, was cloned into the EcoRI-BamHI sites of pJLM400, resulting in pJLM401 or pJLM402, respectively.

#### *Overexpression and purification of F TraM, R100 TraM, and R100 R3K:I5N TraM*

DH5 $\alpha$  cells containing pJLM400, pJLM401, or pJLM402 were grown in 100 mL of LB broth containing ampicillin and 0.4% glucose at 37°C. After 12 hours, the 100 mL culture was added into 1 L of LB with ampicillin and grown under the same conditions for 3 hours. IPTG was added to a final concentration of 1 mM and the culture was grown for another 3 hours before harvesting. All the following steps were performed at 4°C or on ice. The cell pellet was suspended in 75 mL of buffer containing 50mM Tris-HCl, pH7.5 and one tablet of Complete protease inhibitor cocktail (Roche), and sonicated for 30 seconds, 6 times. Cell debris was removed by centrifugation at 39,000 x *g* for 1 hour. Ammonium sulfate (20 grams) was dissolved in the extracted supernatant. After centrifugation at 39,000 x *g* for 1 hour, the supernatant was loaded onto a 25 mL hydrophobic interaction chromatography column (Phenyl Sepharose<sup>TM</sup> 6 Fast Flow, GE healthcare Life Sciences), eluted with Tris-HCl (50 mM, pH 7.5) and a 1 to 0 M ammonium sulfate gradient. The fractions containing TraM (~40 mL) was brought to 135 mL with malonic acid (50 mM, pH 5.5), loaded onto a 25 mL

cation-exchange column (SP Sepharose™ Fast Flow, GE healthcare Life Sciences), and was eluted with malonic acid (50 mM, pH 5.5) and a 0 to 1 M NaCl gradient. The pooled TraM fractions were concentrated to 5 mL using an Amicon® ultracentrifuge filter (Millipore), loaded onto a size exclusion column (Hiload® 26/60 Superdex 75 prep grade, GE Healthcare Life Sciences), and was eluted with 50mM sodium phosphate, 150 mM NaCl, pH 7.5. Fractions containing purified TraM were concentrated and the buffer was exchanged for 0.5 M ammonium acetate using an Amicon® ultracentrifuge filter (Millipore) to a final volume of 1 mL. Protein concentration was determined using BCA protein assays (Pierce) following the manufacturer's instructions.

#### *Expression and purification of pED208 TraM*

BL21-DE3 cells containing pED208 expressing cells were grown in LB media containing 100ug/mL ampicillin at 37°C until the OD 600 reached ~0.7, followed by induction with 0.5 mM IPTG for 5 hours at 27°C. Harvested cells from 1 L of media were suspended in 100 mL of 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT containing 1 Complete, EDTA-free protease inhibitor tablet (Roche). The cells were lysed by sonication for 20 seconds 6 times, followed by centrifugation at 39,000 x g for 45 minutes to remove cell debris. 35 grams of ammonium sulfate was added to the clarified cell lysate, and dissolved by stirring for 15 minutes on ice. Precipitated proteins were removed by centrifugation at 39,000 x g for 30 minutes. 10 grams of ammonium sulfate was dissolved in the resulting supernatant, followed by centrifugation as in the previous step. The resulting pellet containing TraM was dissolved in 100 mL of 50 mM Tris, pH 7.5, 1 M ammonium sulfate. The dissolved pellet was loaded onto a 25 mL HIC

column (Phenyl Sepharose 6 Fast Flow™, GE Healthcare Life Sciences) and eluted by a decreasing gradient of 50 mM Tris pH 7.5, 1 M ammonium sulfate. Ammonium sulfate was added to fractions containing TraM to 45 grams per 100 mL and dissolved by stirring on ice for 15 minutes, followed by centrifugation at 39,000 x g for 30 minutes to precipitate TraM. The resulting pellet was dissolved in 100 mL of 50 mM L-Histidine, pH 5.5, and loaded onto a 25 mL anion exchange column (Q Sepharose Fast Flow™, GE Healthcare Life Sciences) and eluted by increasing gradient of 50 mM L-Histidine, pH 5.5, 1 M NaCl. TraM-containing fractions were concentrated and exchanged into 50 mM MES pH 6.8, 300 mM NaCl by centrifugation at 2,400 x g in 15 mL concentrators (Amicon Ultracel 10K, Millipore). The concentrated fraction was loaded onto a 300 mL size exclusion column (HiLoad 26/60 Superdex75™ prep grade, GE Healthcare Life Sciences) and eluted with the same buffer. TraM-containing fractions were concentrated in 15 mL concentrators (Amicon Ultracel 10K, Millipore) and exchanged into 0.5 M ammonium acetate. Protein concentration was determined using BCA protein assays (Pierce) following the manufacturer's instructions.

*Oligonucleotide DNA purification and annealing for crystallization, titration analysis, and MALLS*

Synthetic DNA oligonucleotides were purified on a 7 mL anion exchange column (Source 15Q) under denaturing conditions (10 mM NaOH) and eluted by increasing gradient of 10mM NaOH, 1M NaCl. DNA-containing fractions were desalted with a Sep-PAK® (C18) cartridge in a volatile buffer (30% acetonitrile in 0.1 M triethylammonium bicarbonate), lyophilized and resuspended in water. Oligonucleotide DNA solutions were quantified by absorbance at 260 nm. Mixed oligos were heated in a heating block to 95°C for 15 minutes and slow cooled to

room temperature in 10 mM Tris pH 7.5 and 100 mM NaCl at a final concentration of 0.5 mM double stranded DNA.

#### *Multi-angle laser light scattering*

All following steps were performed at 25 °C. Purified TraM (200 µg) or purified TraM (120 µg) plus 200 µg of 30-base pair *sbmA* DNA was applied to a Superose™ 12 HR 10/30 column eluted with the same buffer. The effluent from this column was run directly through in-line DAWN EOS™ multi-angle laser light scattering (MALLS) and Optilab rEX™ differential refractive index detectors (Wyatt Technologies, Santa Barbara, CA). ASTRA v.4.90 software was used to process the data.

#### *Titration analysis of pED208 TraM binding to sbmA*

Each binding mixture contained 1.5 µM of 24 base-pair *sbmA*, the indicated molar ratio of TraM, 50mM Tris-HCl (pH 7.5), and 10% glycerol in a 15 µl volume. Samples were incubated for 10 minutes at room temperature. Mixtures were run on TBE-buffered 12% polyacrylamide gel in 1X TBE at 4°C and 200 Volts for 45 minutes. DNA and DNA-protein complexes were visualized by ethidium bromide staining.

#### *Titration analysis of F TraM-sbmA binding*

Each binding mixture contained 1 µM of 30 bp *sbmA*, 50 mM Tris-HCl (pH 7.5), and 10% glycerol. Increasing amount of purified F plasmid TraM was added with different molar ratios to the 30 bp *sbmA*. The final volume of each sample is 10 µl. The resulting mixture was loaded onto a TBE-buffered 12% polyacrylamide gel and was run in 1x TBE (90 mM Tris-borate, 1 mM EDTA) at 4°C and 40 Volts

for 4 hours. DNA and DNA-protein complexes were visualized by ethidium bromide staining.

*K<sub>d</sub> determination of pED208 TraM binding to sbmA by electrophoretic mobility shift assay*

*sbmA* oligos were P<sup>32</sup>-labelled with T4 polynucleotide kinase (Invitrogen) and unincorporated nucleotides were removed by P-30 Micro Bio-Spin columns buffered in 10 mM Tris pH 7.4 (Bio-Rad). TraM-*sbmA* binding buffer was 50 mM Tris pH 7.5, 10% glycerol, 30 ng/μL bovine serum albumin (Pierce), 20 ng/μL polydI-dC (Roche). 0.1 nM of *sbmA* oligo was added to each binding reaction containing the indicated amount of TraM and incubated for 10 minutes at room temperature. TraM-*sbmA* mixtures were run on 1x TBE-buffered 12% 29:1 acrylamide gels at 200 Volts for 45 minutes at 4°C. Bands were visualized by phosphor screen and band intensities were determined with Imagequant. K<sub>d</sub> values were calculated by fitting binding curves to the equation  $y = a * x / (1 + (a * x))$ , where a is the K<sub>a</sub>, y is the ratio of bound to unbound DNA, and x is the protein concentration, using SigmaPlot (<http://www.sigmaplot.com>).

*Crystallization of the pED208 TraM-sbmA complex and TraM apo-N-terminal domain*

Crystals of pED208 TraM-*sbmA* were obtained by vapour diffusion in hanging drops. In order to avoid hybrid electron density of bases around a 2-fold crystallographic axis, the *sbmA* fragment used in crystallization for the final dataset was designed to be palindromic. This DNA fragment, 24BF-24BR, required changing only 2 base pairs and no disruption of the GANTC binding motifs (Table 4-1). No change in binding affinity resulted from this mutation (data not shown). TraM protein at 15 mg/mL in 0.5 M ammonium acetate was mixed

with 0.5 mM *sbmA* DNA in 10 mM Tris pH 7.5, 100 mM NaCl so that the ratio of TraM monomers to DNA was 6:1. 2  $\mu$ L of the protein-DNA mixture was mixed with 1 or 2  $\mu$ L of reservoir solution consisting of 100 mM cacodylic acid pH 6.5, 36% MPD, and 5% PEG2000. Crystals were flash-frozen in liquid nitrogen without any additional cryoprotectant. 1  $\mu$ L of the same TraM-*sbmA* mixture was mixed with 1  $\mu$ L of reservoir solution consisting of 50 mM MES pH 6.0, 5% PEG4000, and 5 mM MgSO<sub>4</sub>. Proteolytic fragments of TraM consisting of the N-terminal domain residues 2-52 crystallized under these conditions, as confirmed by MALDI mass spectrometry (Figure 4-5), in the space group P4<sub>1</sub>2<sub>1</sub>2. These crystals were soaked for 10 minutes in 20% glycerol in reservoir solution prior to flash freezing in liquid nitrogen.

#### *Structure solution*

Native datasets for both crystal forms were collected at ALS 12.3.1 and 8.3.1 at Lawrence Berkeley National Laboratory, Berkeley, California, and Canadian Light Source, Saskatoon. Data from the TraM-*sbmA* complex was collected to 2.90 Å, and from the TraM N-terminal domain to 1.30 Å. The data for the TraM-*sbmA* complex was a merged dataset from 2 original datasets, one at low exposure time and the other at high exposure time to maximize resolution while avoiding overloaded low resolution reflections. Diffraction images were processed with HKL2000 (Otwinowski, 1997).

Selenomethionine-substituted pED208 TraM did not bind DNA or crystallize, ruling out the possibility of determining the structure of the TraM-*sbmA* complex by Se-met MAD phasing. Initial phases were instead determined by molecular replacement utilizing the F TraM C-terminal domain (Lu, *et al.*,

2006) as a search model in MOLREP (Murshudov, *et al.*, 1997). The pED208 C-terminal tetramer model for molecular replacement was obtained by mutation of the F residues to that of pED208 using the “Mutate residue range” function in Coot (Emsley & Cowtan, 2004) followed by model minimization in CNS (Brunger, *et al.*, 1998).

Following refinement with NCS restraints on the 4 chains of the C-terminal tetramer, electron density for helices of the N-terminal domain was visible. The N-terminal domain was solved by manually building the N-terminal helices into visible electron density, followed by refinement with NCS restraints between the 4 copies of  $\alpha 1$  and  $\alpha 2$ . The helices were then replaced with a model based on the high-resolution structure of the MetJ ribbon-helix-helix domain (PDB ID: 1CMB) (Rafferty, *et al.*, 1989) and mutation of residues to poly-Ala with Coot (Emsley & Cowtan, 2004) and refined. The resulting electron density map revealed the DNA phosphate backbone and base pairs, into which a B-DNA model of the sbmA duplex was placed using Coot (Emsley & Cowtan, 2004). The TraM-sbmA model was further refined by cycles of manual building in Coot, followed by TLS and NCS-restrained maximum likelihood refinement and simulated annealing refinement, carried out with REFMAC (Murshudov, *et al.*, 1997) and CNS (Brunger, *et al.*, 1998), respectively.

The high-resolution structure of the N-terminal domain was phased by molecular replacement with MOLREP (Vagin & Teplyakov, 2010) using the the N-terminal domain structure residues #2-47 from the low-resolution structure as a search model. flex-wARP (Cohen, *et al.*, 2008) was used to build the residues of the N-terminal domain into electron density, and the model was refined with

REFMAC (Murshudov, *et al.*, 1997). The partially refined high-resolution pED208 N-term domain structure was used to replace the MetJ model in the TraM-*sbmA* complex, followed by further refinement. Residues of the linker connecting the N- and C-terminal domains were built manually into electron density to complete the TraM-*sbmA* complex structure. The refinement of the N-terminal domain high-resolution structure was completed with anisotropic B-factors in REFMAC (Murshudov, *et al.*, 1997).

The crystallographic asymmetric unit contains one TraM tetramer and one strand of *sbmA* DNA. The biologically relevant complex containing duplex *sbmA* and a pair of TraM tetramers is obtained by a crystallographic two-fold rotation. Figures were made in Pymol (<http://www.pymol.org>). DNA conformation was analyzed with 3DNA (Lu & Olson, 2008).

#### *Accession Numbers*

Coordinates and structure factors have been deposited in the Protein Data Bank (<http://www.rcsb.org/pdb>) under ID codes 3ON0 (TraM-*sbmA* complex) and 3OMY (TraM N-terminal domain).

#### *Donor ability assays*

*E. coli* XK1200 and ED24 were used as donor and recipient strains, respectively. The mating experiments were performed as previously described (Lu, *et al.*, 2002). Donor ability was calculated as the number of transconjugants divided by the number of donors. Each assay was repeated 3 times and the averaged values are reported. Standard deviations of all mating assays were within one log unit.

**Table 4-1 *sbmA* variants for TraM-*sbmA* crystal optimization**

<b>DNA combination</b>	<b>Oligo duplex</b>
23BF-24AR	AGATTCGAATCTGGATTCGAATC TCTAAGCTTAGACCTAAGCTTAGG
24AF-23BR	AGATTCGAATCTGGATTCGAATCC TCTAAGCTTAGACCTAAGCTTAG
24CF-24CR	AAGATTCGAATCTGGATTCGAATC CTAAGCTTAGACCTAAGCTTAGGG
25AF-25AR	AAGATTCGAATCTGGATTCGAATCC TTCTAAGCTTAGACCTAAGCTTAGG
25BF-25BR	AGATTCGAATCTGGATTCGAATCCC TCTAAGCTTAGACCTAAGCTTAGGG
25AF-25BR	AAGATTCGAATCTGGATTCGAATCC TCTAAGCTTAGACCTAAGCTTAGGG
25BF-25AR	AGATTCGAATCTGGATTCGAATCCC TTCTAAGCTTAGACCTAAGCTTAGG
26BF-25AR	AGATTCGAATCTGGATTCGAATCCCT TTCTAAGCTTAGACCTAAGCTTAGG
26BF-25BR	AGATTCGAATCTGGATTCGAATCCCT TCTAAGCTTAGACCTAAGCTTAGGG
25AF-26BR	AAGATTCGAATCTGGATTCGAATCC TCTAAGCTTAGACCTAAGCTTAGGGA
25BF-26BR	AGATTCGAATCTGGATTCGAATCCC TCTAAGCTTAGACCTAAGCTTAGGGA
24BF-24BR	AGATTCGAATCTAGATTCGAATCT TCTAAGCTTAGATCTAAGCTTAGA

<b>Table 4-2 Data Collection and Refinement for pED208 TraM crystals</b>		
	<b>TraM-<i>sbmA</i> complex</b>	<b>TraM N-terminal domain</b>
<b>Data Collection</b>		
Space Group	C222 <sub>1</sub>	P4 <sub>1</sub> 2 <sub>1</sub> 2
Cell Dimensions		
a, b, c (Å)	93.0, 154.7, 167.6	54.1, 54.1, 67.3
α, β, γ (°)	90, 90, 90	90, 90, 90
Wavelength (Å)	1.1158	1.1158
Resolution (Å)	2.90	1.30
R <sub>sym</sub> <sup>b</sup>	0.070 (0.433) <sup>a</sup>	0.044 (0.33)
I/σI	23.0 (3.0)	39.2 (6.6)
Redundancy	6.3 (3.5)	7.5 (7.4)
Completeness (%)	99.3 (94.2)	99.85 (100.0)
Mosaicity	0.729	0.396
<b>Refinement</b>		
Resolution (Å)	50.0-2.90	50.00-1.30
Number of unique reflections	27583 (2568)	25217 (2459)
R <sub>work</sub> <sup>c</sup> / R <sub>free</sub> <sup>d</sup> (%)	25.0 / 27.8	15.1 / 17.3
Number of protein atoms in asymmetric unit	4226	1284
B-factor (overall)	32.4	10.0
Bond angle r.m.s.d. (°)	1.35	1.38
Bond length r.m.s.d. (Å)	0.011	0.011
<sup>a</sup> Data of the highest resolution shell (2.90-3.00 Å for TraM- <i>sbmA</i> complex, 1.30-1.35Å for TraM N-terminal domain) are shown in parentheses		
<sup>b</sup> R <sub>sym</sub> = $\sum_{hkl} \sum_i  I_i(hkl) - \langle I(hkl) \rangle  / \sum_{hkl} \sum_i I_i(hkl)$ , where I <sub>i</sub> (hkl) is the intensity for an observation of a reflection and $\langle I(hkl) \rangle$ is the average intensity of all symmetry-related observations of a reflection		
<sup>c</sup> R <sub>work</sub> = $\sum_{hkl}    F_{obs}  -  F_{calc}    / \sum_{hkl}  F_{obs} $		
<sup>d</sup> R <sub>free</sub> = R <sub>work</sub> calculated for 5% of reflections excluded from refinement		

**Table 4-3 Data Collection for TraM-dIU4 *sbmA* complex**

<b>TraM-dIU4 <i>sbmA</i> complex</b>	
<b>Data Collection</b>	
Space Group	<b>C222<sub>1</sub></b>
Cell Dimensions	
a, b, c (Å)	<b>93.1, 154.9, 167.3</b>
α, β, γ (°)	<b>90, 90, 90</b>
Wavelength (Å)	<b>1.1159</b>
Resolution (Å)	<b>3.1</b>
R <sub>sym</sub> <sup>b</sup>	<b>0.066 (0.399)</b>
I/σI	<b>21.5 (2.4)</b>
Redundancy	<b>4.8 (4.4)</b>
Completeness (%)	<b>97.1 (79.5)</b>
Mosaicity	<b>0.901</b>
<sup>a</sup> Data of the highest resolution shell (3.21-3.10 Å) are shown in parentheses	
<sup>b</sup> R <sub>sym</sub> = $\sum_{hkl} \sum_i  I_i(hkl) - \langle I(hkl) \rangle  / \sum_{hkl} \sum_i I_i(hkl)$ , where I <sub>i</sub> (hkl) is the intensity for an observation of a reflection and $\langle I(hkl) \rangle$ is the average intensity of all symmetry-related observations of a reflection	

**Table 4-4** Plasmids and oligonucleotides

Plasmid & oligos	Description & references
pBAD24	Amp <sup>r</sup> ; (Guzman, <i>et al.</i> , 1995)
pBluescript	Amp <sup>r</sup> ; (Short, <i>et al.</i> , 1988)
pT7-5	Amp <sup>r</sup> ; (Tabor & Richardson, 1985)
pED208	(Falkow & Baron, 1962)
pJLEM200	pT7-5 with pED208 <i>traM</i> ; (Wong, <i>et al.</i> , 2011)
pJLM400	pBluescript KS (+) with F <i>traM</i> expressed from the lac promoter (Lu, <i>et al.</i> , 2004)
pJLM401	pJLM400 with <i>traM</i> from R100; (Wong, <i>et al.</i> , 2011)
pJLM402	pJLM400 with a R100 <i>traM</i> mutant R3K:I5N; (Wong, <i>et al.</i> , 2011)
pJLM403	pJLM400 with R100 <i>traM</i> N5D; (Lu, <i>et al.</i> , 2004)
pRF105	Amp <sup>r</sup> ; R100 <i>oriT</i> and <i>traM</i> cloned in pUC18; (Fekete & Frost, 2000)
pJLU91	TTA GAA T TC TAA TAA GGT TTT TGA AAT GCC
pJLU92	TTT CCC TAC CAC CAG AAC ATT CAA AGT G
pJLU601	TTG AAT TCG AAA GGT TTT ATC TTA TGG CCA GAG TAA TTT TGT ATA TCA G
pJLU602	TTG GAT CCG TGG TTA ATT GTC ATC AAA TTG AAC CAG ATC AAA ATC CTG
pJLU603	TTG AAT TCG AAA GGT TTT ATC TTA TGG CCA AAG TAA ATT TGT ATA TCA GTA ATG ATG TC
30BTA	GAT ACC GCT AGG GGC GCT GCT AGC GGT GCG
30BTA	CGC ACC GCT AGC AGC GCC CCT AGC GGT ATC
pED 4site	AGA TTC GAA TCT AGA TTC GAA TCT



Figure 4-1 TraM binding stoichiometry to *sbmA*

- a) SeW: *Salmonella enterica* subsp. *enterica* serovar Weltevreden str. HI\_N05-537. Ye8010; *Yersinia enterocolitica* str. 8010. Basic residues are highlighted in cyan, acidic residues in red, hydrophobic residues in orange, aromatic residues in magenta, polar aliphatic residues in yellow, and Gly and Pro in green. ClustalW consensus symbols for each residue are shown (“\*”: perfect identity; “.”: conserved substitutions; “.”: semi-conserved substitutions)
- b) *sbm* sequences from F and pED208 plasmids. Inverted repeats are indicated with orange and red arrows. The center of symmetry of the *sbm* site is indicated by a green oval.
- c) MALLS analysis of molecular weight of F TraM-30 bp *sbmA* complex and TraM-75 bp *sbmAB* complex. The traced peaks are the refractive index of the eluted material and the superimposed lines indicate the molecular weight of the protein or protein-DNA complex over the corresponding portion of the elution peaks.
- d) Binding stoichiometry of F TraM to 30bp *sbmA* analyzed by electrophoretic mobility shift assay.
- e) Binding stoichiometry of pED208 TraM to 24bp *sbmA* analyzed by electrophoretic mobility shift assay.

Work in c) and d) was done by Jun Lu

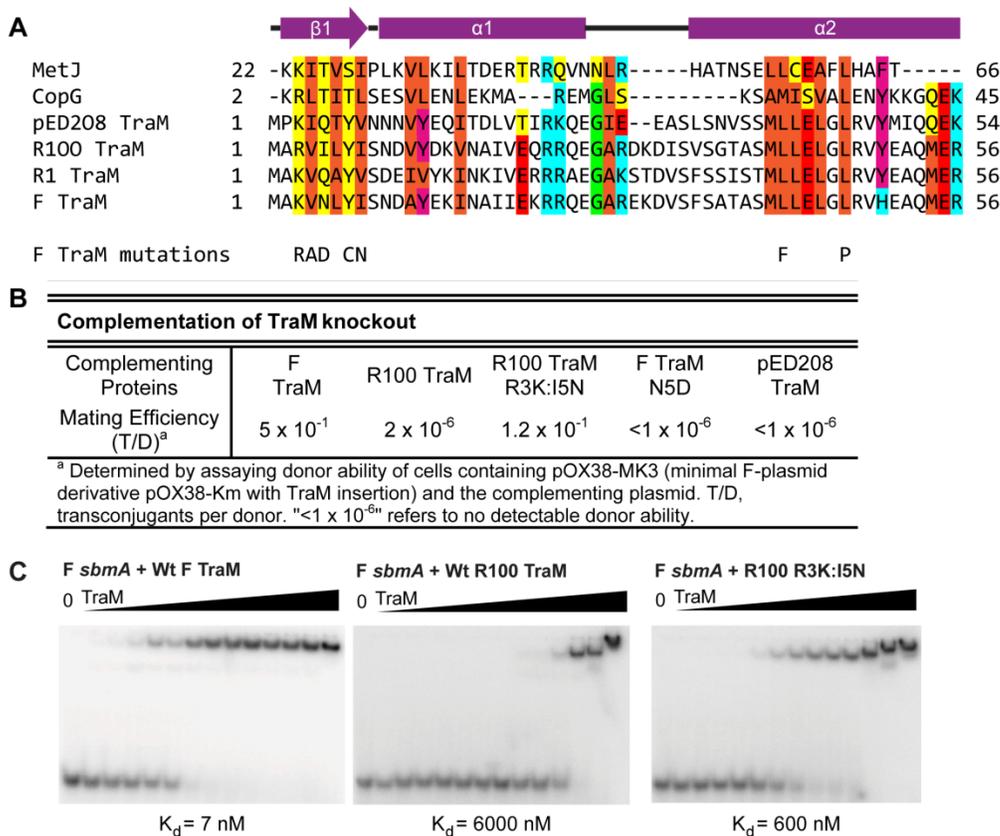


Figure 4-2 Evidence for the N-terminal domain of pED208 TraM being a ribbon-helix-helix fold

- Primary structure alignment of TraM homologues and ribbon-helix-helix fold domains. DNA-contacting  $\beta$ -strand residues are boxed in purple. Conserved residues are highlighted. Hydrophobic: orange. Aromatic: magenta. Acidic: red. Basic: cyan. Gly/Pro: green. Polar aliphatic: yellow. Secondary structure elements of TraM are indicated. F TraM mutations shown in previous studies to disrupt F plasmid conjugation are indicated. R24 and R29, F TraM residues protected from trypsin digestion upon binding to *sbmA*, are boxed in blue.
- Complementation of TraM proteins in a TraM-deficient F-derived plasmid system
- Electrophoretic mobility shift assay of F, R100, and R100 R3K:I5N TraM binding to 30 bp F *sbmA*. Concentrations in each lane are 0, 2, 5, 10, 20, 50, 100, 250, 600, 1000, 2500, 6000, 10000, and 25000 nM.

Work in b) was done by Jun Lu

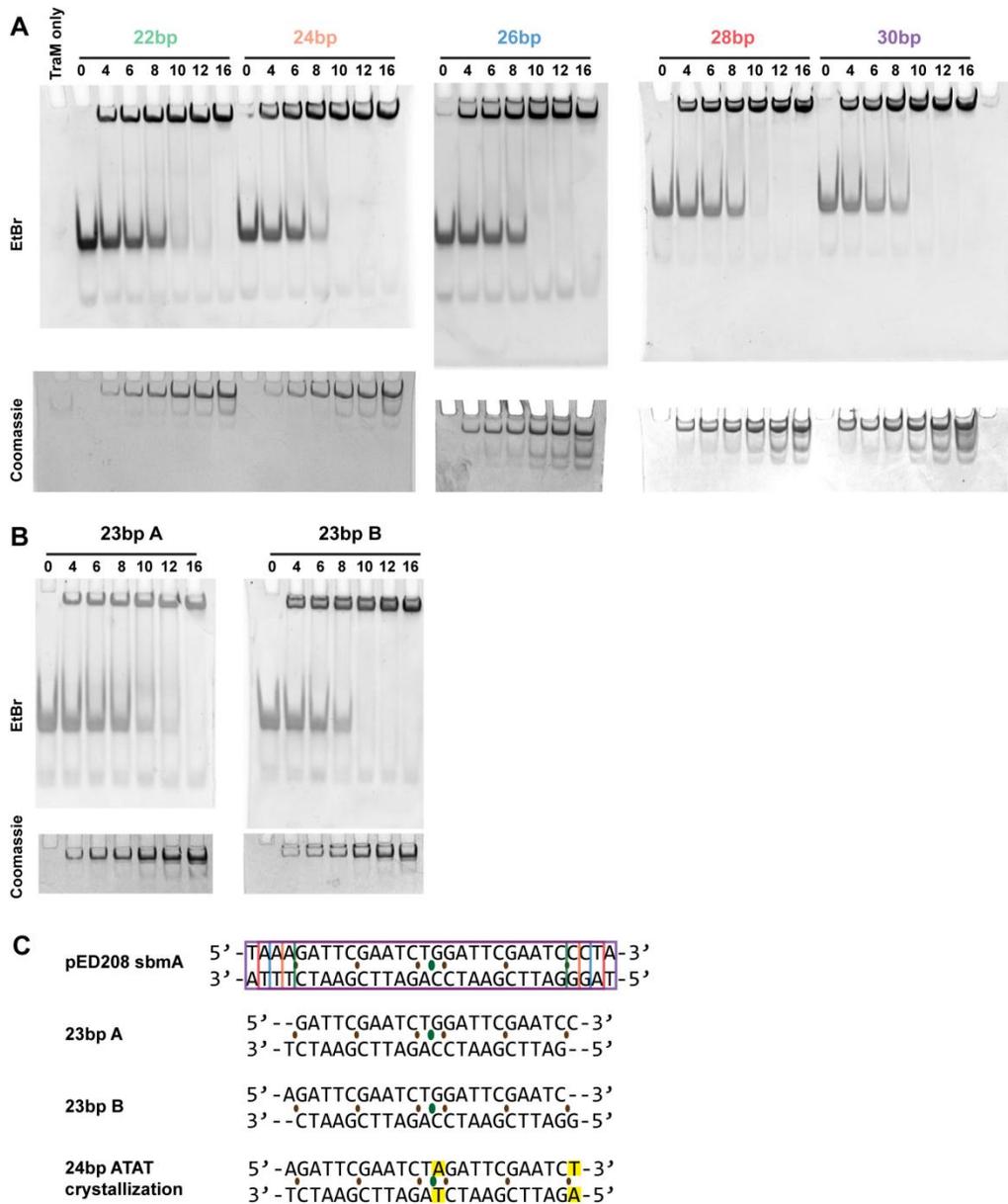


Figure 4-3 Minimal length of *sbmA* required for binding of pED208 TraM

- Relative binding affinity of pED208 TraM to *sbmA* of varying lengths analyzed by electrophoretic mobility shift assay.
- Role of the GANTC flanking nucleotides on pED208 TraM binding to *sbmA*
- pED208 *sbmA* sequences used in minimal TraM binding length determination and crystallization. Sequence lengths are indicated by boxes of the same colour scheme as a). Binding motif boundaries are indicated by brown ovals. The center of symmetry of the *sbm* site is indicated by a green oval. Mutated bases in 24bp ATAT to ensure crystallographic symmetry are boxed in yellow.

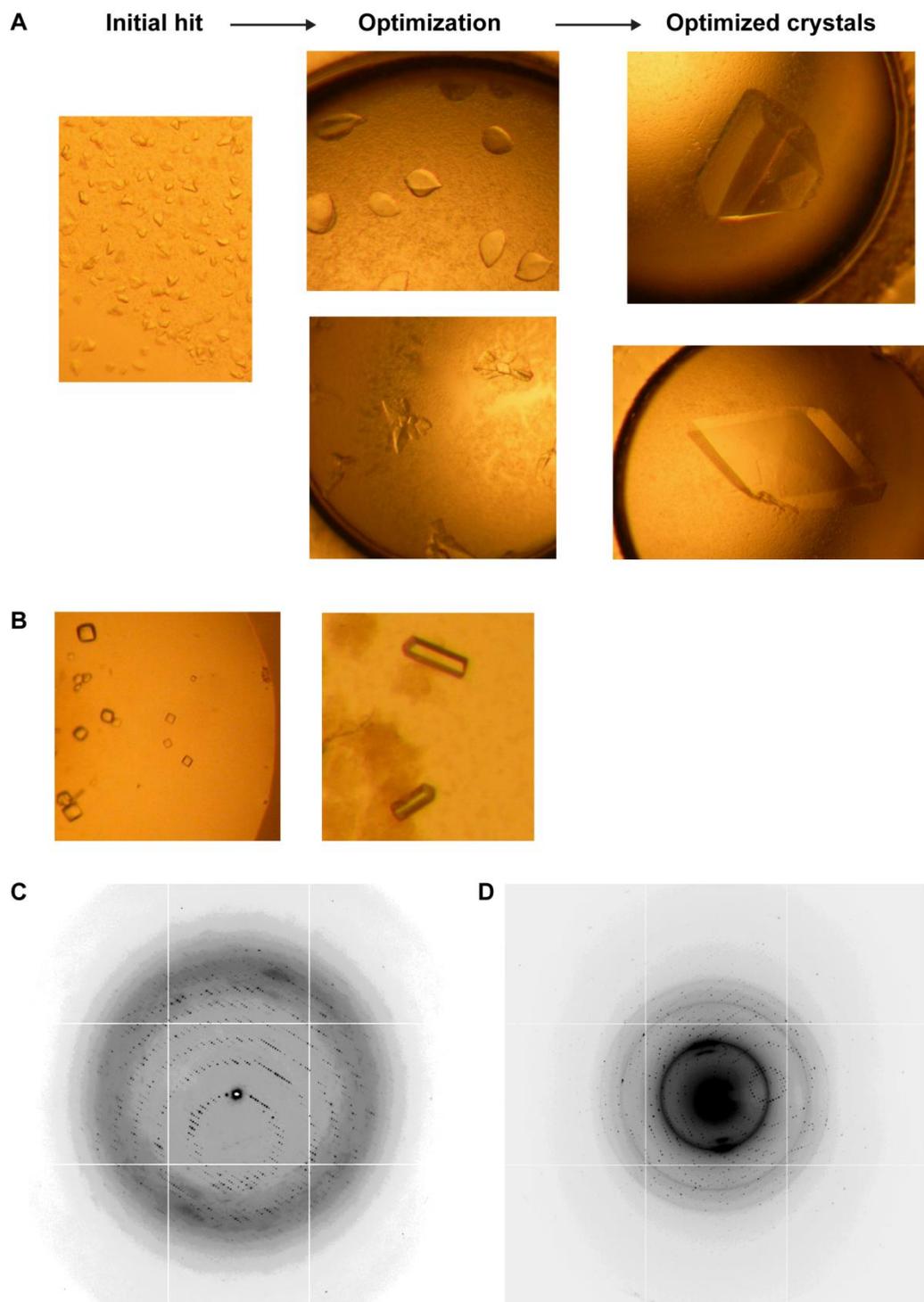
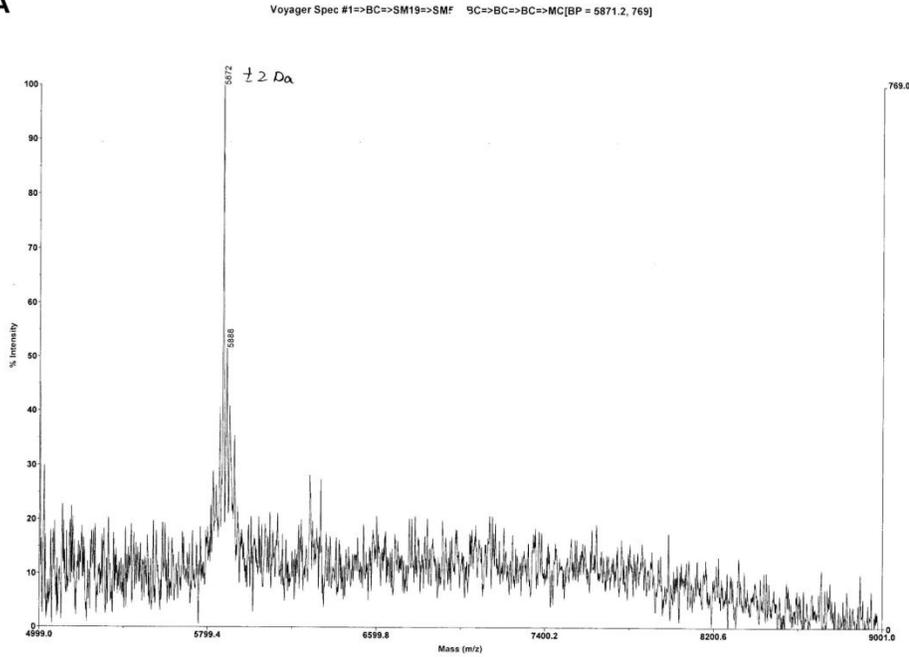


Figure 4-4 Crystals of pED208 TraM-*sbmA* complexes  
a) Optimization of pED208 TraM-*sbmA* crystals  
b) pED208 TraM apo-N-terminal domain #2-52 crystals  
c) pED208 TraM-*sbmA* diffraction image  
d) pED208 apo-N-terminal domain #2-52 diffraction image

**A****B**

Database: User Protein  
 Considered modifications: | Acetyl (Protein N-term) | Oxidation (M) |  
 Digest Used: Trypsin  
 Max. # Missed Cleavages: 5  
 User AA Formula 1: C2 H3 N1 O1  
 Minimum Digest Fragment Mass: 5000  
 Maximum Digest Fragment Mass: 7000  
 Minimum Digest Fragment Length: 5  
 Index Number: 1  
 pI of Protein: 4.7  
 Protein MW: 5871  
 Amino Acid Composition: A1 D1 E5 G2 I5 K2 L6 M2 N4 P1 Q5 R2 S4 T3 V5 Y3

1 PKIQTYVNNN VYEQITDLVT IRKQEGIEEA SLSNVSSMLL ELGLRVYMIQ Q

Number	m/z (mi)	m/z (av)	Modifications	Start	End	Missed Cleavages	Sequence
1	5105.6827	5108.8568		1	45	3	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLR(V)
1	5121.6776	5124.8562	1Oxidation	1	45	3	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLR(V)
1	5147.6933	5150.8944	1Acetyl	1	45	3	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLR(V)
1	5163.6882	5166.8937	1Acetyl 1Oxidation	1	45	3	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLR(V)
1	5642.9085	5646.4962		3	51	3	(K)IQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5658.9034	5662.4955	1Oxidation	3	51	3	(K)IQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5674.8983	5678.4949	2Oxidation	3	51	3	(K)IQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5868.0562	5871.7888		1	51	4	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5884.0511	5887.7882	1Oxidation	1	51	4	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5900.0460	5903.7876	2Oxidation	1	51	4	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5910.0667	5913.8264	1Acetyl	1	51	4	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5926.0617	5929.8257	1Acetyl 1Oxidation	1	51	4	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)
1	5942.0566	5945.8251	1Acetyl 2Oxidation	1	51	4	(-)PKIQTYVNNNVEQITDLVTIRKQEGIEEASLSNVSSMLLEGLRVYMIQ(-)

Figure 4-5 Mass spectrometry confirmation of protein fragment in apo-N-terminal domain #2-52 crystals

a) MALDI-TOF Mass spectrum of N-terminal domain #2-52 crystals

b) Predicted mass of pED208 TraM #2-52 by Protein Prospector MS-Digest (<http://prospector.ucsf.edu/prospector/mshome.htm>)

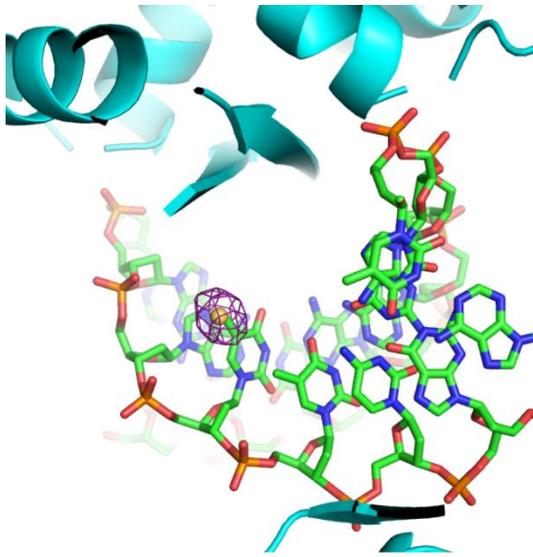


Figure 4-6 Iodine peak from difference fourier map  
Map was calculated after molecular replacement and refinement of the pED208 TraM C-terminal domain. Peak was contoured at  $5.5\sigma$ . Iodine shown as yellow sphere.

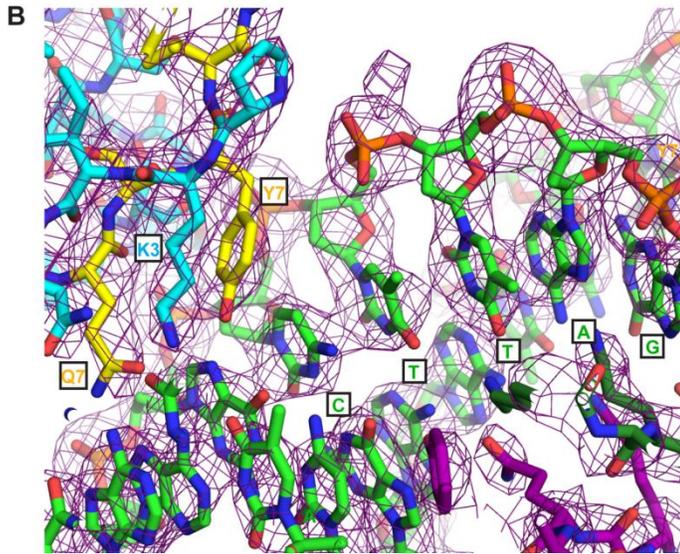
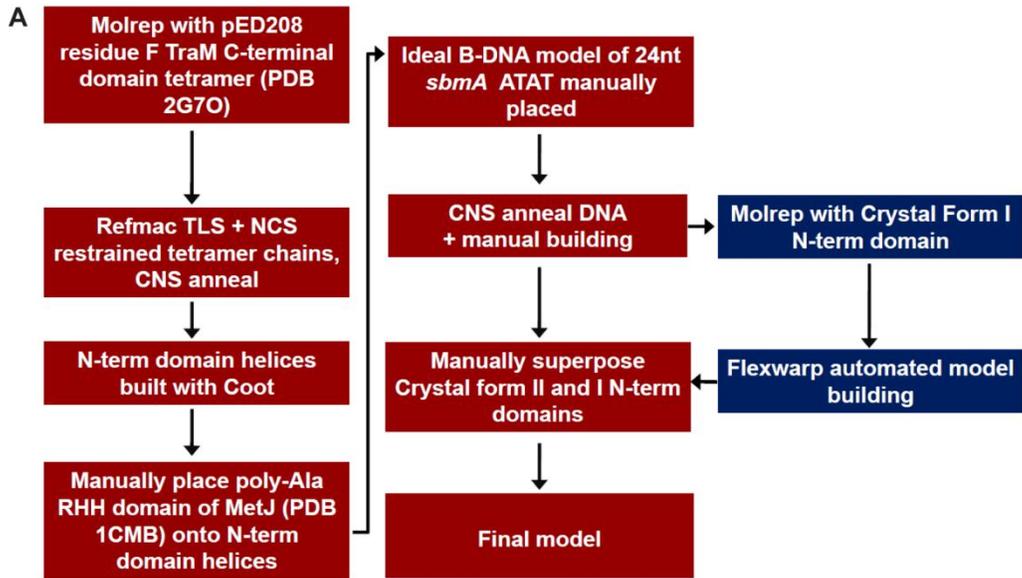


Figure 4-7 Structure Solution of pED208 TraM-*sbmA* complex and TraM N-terminal domain

a) Overview of structure solution process

b) Prime-and-switch map showing electron density at the protein-DNA interface contoured at  $1.9\sigma$ .

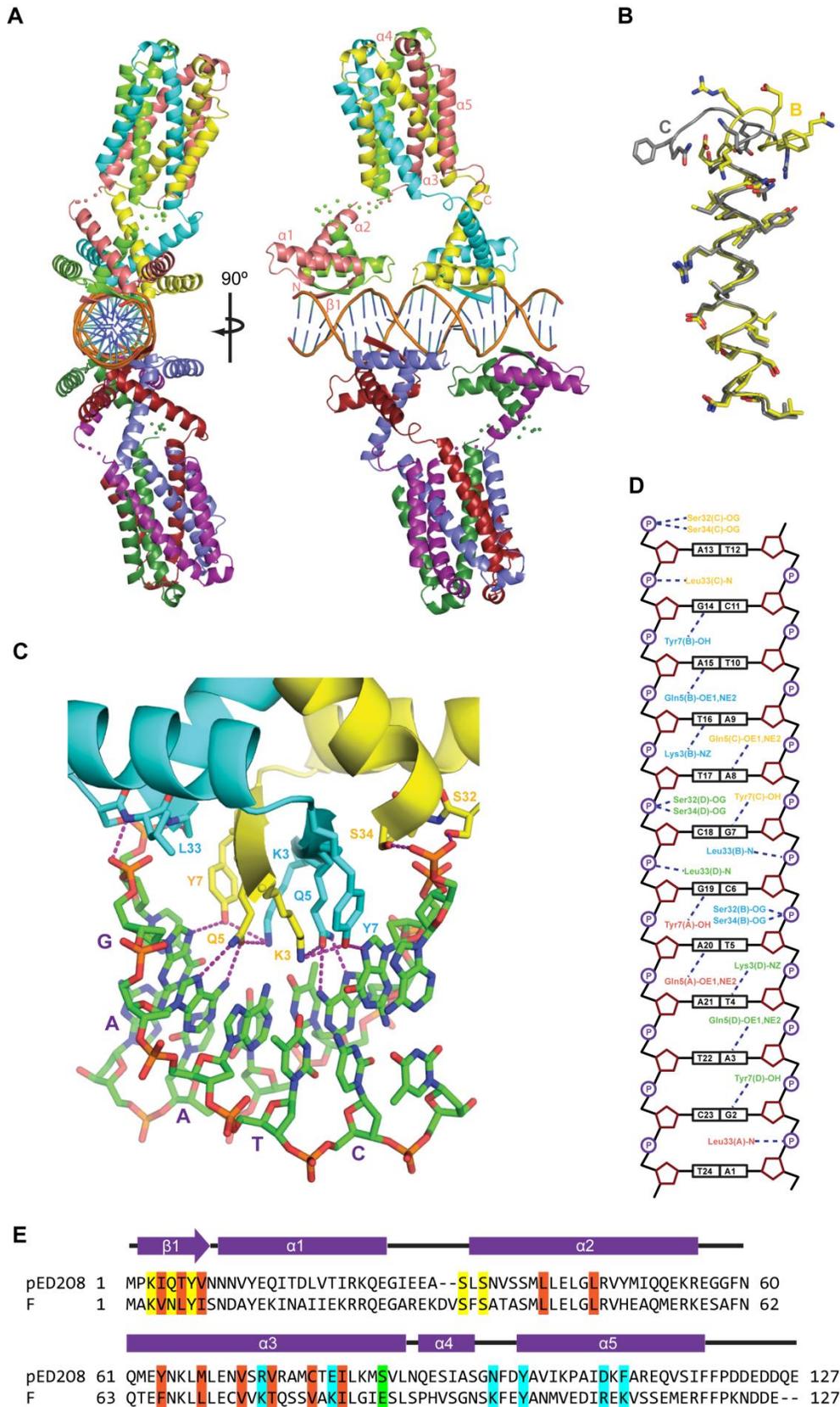


Figure 4-8 Crystal structure of pED208 TraM bound to *sbmA*

- a) Orthogonal views of the overall structure of the TraM-*sbmA* complex.  $\alpha$ -helices and  $\beta$ -strands are indicated. Disordered linkers are indicated by spheres, one for each C $\alpha$  that could not be refined.
- b) Superimposition of C $\alpha$ s of two chains from the same N-terminal domain bound to *sbmA*, residues 33-60 of Chain B and Chain C, showing that  $\alpha$ 2 of Chain C becomes unwound relative to Chain B.
- c) Interactions between the N-terminal domain of TraM and DNA. Hydrogen bonds are indicated by purple dashed lines. The GAATC binding motif consisting of bases G7 to T11 of *sbmA* is indicated with purple letters.
- d) Schematic diagram of TraM-*sbmA* interactions for one TraM tetramer. Hydrogen bonds are indicated with dashed lines.
- e) Sequence alignment of pED208 TraM and F TraM. Secondary structure elements are indicated. Conserved hydrophobic core residues are highlighted in orange, DNA-contacting residues in yellow, TraD C-terminal tail-contacting residues in cyan, and position 88, responsible for protonation-mediated destabilization of the F TraM tetramer, in green.

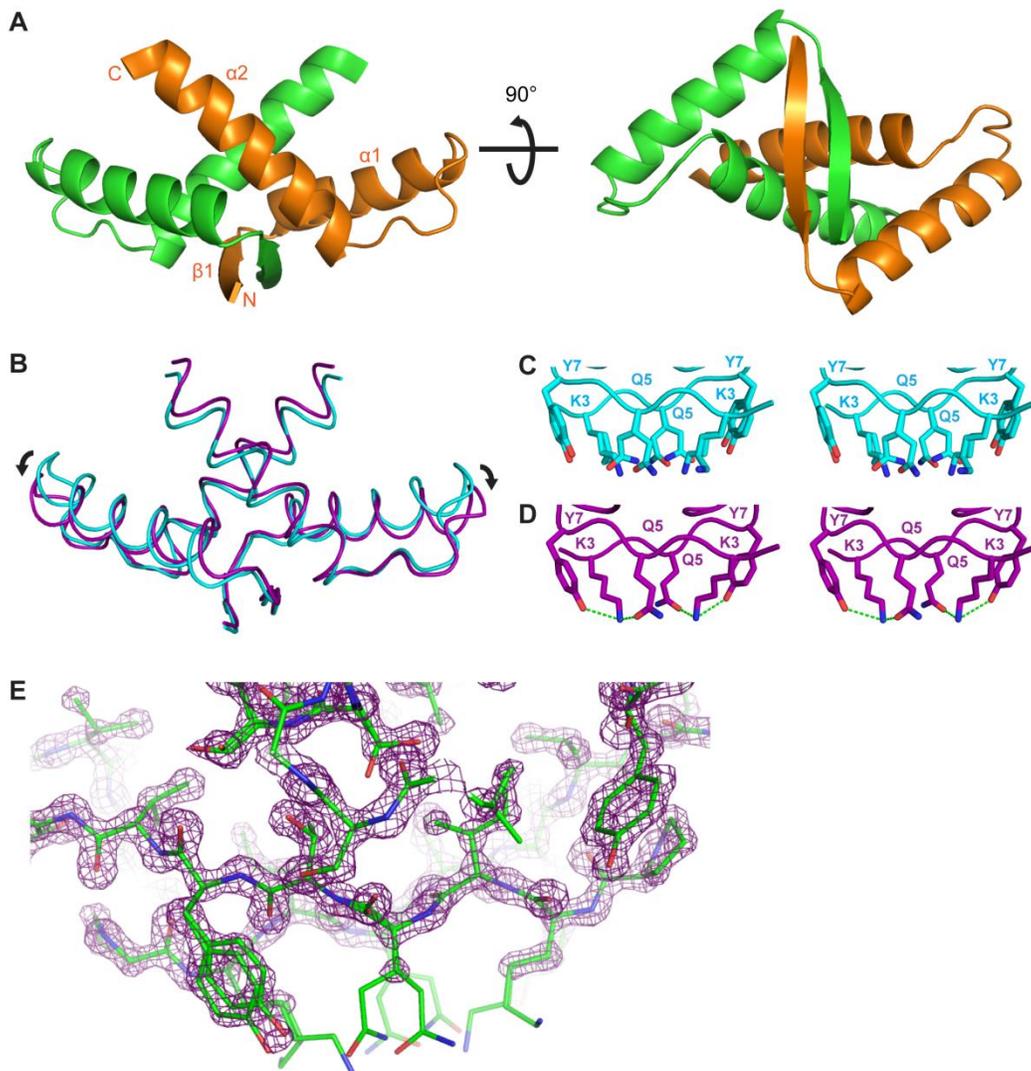


Figure 4-9 Crystal Structure of pED208 apo-N-terminal domain #2-52

- Orthogonal views of the N-terminal domain of pED208 TraM, solved to 1.3Å resolution
- Superimposition of Cas of the apo 1.30 Å N-terminal domain pED208 TraM structure (cyan) with the 2.90Å N-terminal domain TraM structure bound to *sbmA* (purple). Shift in  $\alpha 1$ - $\alpha 2$  loop upon binding to DNA is indicated with arrows.
- Stereo view of the  $\beta$ -sheet of pED208 TraM when unbound, showing alternate conformers of Lys3, Glu5, and Tyr7
- Stereo view of the  $\beta$ -sheet DNA-binding residues of pED208 TraM when bound to *sbmA*, showing hydrogen bond interactions between Lys3, Glu5, and Tyr7 in their DNA-bound conformation
- $2mF_o - F_c$  electron density map contoured at  $2.3\sigma$

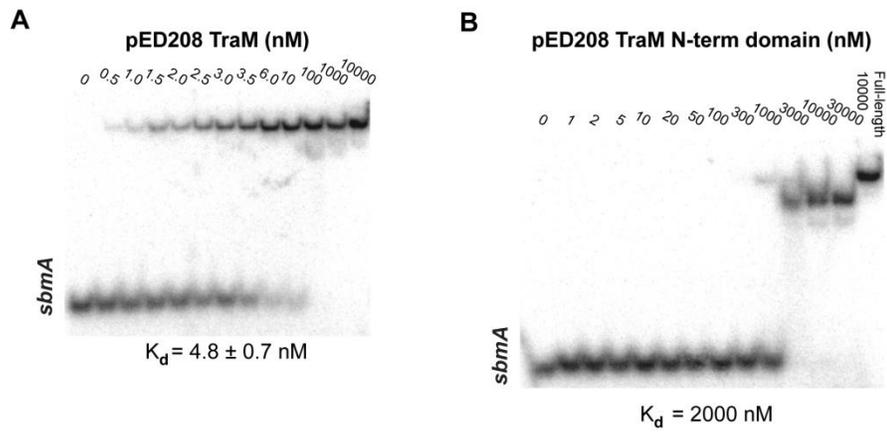


Figure 4-10 Comparison of full-length and N-terminal domain pED208 TraM binding to 24bp *sbmA*

- a) Binding affinity of full-length TraM for 24 bp *sbmA*
- b) Binding affinity of the TraM N-terminal domain for 24bp *sbmA*

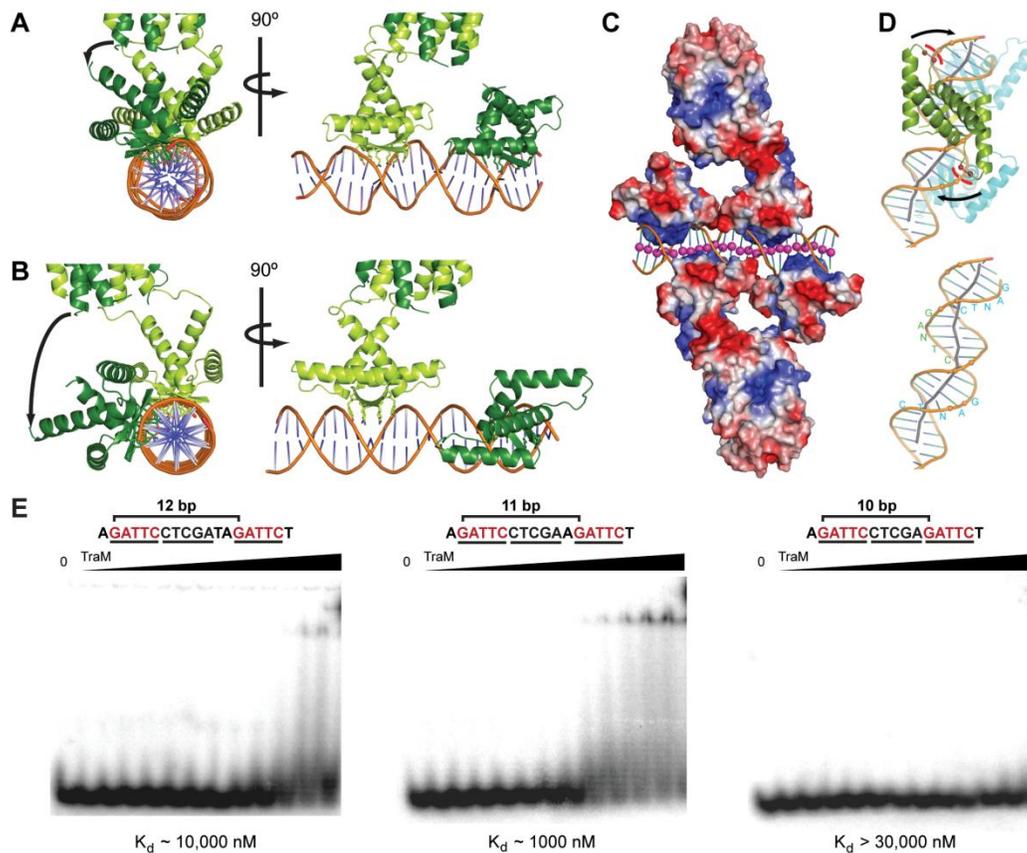


Figure 4-11 Cooperative recognition of *sbmA* DNA by TraM is mediated by DNA unwinding and kinking.

- Alignment of N-terminal domains of one TraM tetramer, as observed in the crystal structure. Binding of TraM to *sbmA* results in DNA unwinding and alignment of the N-terminal DNA binding domains of one tetramer on the same side of the DNA.
- Model of TraM N-terminal domains bound to *sbmA* in an ideal B-DNA conformation.
- Electrostatic surface potential map of TraM bound to *sbmA*. The DNA helix axis (indicated by pink spheres) appears to be bent by attraction to the basic B-sheet surface and repulsion by the acidic loops between  $\alpha 1$  and  $\alpha 2$ .
- DNA kinking induced in *sbmA* is driven by repulsion of the DNA backbone by the  $\alpha 1$ - $\alpha 2$  loop. The negatively charged side chains of Glu29 and Glu30, indicated by red spheres, repel the phosphate backbone. The DNA helix axis is indicated by a grey line. GANTC binding motifs are indicated in blue and green letters.
- Effect of varying the number of base pairs between two GANTC motifs in *sbmA* on the binding of pED208 TraM measured by EMSA. Each DNA contains only 2 GANTC motifs (highlighted in red) separated by either 12 (left panel), 11 (centre panel), or 10 bp (right panel). pED208 TraM concentrations in each lane are 0, 2, 5, 10, 20, 50, 100, 250, 600, 1000, 3000, 10000, and 30000 nM.

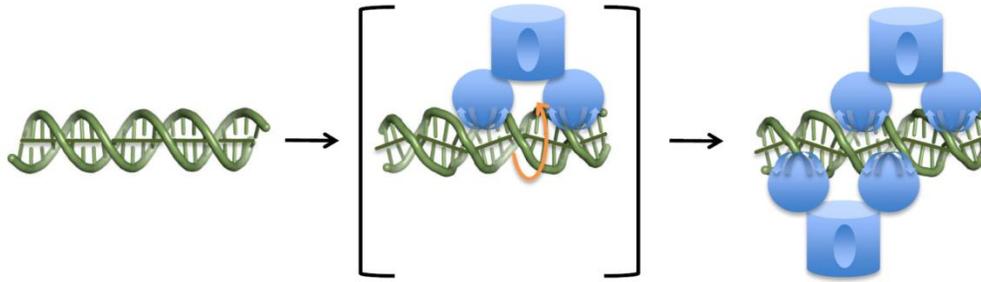


Figure 4-12 Model for cooperative recognition of DNA by TraM  
*sbmA* DNA exists in a B-like conformation in the absence of TraM (left). A single TraM tetramer (blue) binds a pair of GANTC elements via its two RHH domains, thereby unwinding and kinking the DNA to form an unstable intermediate complex (center). Binding of the first tetramer induces a DNA conformation that aligns the remaining free pair of GANTC elements on the opposite side of the DNA helix, which facilitates binding of the second tetramer and stabilization of the complex (right).

This figure was prepared by Mark Glover

## References

- Bolinger D, Sulkowska JI, Hsu HP, Mirny LA, Kardar M, Onuchic JN & Virnau P (2010) A Stevedore's protein knot. *PLoS Comput Biol* **6**: e1000731.
- Brunger AT, Adams PD, Clore GM, *et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**: 905-921.
- Cohen SX, Ben Jelloul M, Long F, *et al.* (2008) ARP/wARP and molecular replacement: the next generation. *Acta Crystallogr D Biol Crystallogr* **64**: 49-60.
- Di Laurenzio L, Frost LS, Finlay BB & Paranchych W (1991) Characterization of the oriT region of the IncFV plasmid pED208. *Mol Microbiol* **5**: 1779-1790.
- Di Laurenzio L, Scraba DG, Paranchych W & Frost LS (1995) Studies on the binding of integration host factor (IHF) and TraM to the origin of transfer of the IncFV plasmid pED208. *Mol Gen Genet* **247**: 726-734.
- Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126-2132.
- Falkow S & Baron LS (1962) EPISOMIC ELEMENT IN A STRAIN OF SALMONELLA TYPHOSA. *J Bacteriol* **84**: 581-589.
- Fekete RA & Frost LS (2000) Mobilization of chimeric oriT plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. *J Bacteriol* **182**: 4022-4027.
- Fekete RA & Frost LS (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.
- Finlay BB, Paranchych W & Falkow S (1983) Characterization of conjugative plasmid EDP208. *J Bacteriol* **156**: 230-235.
- Finlay BB, Frost LS & Paranchych W (1986) Nucleotide sequence of the tra YALE region from IncFV plasmid pED208. *J Bacteriol* **168**: 990-998.
- Frost LS, Ippen-Ihler K & Skurray RA (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* **58**: 162-210.
- Fu YH, Tsai MM, Luo YN & Deonier RC (1991) Deletion analysis of the F plasmid oriT locus. *J Bacteriol* **173**: 1012-1020.
- Gomis-Ruth FX, Sola M, Acebo P, *et al.* (1998) The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J* **17**: 7404-7415.
- Guzman LM, Belin D, Carson MJ & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Hanahan D (1983) Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**: 557-580.
- Howard MT, Nelson WC & Matson SW (1995) Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem* **270**: 28381-28386.
- Inamoto S, Fukuda H, Abo T & Ohtsubo E (1994) Site- and strand-specific nicking at oriT of plasmid R100 in a purified system: enhancement of the nicking activity of Tral (helicase I) with TraY and IHF. *J Biochem* **116**: 838-844.
- Itou H, Watanabe N, Yao M, Shirakihara Y & Tanaka I (2010) Crystal structures of the multidrug binding repressor Corynebacterium glutamicum CgmR in complex with inducers and with an operator. *J Mol Biol* **403**: 174-184.
- Jeng WY, Ko TP, Liu CI, Guo RT, Liu CL, Shr HL & Wang AH (2008) Crystal structure of IcaR, a repressor of the TetR family implicated in biofilm formation in Staphylococcus epidermidis. *Nucleic Acids Res* **36**: 1567-1577.

- Karl W, Bamberger M & Zechner EL (2001) Transfer protein TraY of plasmid R1 stimulates Tral-catalyzed oriT cleavage in vivo. *J Bacteriol* **183**: 909-914.
- Kupelwieser G, Schwab M, Hogenauer G, Koraimann G & Zechner EL (1998) Transfer protein TraM stimulates Tral-catalyzed cleavage of the transfer origin of plasmid R1 in vivo. *J Mol Biol* **275**: 81-94.
- Llosa M, Zunzunegui S & de la Cruz F (2003) Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc Natl Acad Sci U S A* **100**: 10465-10470.
- Lu D, Fillet S, Meng C, *et al.* (2010) Crystal structure of TtgV in complex with its DNA operator reveals a general model for cooperative DNA binding of tetrameric gene regulators. *Genes Dev* **24**: 2556-2565.
- Lu J, Fekete RA & Frost LS (2003) A rapid screen for functional mutants of TraM, an autoregulatory protein required for F conjugation. *Mol Genet Genomics* **269**: 227-233.
- Lu J, Zhao W & Frost LS (2004) Mutational analysis of TraM correlates oligomerization and DNA binding with autoregulation and conjugative DNA transfer. *J Biol Chem* **279**: 55324-55333.
- Lu J, den Dulk-Ras A, Hooykaas PJ & Glover JN (2009) *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. *Proc Natl Acad Sci U S A* **106**: 9643-9648.
- Lu J, Manchak J, Klimke W, Davidson C, Firth N, Skurray RA & Frost LS (2002) Analysis and characterization of the IncFV plasmid pED208 transfer region. *Plasmid* **48**: 24-37.
- Lu J, Edwards RA, Wong JJ, Manchak J, Scott PG, Frost LS & Glover JN (2006) Protonation-mediated structural flexibility in the F conjugation regulatory protein, TraM. *EMBO J* **25**: 2930-2939.
- Lu XJ & Olson WK (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc* **3**: 1213-1227.
- Luscombe NM, Laskowski RA & Thornton JM (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res* **29**: 2860-2874.
- Miller DL & Schildbach JF (2003) Evidence for a monomeric intermediate in the reversible unfolding of F factor TraM. *J Biol Chem* **278**: 10400-10407.
- Moncalian G & de la Cruz F (2004) DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* **1701**: 15-23.
- Moore D, Wu JH, Kathir P, Hamilton CM & Ippen-Ihler K (1987) Analysis of transfer genes and gene products within the traB-traC region of the *Escherichia coli* fertility factor, F. *J Bacteriol* **169**: 3994-4002.
- Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**: 240-255.
- Otwinowski ZaM, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology Volume 276: Macromolecular Crystallography, part A*: 307-326.
- Penfold SS, Simon J & Frost LS (1996) Regulation of the expression of the traM gene of the F sex factor of *Escherichia coli*. *Mol Microbiol* **20**: 549-558.

- Pohl E, Holmes RK & Hol WG (1999) Crystal structure of the iron-dependent regulator (IdeR) from *Mycobacterium tuberculosis* shows both metal binding sites fully occupied. *J Mol Biol* **285**: 1145-1156.
- Pohl E, Holmes RK & Hol WG (1999) Crystal structure of a cobalt-activated diphtheria toxin repressor-DNA complex reveals a metal-binding SH3-like domain. *J Mol Biol* **292**: 653-667.
- Rafferty JB, Somers WS, Saint-Girons I & Phillips SE (1989) Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. *Nature* **341**: 705-710.
- Schreiter ER & Drennan CL (2007) Ribbon-helix-helix transcription factors: variations on a theme. *Nat Rev Microbiol* **5**: 710-720.
- Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA & Brennan RG (2002) Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J* **21**: 1210-1218.
- Schwab M, Reisenzein H & Hogenauer G (1993) TraM of plasmid R1 regulates its own expression. *Mol Microbiol* **7**: 795-803.
- Short JM, Fernandez JM, Sorge JA & Huse WD (1988) Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Res* **16**: 7583-7600.
- Stayrook S, Jaru-Ampornpan P, Ni J, Hochschild A & Lewis M (2008) Crystal structure of the lambda repressor and a model for pairwise cooperative operator binding. *Nature* **452**: 1022-1025.
- Stockner T, Plugariu C, Koraimann G, Hogenauer G, Bermel W, Prytulla S & Sterk H (2001) Solution structure of the DNA-binding domain of TraM. *Biochemistry* **40**: 3370-3377.
- Tabor S & Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* **82**: 1074-1078.
- Tato I, Matilla I, Arechaga I, Zunzunegui S, de la Cruz F & Cabezon E (2007) The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J Biol Chem* **282**: 25569-25576.
- Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**: 22-25.
- Verdino P, Keller W, Strohmaier H, Bischof K, Lindner H & Koraimann G (1999) The essential transfer protein TraM binds to DNA as a tetramer. *J Biol Chem* **274**: 37421-37428.
- White A, Ding X, vanderSpek JC, Murphy JR & Ringe D (1998) Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex. *Nature* **394**: 502-506.
- Willetts NS & Finnegan DJ (1970) Characteristics of *E. coli* K12 strains carrying both an F prime and an R factor. *Genet Res* **16**: 113-122.
- Wisedchaisri G, Holmes RK & Hol WG (2004) Crystal structure of an IdeR-DNA complex reveals a conformational change in activated IdeR for base-specific interactions. *J Mol Biol* **342**: 1155-1169.
- Wong JJ, Lu J, Edwards RA, Frost LS & Glover JNM (2011) Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**: 6775-6788.

## Chapter 5

### Specificity in pED208 TraM-DNA and TraM-TraD interactions<sup>1</sup>

#### Overview

The crystal structure of pED208 TraM bound to *sbmA* shows how the N-terminal domain RHH domain of TraM recognizes the conserved bases of the GANTC binding motif. In this chapter, the role of the G and A bases in TraM binding are investigated by competition assay. Both bases contribute significantly to binding affinity, with A having a larger role. A mixture of pED208 and F TraM was able to bind with high affinity and cooperativity to a hybrid *sbmA* fragment with the binding motifs of F on one side of the DNA helix and those of pED208 on the other. This shows that the mechanism of TraM cooperative binding is conserved between F and pED208. A model of the TraM-bound pED208 TraD C-terminus was made, based on the crystal structure of the F TraD C-terminus bound to TraM and the C-terminal tetramer of pED208 TraM from the complex structure. A charge exchange exists between F and pED208 in two residues that form electrostatic interactions with each other. pED208 TraD has Arg734 in place of F Asp715, and pED208 TraM has Glu81 in place of Lys83. The TraM-TraD C-terminus interaction and pED208 Arg734 are shown *in vivo* to be a specificity determinants between pED208 and F.

---

<sup>1</sup> Part of this work was previously published: Wong JJ, Lu J, Edwards RA, Frost LS & Glover JNM (2011) Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**: 6775-6788.

## Introduction

It has long been observed that F-like plasmids are able to mobilize only their cognate plasmid. Kupelwieser et. al. show that the F-plasmid-derived pOX38-Km can mobilize a plasmid containing R1 *oriT* and expressing R1 TraM, but not a variant of the same plasmid that cannot endogenously express TraM. They also show that the R1-16 plasmid cannot mobilize plasmids with the F-like P307 *oriT*, whereas pOX38-Km can. The ability of plasmids to self-mobilize is well-correlated to whether the endogenously-produced TraM can bind to the *oriT* sequence, and they show that the DNA-binding ability resides in the N-terminal residues of TraM (Kupelwieser, et al., 1998). A SELEX experiment for R1 TraM binding resulted in *in vitro*-selected sequences that were very close to the native binding motif of R1 TraM, providing evidence that TraM binding is DNA sequence-specific (Geist & Brantl, 2008). Fekete and Frost showed that R100-1 can mobilize a plasmid with R100 *oriT* but not one with F *oriT* (Fekete & Frost, 2000). Lu et.al. showed that pED208 is unable to mobilize plasmids with the F or R100 *oriT*, and that pOX38-Km cannot mobilize pED208 (Lu, et al., 2002).

Previous studies have suggested that TraD can interact with TraM from different plasmids within the F-like plasmid family, but it is fairly selective for F-like relaxosomes rather than those from other incompatibility groups. F TraD is able to interact with R1 TraM *in vitro*. However, overexpression of the last 38 amino acids of the R1 TraD had a negative dominant effect on R1-16 transfer but not on pSU2007 (R388 derivative) transfer (Beranek, et al., 2004). Compared to coupling proteins from IncP plasmids like RP4 and IncW plasmids like R388, TraD of F-like plasmids has a substantial C-terminal extension that confers specificity and efficiency in self-transmission at the expense of being able to

mobilize a wider range of plasmid Inc groups (Sastre, *et al.*, 1998). These studies as a whole strongly suggest that a great deal of this specificity results from interaction between TraM and *oriT* DNA, and interaction between TraD and TraM.

The pED208 TraM-*sbmA* crystal structure reveals in detail how TraM of pED208 interacts with *sbmA* DNA in a specific manner. The N-terminal domains of TraM dimerize to form ribbon-helix-helix (RHH) DNA recognition modules, where, as is typical of RHH domains, the N-terminal  $\beta$ -sheet residues contact DNA bases in the major groove (Chapter 4). In general, plasmids tend to have in their *oriT* region DNA sequence motifs unique to the individual plasmid and for their cognate TraM RHH domain, and we confirmed the role of individual bases within the strictly conserved and palindromic GANTC binding motif for pED208 TraM.

The pED208 TraM-*sbmA* crystal structure also provides a scaffold from which the binding of pED208 TraD can be modeled. We propose a mode of specific interaction between the pED208 TraM tetramerization domain and the C-terminal tail of its cognate TraD, which is a critical determinant of allelic specificity among F-like systems. We demonstrate *in vivo* the first instance of plasmid specificity in coupling protein-relaxosome interaction within the F-like plasmid family, which is that between pED208 and F. The differences in key residues in the TraD C-terminus and TraM binding pocket provide a rationale for the specificity of pED208 and F TraD in mobilizing their cognate plasmids. The mechanism of cooperative binding allows TraM to dictate plasmid specificity by

binding to *oriT* and mediating relaxosome-coupling protein interactions simultaneously.

## Results

### *F and pED208 TraM bind specifically to their cognate TraM sites*

F and pED208 TraM bind to their cognate TraM sites with exquisite specificity, which is remarkable given the only difference in their DNA-contacting  $\beta$ -strand residues is an Asn to Gln substitution in pED208. In the EMSA experiment of Figure 5-1A, the length of *sbmA* is controlled for each DNA to prevent any effects from differences in length –pED208 *sbmA*, like the F *sbmA* fragment used, was 30 bp in length. No binding is detectable for F TraM to pED208 *sbmA*, and only non-specific binding at very high concentrations occurs for pED208 TraM to F *sbmA*. However, both bind with nanomolar affinity to DNA from the same plasmid (Figure 5-1A). The additional band occurring at high concentrations for pED208 binding to pED208 *sbmA* is non-specific binding of TraM to residual single-stranded DNA. Single-stranded DNA migrating below the double stranded DNA clearly disappears as TraM concentration increases (data not shown). The portion of TraM that binds DNA is the N-terminal domain, as shown in several previous studies (Schwab, *et al.*, 1993, Kupelwieser, *et al.*, 1998, Miller & Schildbach, 2003, Lu, *et al.*, 2004) and by the ability of a chimeric TraM protein with the F N-terminal domain and pED208 C-terminal to bind to F *sbmA* with wild-type affinity, but not significantly to pED208 *sbmA* (Figure 5-1B).

*Role of TraM-sbmA interactions observed in the TraM-DNA crystal structure in TraM-sbmA binding*

The N-terminal  $\beta$ -sheet enters the DNA major groove and contacts the GANTC repeat with the side chains of residues on the surface. Gln5 contacts the adenine base of the binding motif, and Tyr7 contacts the guanine base. Lys3 forms a hydrogen bond network with Gln5 and Tyr7 (Figure 5-2A and 5-2B). The crystal structure strongly suggest that the conserved GC and AT base pairs are essential for binding.

We used a competitive EMSA assay to test the relative importance of G and A within the GANTC motifs for TraM binding (Figure 5-3A and 5-3B). TraM-*sbmA* complexes were formed with labeled *sbmA* DNA, and unlabelled competitor DNA was added to compete with the labeled DNA for TraM binding. Either the G or A was mutated in each double-stranded oligo and in various numbers of GANTC sites in order to test their effect on binding. Mutation of GtoC or AtoT is deleterious to TraM binding, with AtoT mutations have a greater binding defect. These results demonstrate that TraM recognizes the four GANTC motifs in *sbmA* in a highly specific and cooperative manner.

*Cooperative binding at sbmA is conserved in F and pED208*

We asked whether the deformations in DNA structure observed in the pED208 TraM-*sbmA* crystal structure are involved in the cooperative DNA interactions by other TraM proteins. Although the binding motifs are not conserved between the F and pED208 plasmids, the 12 base-pair spacing between sites bound by the same TraM tetramer is (Figure 5-6D). To ascertain whether the DNA deformations induced by pED208 TraM are similar to those

induced by F TraM, we tested the ability of pED208 and F TraM to bind a hybrid *sbmA* DNA in which the 1<sup>st</sup> and 3<sup>rd</sup> binding motifs correspond to those found in pED208 *sbmA* (GANTC), while the 2<sup>nd</sup> and 4<sup>th</sup> motifs correspond to F *sbmA* (A(G/C)CG(G/C)T) (Figure 5-4). Since the steric hindrance effects of placing the F and pED208 motifs side-by-side in an unnatural DNA were unknown, 3 versions of the hybrid *sbmA* sequence with different F DNA fragments were tested for ability to bind F TraM, pED208 TraM, or a 1:1 mixture of the two. Hybrid *sbmA* has an F motif that includes 1 base 3' to the binding motif, and Hybrid *sbmA* B has an F motif that is shifted to include 1 base 5' to the binding motif. Hybrid *sbmA* C has an extra base pair on the end to ensure that 3 bases are flanking each F motif, and one base pair taken off the front to ensure the length of the DNA remains at 30 bp. All 3 DNAs showed defects in binding affinity to F or pED208. However, high affinity and cooperative binding could be restored in Hybrid *sbmA* B and C, but not A. The presence of two base pairs at the end of B instead of three does not appear to cause any significant defect in binding, which is inconsistent with 30 bp being the minimum length of F *sbmA* required for high affinity binding (Figure 5-4). This may indicate there may be a subtler asymmetry in F *sbmA* binding than has been previously investigated. Hybrid *sbmA* C is chosen to be the subject of our discussion of TraM cooperativity.

While either F TraM or pED208 TraM bound this DNA with significantly reduced affinity in EMSA, an equimolar mixture of F and pED208 TraM tetramers bound this DNA at high affinity, similar to the affinity of either F TraM or pED208 TraM for their cognate *sbmA*. Moreover, the mobility of the shifted species derived from the pED208/F TraM mixture was distinct from that of either the F TraM complex or the pED208 complex, demonstrating that the complex derived

using the protein mixture contains both F and pED208 TraM tetramers (Figure 5-4C). Thus, this result indicates that the structural changes induced in the DNA by TraM from these two plasmids are similar enough to facilitate cooperative binding on the hybrid *sbmA*, and suggest that similar distortions will facilitate DNA recognition within the family of TraM proteins. This experiment also further demonstrates that the binding motifs of pED208 and F *sbmA* are specific to their cognate TraM. High affinity (nanomolar level), cooperative binding does not occur with pED208 or F TraM alone, so only one set of binding motifs in the hybrid DNA is occupied by either TraM.

*The  $\alpha 1$ - $\alpha 2$  loop may provide additional binding specificity*

Unlike pED208 TraM, the F plasmid TraM requires a longer *sbmA* fragment in order to attain maximum binding affinity. Whereas the pED208 TraM requires a 24 bp *sbmA* and 1 base pair flanking its TraM binding motif (GANTC), the F plasmid TraM requires a 30 bp *sbmA* (Figure 5-5) and 3 base pairs flanking its TraM binding motif (A(G/C)CG(G/C)T) (Figure 5-6D). This finding led us to hypothesize a rationale for the requirement of extra flanking bases in the F plasmid. In light of previous unpublished results suggesting their importance, we hypothesized that the residues in the loop between  $\alpha 1$  and  $\alpha 2$  play a role in the specificity of F-like plasmid TraM binding

Kinking of the pED208 *sbmA* is mediated by repulsion of the phosphate backbone by Glu29 and Glu30 on the  $\alpha 1$ - $\alpha 2$  loop (Figure 5-6A and 5-6B). In F TraM, Glu29 is replaced by Arg29 and there is an additional basic residue, Lys31, in the  $\alpha 1$ - $\alpha 2$  loop, opening up the possibility that it may form ionic interactions with the phosphate backbone (Figure 5-6C). A role for F TraM Arg 29

in binding to F *sbmA* is further suggested by its protection from trypsin digestion when bound to DNA (J.Lu, unpublished data). This may explain why the minimal length of *sbmA* needed for maximal binding is 30 bp for the F plasmid, instead of 24 bp for the pED208 plasmid—the additional DNA bases of the F minimal *sbmA* may allow the basic residues of the  $\alpha$ 1- $\alpha$ 2 loop to form additional interactions with the phosphate backbone which may contribute to binding affinity.

*Selective TraM-TraD interactions also govern allelic specificity*

In addition to DNA binding, TraM also must contact the conjugative pore protein, TraD, to effect conjugation (Beranek, *et al.*, 2004). TraM-TraD interactions critically depend on the recognition of the C-terminal tail of TraD by a groove on the surface of the TraM tetramerization domain (Lu, *et al.*, 2008). To test the relative importance of TraM-DNA and TraM-TraD interactions for plasmid specificity, we assessed the ability of F or pED208 TraM and TraD, as well as chimeric molecules derived from these proteins, to rescue conjugative transfer of a TraM- and TraD-deficient F plasmid derivative (Figure 5-7A and Table 5-1). A critical role for plasmid specific TraM-TraD interactions was demonstrated by the finding that a chimeric TraM with an F N-terminal domain and a pED208 C-terminal domain, TraM[F<sup>1-55</sup>:pED208<sup>56-127</sup>], does not complement a TraM-deficient F plasmid (Figure 5-7A), despite the fact that the chimeric protein binds to F *sbmA* with wild type affinity (Figure 5-1B). Likewise, substitution of the eight C-terminal residues of F TraD with those of pED208 TraD in TraD[F<sup>1-709</sup>:pED208<sup>729-736</sup>] also disrupts conjugation when co-expressed with F TraM. Significantly, mating is rescued when both chimeric proteins, TraM[F<sup>1-55</sup>:pED208<sup>56-127</sup>] and TraD[F<sup>1-709</sup>:pED208<sup>729-736</sup>], are co-expressed, demonstrating plasmid specific

TraM-TraD interactions rely on specific interactions between the C-terminal TraD tail and the TraM tetramerization domain (Figure 5-7A).

A comparison of the sequences of the TraD tails and the structures of TraD binding pockets on TraM in the two plasmid systems suggests an explanation for this specificity. pED208 TraD contains a single positively charged residue (Arg734) in its otherwise highly negatively charged tail at a position that is negatively charged in F TraD (Figure 5-7B). Modeling of the pED208 TraM-TraD interaction based on the structure of F TraM-TraD complex suggests that pED208 TraD Arg734 will be juxtaposed with a negatively charged residue (Glu81) in the pED208 TraM pocket (Figure 5-7D). In F TraM, this residue is positively charged (Lys83) (Figure 5-7C). To test the hypothesis that complementary charge interactions between TraM and TraD may help define specificity, we mutated Arg734 to Asp in the chimeric TraD protein TraD[F<sup>1-709</sup>:pED208<sup>729-736</sup>:R734D]. This protein significantly rescued conjugation compared to TraD[F<sup>1-709</sup>:pED208<sup>729-736</sup>] providing further support that side chain-specific interactions analogous to those observed in the F plasmid system (Lu, *et al.*, 2008) are necessary *in vivo* for conjugation in pED208 (Figure 5-7A). These interactions define the binding specificity of the F and pED208 TraD C-terminal tail for their cognate TraM.

## **Discussion**

### **Role of TraM in the definition of allelic specificity**

Core components of conjugative type IV secretion pores can transfer proteins with the appropriate translocation signals (Alvarez-Martinez & Christie, 2009), whereas the conjugative pore requires both a T4SS and a coupling

protein that recognizes the relaxosome in preparation for DNA transfer. In F-like plasmids, this recognition event prepares the relaxase, which contains complex internal translocation signals and is covalently bound to the *nic* site, for translocation to the recipient cell (Lang, *et al.*, 2010, Lang, *et al.*, 2011). Thus, the relaxosome accessory protein, TraM, and the coupling protein, TraD, confer a high level of selectivity for the cognate relaxosome (Sastre, *et al.*, 1998), with the DNA binding specificity of TraM being key for this selectivity. Both R1 and pED208 TraM RHH modules bind DNA elements containing GANTC motifs, which is explained by the critical DNA contacts made by residues at positions 3, 5 and 7 in the  $\beta$  ribbon that are conserved in these two proteins (Figure 5-4). In contrast, the F TraM RHH binds an unrelated, 6 bp DNA motif, A(G/C)CG(C/G)T. Interestingly, this difference is likely largely dictated by a single, conservative Gln-Asn substitution in the RHH  $\beta$  ribbon. Gln5 is critical for recognition of the AT base pairs in the GANTC in the pED208 system. In F, the shorter Asn5 side chain is not expected to recognize an AT pair in the same manner, and may also interact differently with Lys3, leading to a rearrangement at the protein-DNA interface. In addition, subtle changes to the way in which the RHH contacts the DNA backbone may also impact the DNA geometry, which may in turn modulate sequence specificity. For example, Glu29 and Glu30 in the  $\alpha$ 1- $\alpha$ 2 loop repel the phosphodiester backbone and contribute to DNA kinking in the pED208 TraM (Figure 5-6A and 5-6B). In F, R1 and R100, this loop is two residues longer and Glu29 is substituted with either an Arg or Lys residue (Figure 5-6D), opening up the possibility of ionic interaction with the DNA phosphate backbone. Thus, a combination of direct readout, mediated through interactions involving residues of the RHH  $\beta$ -ribbon, and indirect, structural effects, mediated by interactions

between the RHH and the DNA backbone, together govern DNA binding specificity.

In addition to specificity at the level of TraM-DNA interactions, our work also reveals that TraM-TraD contacts are also plasmid-specific. Central to TraM-TraD interaction is the recognition of the C-terminal Phe and main chain carboxylate of the TraD tail by a hydrophobic pocket in TraM (Lu, *et al.*, 2008). The carboxylate group is recognized by conserved positively charged residues that form one side of this otherwise hydrophobic pocket (Figure 5-7C). Negatively charged residues in the tail make a number of long-range electrostatic interactions with the overall positively charged surface of the TraM interaction site. The putative TraD binding pocket in pED208 TraM is well conserved with F TraM, and includes a pocket that could bind the C-terminal Tyr of the pED208 TraD tail, positively charged residues to recognize the TraD C-terminal carboxylate, as well as a number of charged residues surrounding the binding site (Figure 5-7C and 5-7D). The reversal of a potential charge-charge interaction in the pED208 TraM-TraD complex (TraM Glu81 to TraD Arg734) compared to the F complex (TraM Lys83 to TraD Asp715) appears to play a role in helping to define binding specificity between these two plasmid systems (Figure 5-7B).

The precise nature of the pED208 TraD-TraM interaction has yet to be determined experimentally. The model proposed in (Figure 5-7D) is based on the assumption that the pED208 TraD C-terminus would adopt the same  $\beta$ -hairpin conformation as the F TraD C-terminus, and was created by simply mutagenizing the non-conserved residues in the F TraD peptide to those of pED208. However, this model resulted in a steric clash between the experimentally obtained

pED208 Lys84 conformation and the TraD peptide backbone, and the lysine side chain was moved out of the way manually to reduce the steric clash. In addition, there are numerous differences in the pED208 TraM binding pocket compared to F TraM that result in non-conservative amino acid substitutions, and the proline that makes up the Pro-Gly turn of the  $\beta$ -hairpin is not present in pED208 TraD (Figure 5-7E). The result of an Autodock Vina docking experiment suggests that the binding of the pED208 TraD C-terminus to its TraM binding pocket may be fairly different – instead of the aromatic side chain of the C-terminal Tyr736 in the pocket which is occupied by that of Phe717 in F, the docking algorithm placed either an Asp732 or Glu735 in the pocket (Figure 5-8 and 5-9), perhaps due to the higher electropositive nature of that pocket in pED208 relative to F (Figure 5-7C and 5-7D). As suggested by the placement of alternate acidic residues in the pocket, there is greater range in the TraD conformations generated for pED208 TraM binding relative to the positive control, F. (Autodock Vina-generated conformations for the F TraD peptide were remarkably consistent with each other and with the conformation in the crystal structure) (Figure 5-8). Arg 734, which is in close proximity for complementary-charge interaction with Glu81 in the manually-created model is not within interaction distance of Glu81 in the majority of the models. A possible source of bias in the docking arises from the greater weight that Autodock places on satisfying ionic interactions than hydrophobic interactions (Trott & Olson, 2009). In addition, Autodock is meant for docking small molecules, and not large, flexible peptides. Inapplicabilities inherent to the docking algorithm aside, the docking results, in combination with the many amino acid differences in the TraD C-terminus and its binding pocket in TraM, suggest that we cannot assume with a firm degree of certainty that the pED208 TraD C-terminus adopts the same conformation as that of F. Even if so, it would not

invalidate our *in vivo* results showing that the residue in the position of pED208 Arg 734 is an important specificity determinant—it just may not interact with TraM in quite the way initially hypothesized.

## **Materials and Methods**

### *Growth media and bacterial strains*

Media and antibiotics for bacterial growth and bacterial strains used are described in (Wong, *et al.*, 2011).

### *Primers, Plasmids, and cloning of TraM and TraD mutants*

Primers, plasmids, and cloning of all TraM and TraD constructs used are described in (Wong, *et al.*, 2011).

pACYC184 (Chang & Cohen, 1978), pBad24 (Guzman, *et al.*, 1995), pBluescript Amp<sup>r</sup>; (Short, *et al.*, 1988), pOX38-MK38 (Penfold, *et al.*, 1996), pOX38-DM (Lu, *et al.*, 2008), pJMTraD (Lu, *et al.*, 2008), pJLM400 (Lu, *et al.*, 2004), pRF105 (Fekete & Frost, 2000), pRF911 (Fekete & Frost, 2002), pED208 (Falkow & Baron, 1962), pRFM200 (Lu, *et al.*, 2003), and pT7-5 (Tabor & Richardson, 1985) have been described previously.

The 0.5-kb *EcoRI-BamHI* fragment of the PCR products amplified from pED208 using primer pair JLU91 was cloned into the *EcoRI-BamHI* sites of pT7-5 or pBluescript KS (+), resulting in pJLEM200 or pJLM404, respectively. The 0.5-kb *EcoRI-BamHI* fragment of the PCR products amplified from pRF105 using primer pair JLU601 and JLU602, or JLU603 and JLU602, was cloned into the *EcoRI-BamHI* sites of pJLM400, resulting in pJLM401 or pJLM402, respectively.

Overlap extension (Ho, *et al.*, 1989) was used to construct pJLM407 expressing a hybrid TraM (F TraM<sup>1-55</sup>:pED208 TraM<sup>56-127</sup>). PCR primer pair JLU3 and JLU612 were used to amplify a fragment containing F *traM*<sup>1-55</sup> from pRFM200. Primer pair JLU611 and JLU608 were used to amplify a fragment containing pED208 *traM*<sup>56-127</sup> from pJLEM200. Primer pair JLU3 and JLU608 were used to amplify the full-length hybrid *traM* (F TraM<sup>1-55</sup>:pED208 TraM<sup>56-127</sup>) fragment, which was further digested by *EcoRI* and *BamHI* and cloned into the *EcoRI*-*BamHI* sites of pBluescript KS (+), resulting pJLM407.

The 1.5-kb *EcoRI*-*Scal* fragment from pJLM400 or pJLM407 was cloned into the *EcoRI*-*Scal* sites of pACYC184, resulting in pACM400 or pACM407, respectively. The 1.1-kb *PstI*-*HindIII* fragment of PCR products amplified from pJMtraD using primer pair JLU262 and JLU263, or JLU262 and JLU264, was cloned into the *PstI*-*HindIII* sites of pJMtraD, resulting in pJLD263 or pJLD264, respectively.

#### *Donor ability assays*

*E. coli* XK1200 and ED24 were used as donor and recipient strains, respectively. The mating experiments were performed as previously described (Lu, *et al.*, 2002). Donor ability was calculated as the number of transconjugants divided by the number of donors. Each assay was repeated 3 times and the averaged values are reported. Standard deviations of all mating assays were within one log unit.

#### *pED208 TraM binding to sbmA by electrophoretic mobility shift assay*

*sbmA* oligos were P<sup>32</sup>-labelled with T4 polynucleotide kinase (Invitrogen) and unincorporated nucleotides were removed by P-30 Micro Bio-Spin columns buffered in 10 mM Tris pH 7.4 (Bio-Rad). TraM-*sbmA* binding buffer was 50 mM Tris pH 7.5, 10% glycerol, 30 ng/μL bovine serum albumin (Pierce), 20 ng/μL polydI-dC (Roche). 0.1 nM of *sbmA* oligo was added to each binding reaction containing the indicated amount of TraM and incubated for 10 minutes at room temperature. TraM-*sbmA* mixtures were run on 1x TBE-buffered 12% 29:1 acrylamide gels at 200 Volts for 45 minutes at 4°C.

#### *Competition of Wild-type pED208 sbmA by pED208 sbmA mutants*

0.1 nM of P<sup>32</sup>-labelled pED208 24 bp wild-type *sbmA* was mixed with 50 nM of TraM (enough to effectively bind all unbound DNA) in 50 mM Tris pH 7.5, 10% glycerol, and 30 ng/μL bovine serum albumin (Pierce). TraM-*sbmA* complexes were incubated for 10 minutes, followed by addition of unlabelled competitor *sbmA* oligos in the amount of times the concentration of labeled oligo indicated. Mixtures with competitor DNA were incubated for 48 hours. All incubations were done at room temperature. Mixtures were run at 200 V on a 1x TBE-buffered 12% acrylamide gel for 45 minutes at 4°C. Gel bands were visualized by phosphor screen. Proportion of bound radioligand was fitted to a 3-parameter logistic curve for obtaining the IC<sub>50</sub>.

**Table 5-1 *In vivo* TraM-TraD tail interaction specificity between F and pED208 plasmids**

Complementing Proteins		Mating Efficiency (T/D) <sup>a</sup>
No TraM	TraD [F <sup>1-709</sup> :pED <sup>729-736</sup> ]	<1 x 10 <sup>-6</sup>
No TraM	TraD [F <sup>1-709</sup> :pED <sup>729-736</sup> .R734D]	<1 x 10 <sup>-6</sup>
F TraM	TraD [F <sup>1-709</sup> :pED <sup>729-736</sup> ]	1 x 10 <sup>-6</sup>
F TraM	TraD [F <sup>1-709</sup> :pED <sup>729-736</sup> .R734D]	3 x 10 <sup>-6</sup>
F TraM	F TraD	1 x 10 <sup>-1</sup>
TraM [F <sup>1-55</sup> :pED <sup>56-127</sup> ]	TraD [F <sup>1-709</sup> :pED <sup>729-736</sup> ]	1 x 10 <sup>-1</sup>
TraM [F <sup>1-55</sup> :pED <sup>56-127</sup> ]	TraD [F <sup>1-709</sup> :pED <sup>729-736</sup> .R734D]	<1 x 10 <sup>-6</sup>
TraM [F <sup>1-55</sup> :pED <sup>56-127</sup> ]	F TraD	<1 x 10 <sup>-6</sup>

<sup>a</sup> Determined by assaying donor ability of cells containing pOX38-MK3 or pOX38-DM and the complementing plasmid(s). T/D, transconjugants per donor. "<1 x 10<sup>-6</sup>" refers to no detectable donor ability.

This work was done by Jun Lu

**Table 5-2** Plasmids and oligonucleotides

Plasmid & oligos	Description & references
pACYC184	Tc <sup>r</sup> ; (Chang & Cohen, 1978)
pBAD24	Amp <sup>r</sup> ; (Guzman, <i>et al.</i> , 1995)
pBluescript	Amp <sup>r</sup> ; (Short, <i>et al.</i> , 1988)
pT7-5	Amp <sup>r</sup> ; (Tabor & Richardson, 1985)
pED208	(Falkow & Baron, 1962)
pJLEM200	pT7-5 with pED208 <i>traM</i> ; (Wong, <i>et al.</i> , 2011)
pJLM400	pBluescript KS (+) with F <i>traM</i> expressed from the lac promoter (Lu, <i>et al.</i> , 2004)
pJLM404	pJLM400 with <i>traM</i> from pED208; (Wong, <i>et al.</i> , 2011)
pJLM407	pJLM400 with <i>traM</i> hybrid F TraM1-55:pED208 TraM56-127; (Wong, <i>et al.</i> , 2011)
pACM404	pACYC184 with a <i>traM</i> ; (Wong, <i>et al.</i> , 2011)
pACM407	pACYC184 with a <i>traM</i> hybrid ; F TraM <sup>1-55</sup> :pED208 TraM <sup>56-127</sup> ; (Wong, <i>et al.</i> , 2011)
pJMTraD	pBAD24 with <i>traD</i> from F; (Lu, <i>et al.</i> , 2008)
pOX38-DM	Km <sup>r</sup> Cm <sup>r</sup> , TraD <sup>-</sup> TraM <sup>-</sup> derivative of pOX38-Km; (Lu, <i>et al.</i> , 2008)
pOX38-MK3	Km <sup>r</sup> Cm <sup>r</sup> , TraM <sup>-</sup> derivative of pOX38-Km; (Penfold, <i>et al.</i> , 1996)
pRFM200	pT7-5 with an F <i>BstBI</i> - <i>BglII</i> fragment from <i>traM</i> to <i>P<sub>FinP</sub></i> ; (Lu, <i>et al.</i> , 2003)
pRF911	Amp <sup>r</sup> ; F plasmid <i>sbmA</i> cloned in pBEND2; (Fekete & Frost, 2002)
JLU3	CTA TAG GGA GAC CGG AAT TCG
JLU262	GCC ATC CGT TAC CTG CAG G
JLU263	ATA TAT AAG CTT TCA GTA TTC CCT TCC GTC ATC CAT ATC CTC CCC GCG CTC C
JLU264	ATA TAT AAG CTT TCA GTA TTC ATC TCC GTC ATC CAT ATC CTC CCC GCG CTC C
JLU608	ATG GAT CCA CCA GAA CAT TCA AAG TG
JLU611	GAA GGA GGC TTT AAT CAG ATG GAG TAC AAC AAG CTC ATG CTG GAA AAC G
JLU612	CTC CAT CTG ATT AAA GCC TCC TTC CTC CAT CTG AGC CTC ATG TAC AC
30BTA	GAT ACC GCT AGG GGC GCT GCT AGC GGT GCG
30BTA	CGC ACC GCT AGC AGC GCC CCT AGC GGT ATC
pED 4site	AGA TTC GAA TCT AGA TTC GAA TCT

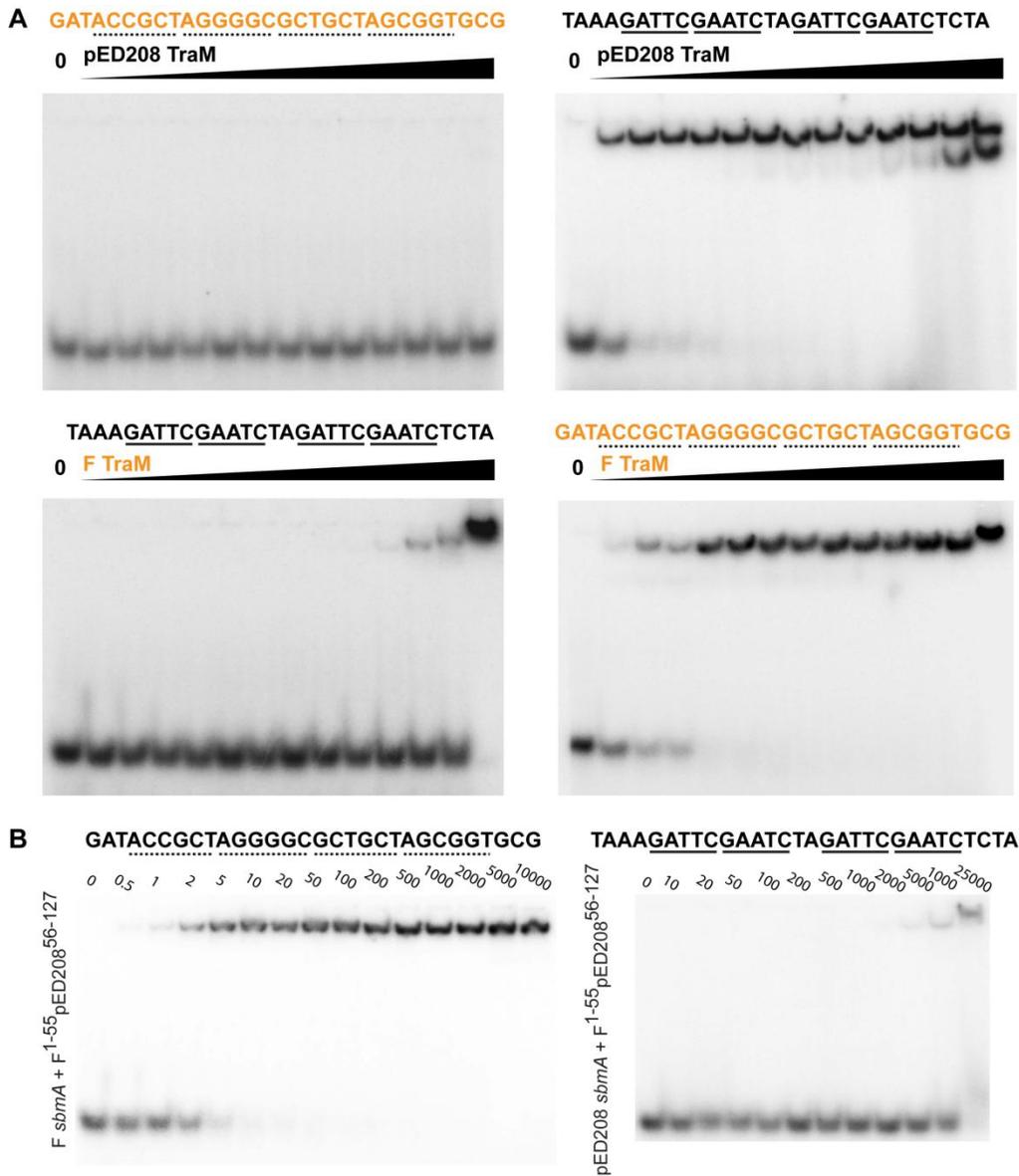


Figure 5-1 Specificity of F-like plasmid TraM in binding to *sbmA*

- a) Electrophoretic mobility shift assay (EMSA) of F and pED208 TraM binding to F or pED208 *sbmA*. DNA sequences from the F plasmid are coloured in orange. DNA sequences from the pED208 plasmid are in black. Concentrations in each lane are 0, 2, 5, 10, 20, 50, 100, 250, 600, 1000, 2500, 6000, 10000, and 25000 nM.
- b) EMSA of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] binding to F and pED208 *sbmA*. For F *sbmA*, concentrations in each lane are 0, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, and 10000 nM. For pED208 *sbmA*, concentrations in each lane are 0, 10, 20, 50, 100, 200, 500, 1000, 2000, 5000, and 10000 nM.

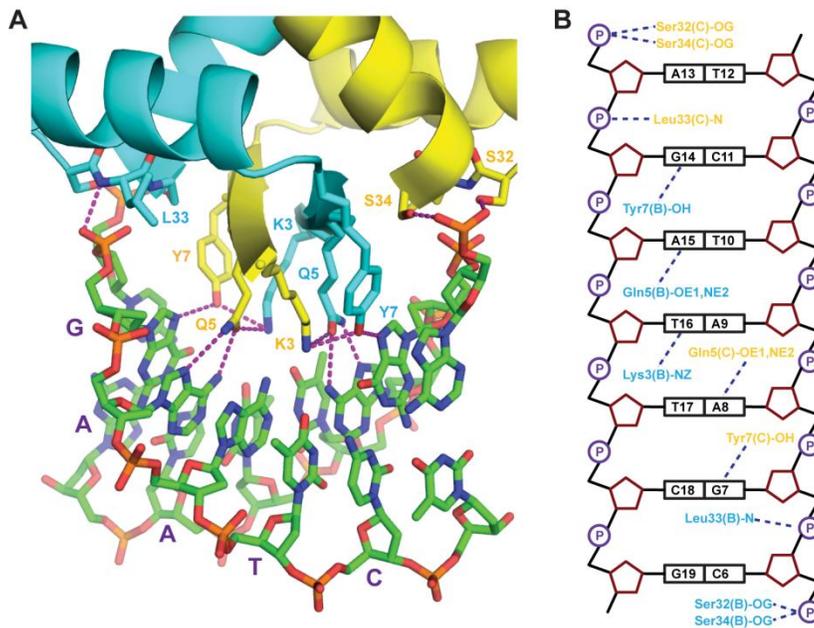


Figure 5-2 DNA contacts of the pED208 TraM N-terminal domain

- b) Interactions between the N-terminal domain of TraM and DNA. Hydrogen bonds are indicated by purple dashed lines. The GAATC binding motif consisting of bases G7 to T11 of *sbmA* is indicated with purple letters.
- c) Schematic diagram of TraM-*sbmA* interactions for one TraM tetramer. Hydrogen bonds are indicated with dashed lines.

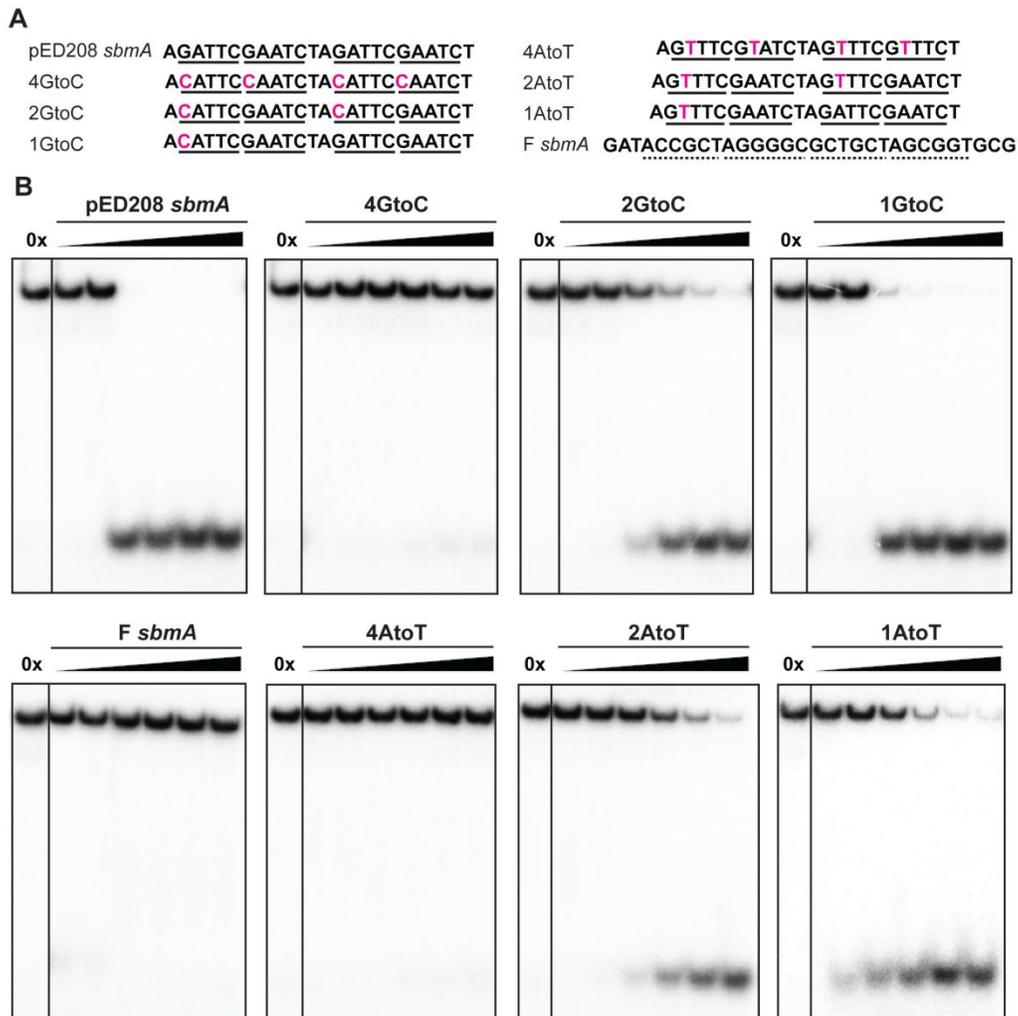


Figure 5-3 Role of residue-specific interactions in TraM-*sbmA* interaction  
a) *sbmA* oligos used in competition assays.  
b) Effect of mutation of G or A in the GANTC motifs of *sbmA* on competition of wild-type pED208 TraM from TraM-*sbmA* complexes. Oligos with mutations in similar numbers of GANTC motifs are aligned vertically for comparison. Amounts of competitor in each lane, as number of times labeled DNA are 0, 100x, 1000x, 2500x, 7500x, 15000x, and 25000x.

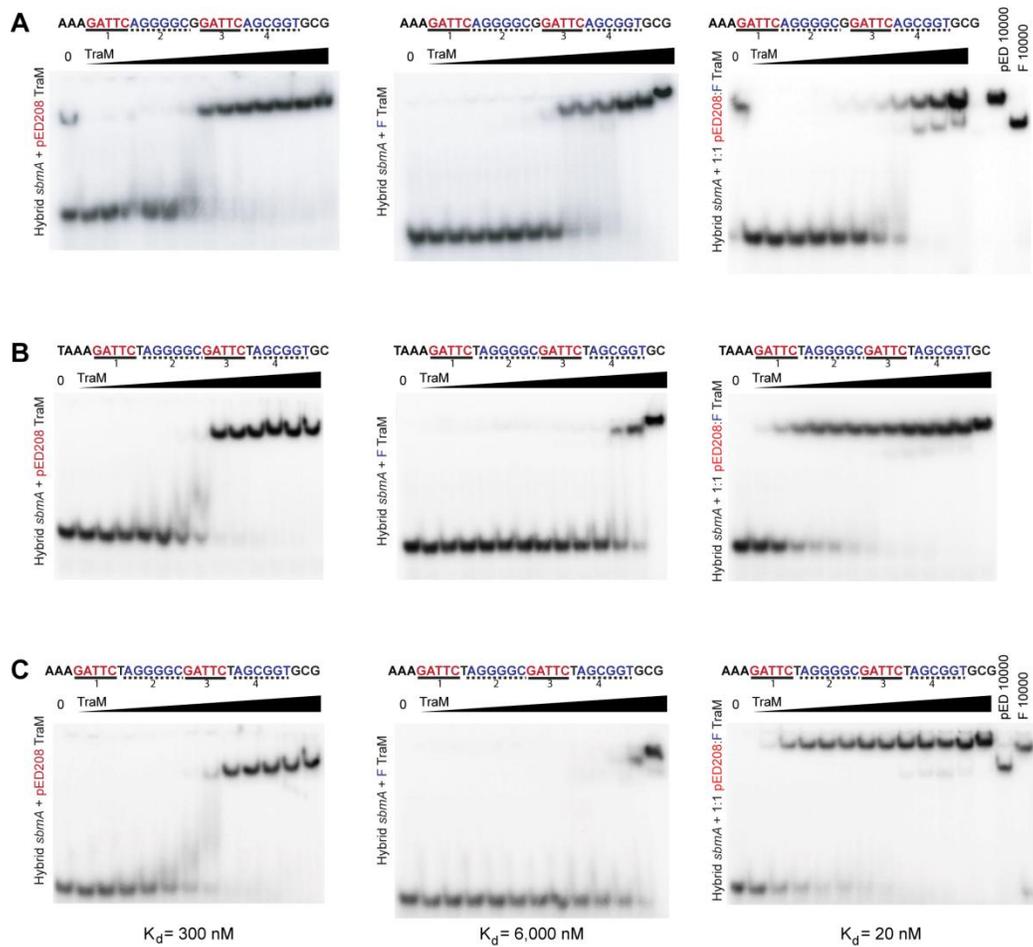


Figure 5-4 Cooperative binding of F TraM and pED208 TraM to a hybrid *sbmA*. Binding of pED208 TraM (left), F TraM (centre), and a 1:1 mixture of F and pED208 TraM (right) to a hybrid *sbmA* containing the GATC motifs of pED208 in positions 1 and 3 (highlighted in red), and the A(G/C)CG(G/C)T motifs of F in positions 2 and 4 (highlighted in blue) of *sbmA* were assessed by EMSA. TraM concentrations in each lane are 0, 2, 5, 10, 20, 50, 100, 250, 600, 1000, 3000, 10000, and 25000 nM.

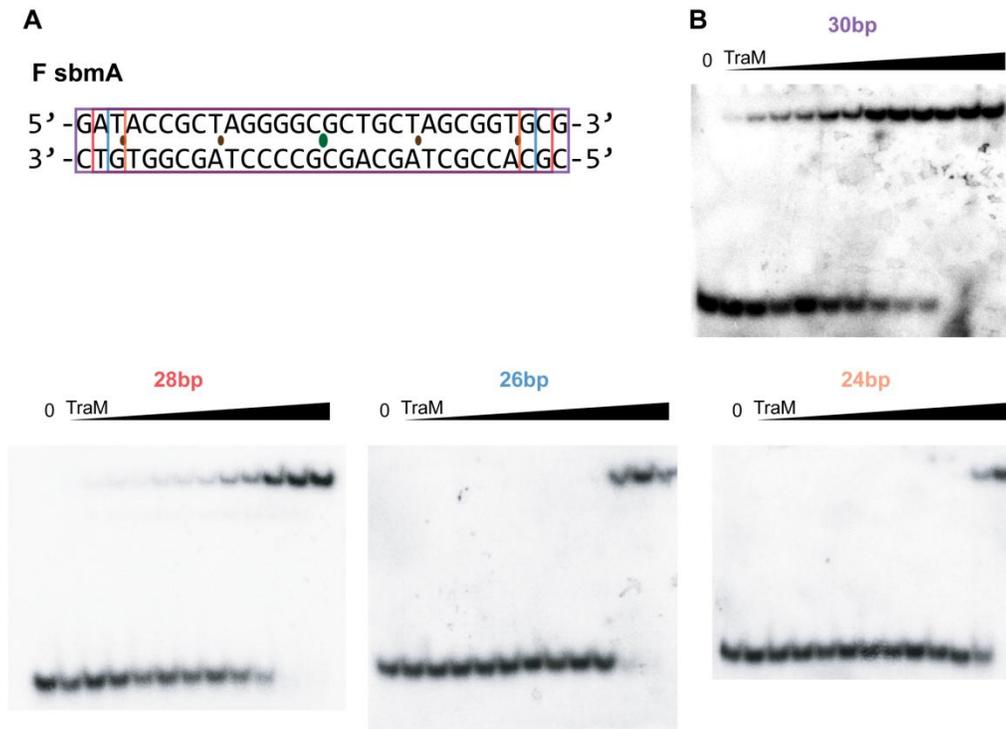


Figure 5-5 Minimum length of F *sbmA* required for F TraM binding with full affinity

a) F *sbmA* sequences used in minimal TraM binding length determination. Sequence lengths are indicated by boxes of the same colour scheme as b). Binding motif boundaries are indicated by brown ovals. The center of symmetry of the *sbm* site is indicated by a green oval. Mutated bases in 24bp ATAT to ensure crystallographic symmetry are boxed in yellow.

b) EMSA of F TraM binding to F *sbmA* sequences of varying lengths

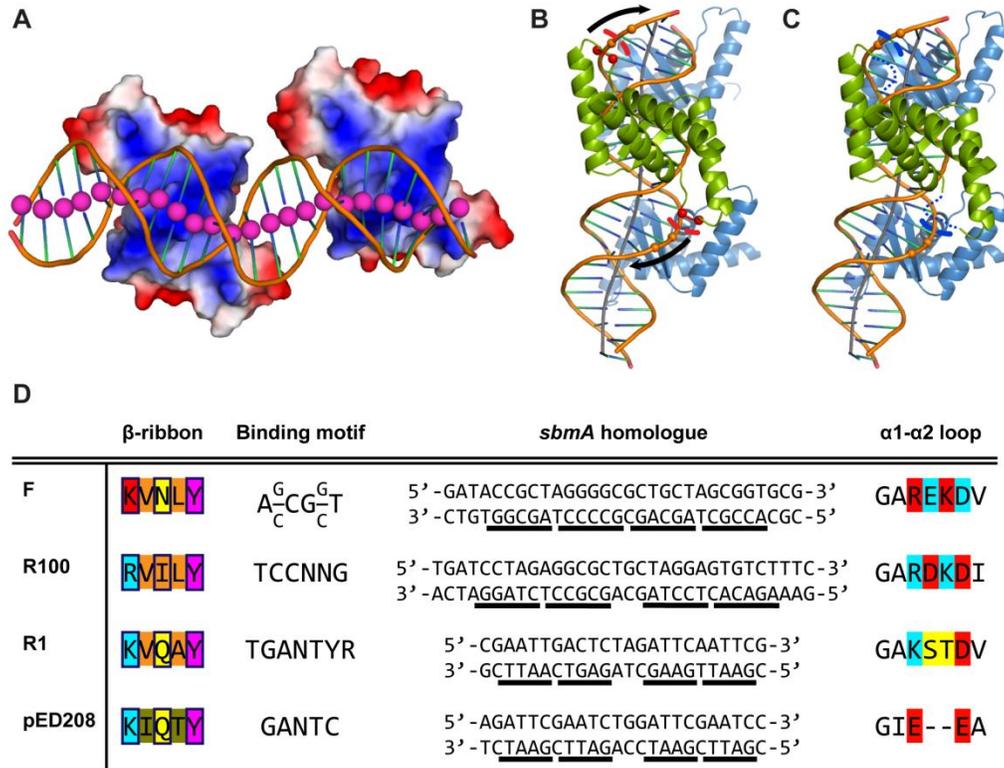


Figure 5-6 Possible specificity in TraM-*sbmA* interaction in from the TraM  $\alpha$ 1- $\alpha$ 2 loop

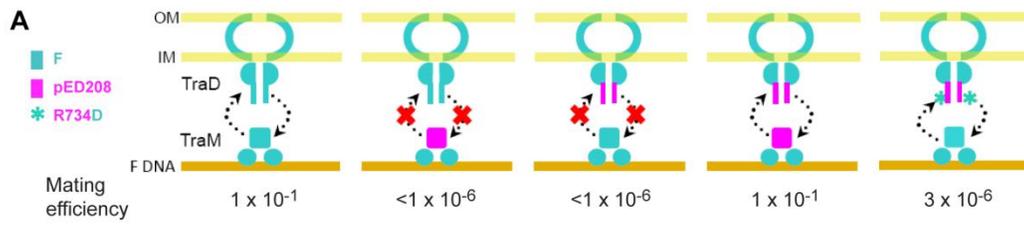
a) 2 RHH domains from one TraM tetramer bound to *sbmA*. The DNA helix axis (indicated by pink spheres) appears to be bent by attraction to the basic  $\beta$ -sheet surface and repulsion by the acidic loops between  $\alpha$ 1 and  $\alpha$ 2.

b) Kinking of *sbmA* DNA by the pED208 TraM  $\alpha$ 1- $\alpha$ 2 loop. Acidic residues Glu29 and Glu30 are shown by red spheres. The DNA axis is shown by a grey line.

c) Putative binding of F *sbmA* phosphate backbone by the F TraM  $\alpha$ 1- $\alpha$ 2 loop. The basic loop is shown by a blue dotted line.

d) Comparison of DNA-binding specificity determinants in F-like plasmids. Residues of the RHH  $\beta$ -sheet that contact DNA bases are boxed in dark blue.

Illustration in b) was prepared by Mark Glover



**B**

pED208	725	DKTHEMDDGREY	736
F	706	ERGEDVEPGDDF	717

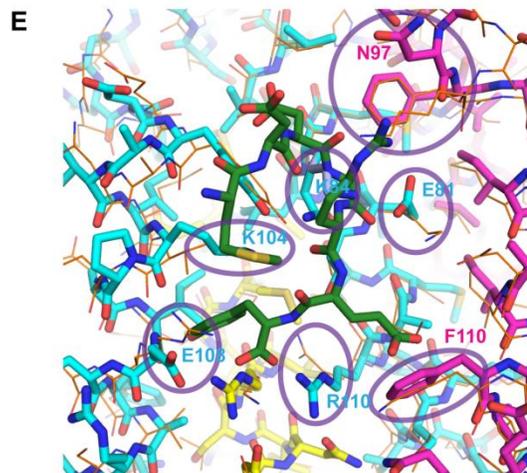
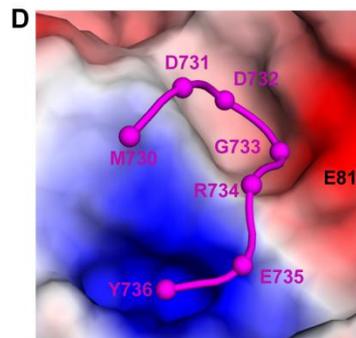
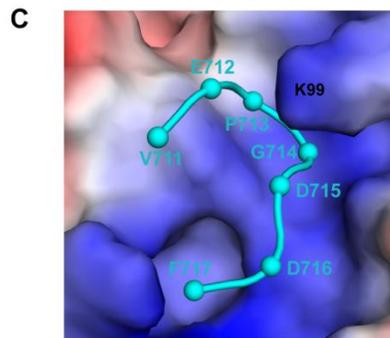


Figure 5-7 Conserved mechanisms of allelic specificity of TraM-*sbmA* and TraM-TraD interaction

- a) Double complementation of a TraM and TraD-deficient F-derived plasmid system by F and pED208 TraM and TraD mutants. F plasmid components are in cyan, pED208 components are in magenta.
- b) Sequence alignment of C-terminal tails of TraD. Conserved acidic residues are highlighted in red, basic residues in cyan, glycines in green, and aromatic residues in magenta. The residues that show charge exchange between pED208 and F are boxed.
- d) View of the TraD binding pocket in the F TraM C-terminal tetramer (Lu, *et al.*, 2008)
- e) Model of pED208 TraM C-terminal domain – TraD C-terminal tail interaction based on the F structure.
- f) Comparison of the TraD binding pocket of TraM in F and pED208. pED208 TraD residues are shown as green sticks. pED208 TraM residues are shown as cyan, pink, and yellow sticks to indicate different chains. F TraM residues are shown as orange lines. The pED208 Lys84 is circled in red, and its conformation in the crystal structure is shown in pale blue, while its conformation after manual modeling is shown in cyan. Non-conservative amino acid substitutions are circled in purple.

Work in a) was done by Jun Lu, and illustrations in a) was done by Mark Glover

mode	affinity	dist from best mode	
	(kcal/mol)	rmsd l.b.	rmsd u.b.
1	-7.6	0.000	0.000
2	-7.6	2.063	7.694
3	-7.5	2.137	8.464
4	-7.3	2.075	7.833
5	-7.2	2.574	9.050
6	-7.1	2.269	8.040
7	-7.1	2.004	8.120
8	-7.1	2.668	4.895
9	-7.0	2.181	7.913

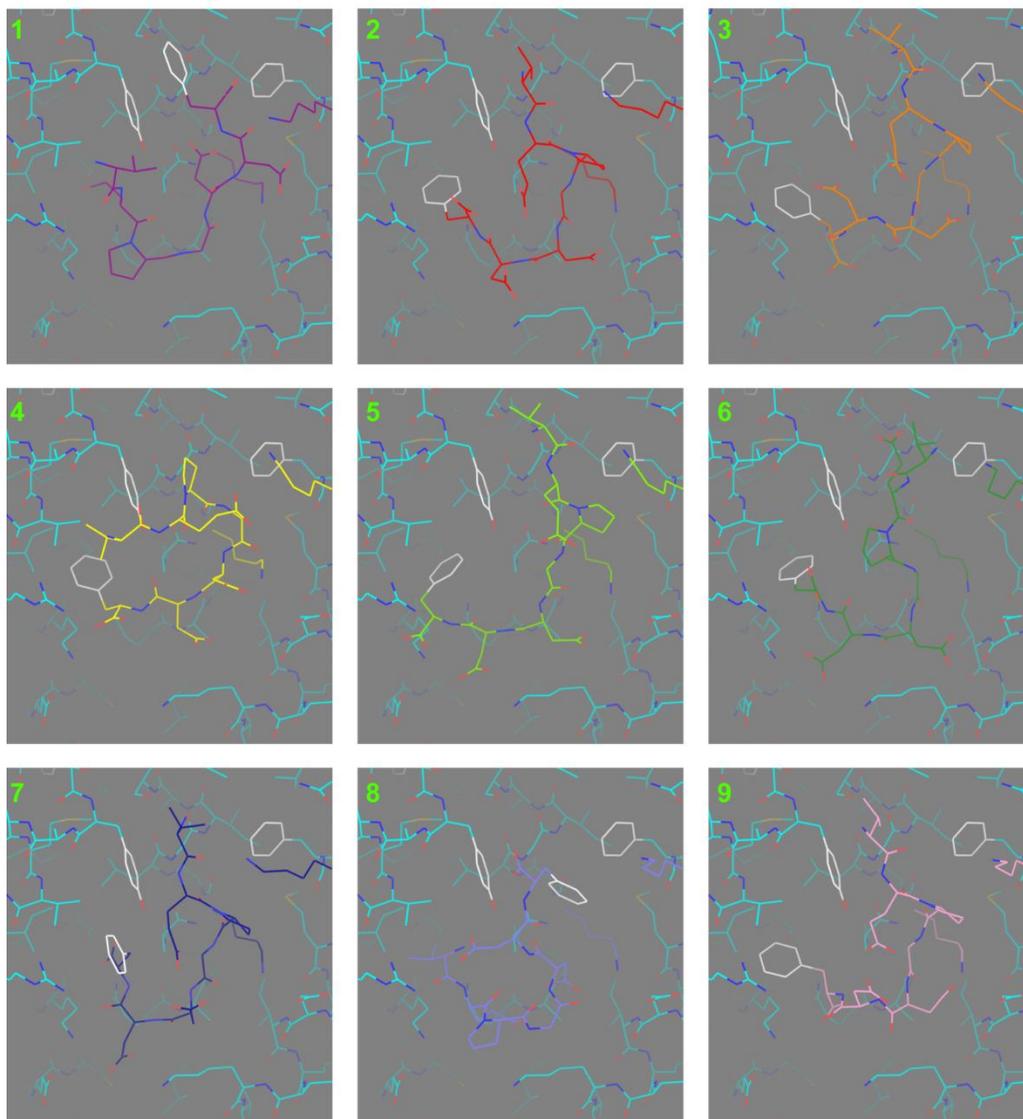


Figure 5-8 Top 9 conformations from docking of the F TraD tail to the F TraM tetramer in Autodock Vina

mode	affinity	dist from best mode	
	(kcal/mol)	rmsd l.b.	rmsd u.b.
1	-6.1	0.000	0.000
2	-5.9	2.812	7.720
3	-5.8	3.606	7.597
4	-5.8	2.676	7.669
5	-5.7	3.343	7.684
6	-5.7	3.052	7.877
7	-5.7	3.212	7.821
8	-5.6	3.797	8.949
9	-5.6	5.062	9.975

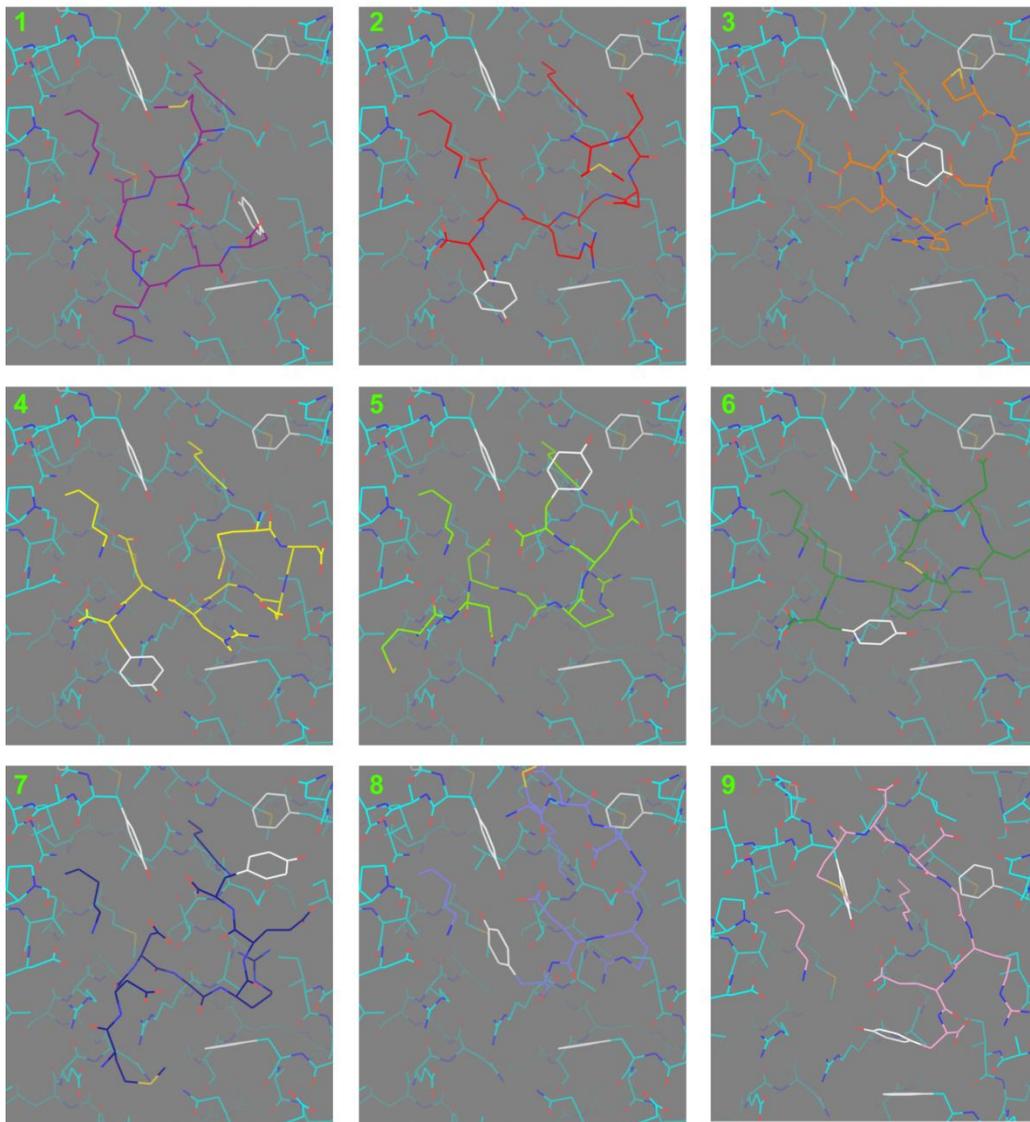


Figure 5-9 Top 9 conformations from docking of the pED208 TraD tail to the pED208 TraM tetramer in Autodock Vina

## References

- Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**: 775-808.
- Beranek A, Zettl M, Lorenzoni K, Schauer A, Manhart M & Koraimann G (2004) Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. *J Bacteriol* **186**: 6999-7006.
- Chang AC & Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**: 1141-1156.
- Falkow S & Baron LS (1962) EPISOMIC ELEMENT IN A STRAIN OF SALMONELLA TYPHOSA. *J Bacteriol* **84**: 581-589.
- Fekete RA & Frost LS (2000) Mobilization of chimeric oriT plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. *J Bacteriol* **182**: 4022-4027.
- Fekete RA & Frost LS (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.
- Geist C & Brantl S (2008) TraM protein of plasmid R1: in vitro selection of the target region reveals two consensus 7 bp binding motifs spaced by a 4 bp linker of defined sequence. *Plasmid* **59**: 20-35.
- Guzman LM, Belin D, Carson MJ & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Ho SN, Hunt HD, Horton RM, Pullen JK & Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**: 51-59.
- Kupelwieser G, Schwab M, Hogenauer G, Koraimann G & Zechner EL (1998) Transfer protein TraM stimulates TraI-catalyzed cleavage of the transfer origin of plasmid R1 in vivo. *J Mol Biol* **275**: 81-94.
- Lang S, Kirchberger PC, Gruber CJ, *et al.* (2011) An activation domain of plasmid R1 TraI protein delineates stages of gene transfer initiation. *Mol Microbiol* **82**: 1071-1085.
- Lang S, Gruber K, Mihajlovic S, *et al.* (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol Microbiol* **78**: 1539-1555.
- Lu J, Fekete RA & Frost LS (2003) A rapid screen for functional mutants of TraM, an autoregulatory protein required for F conjugation. *Mol Genet Genomics* **269**: 227-233.
- Lu J, Zhao W & Frost LS (2004) Mutational analysis of TraM correlates oligomerization and DNA binding with autoregulation and conjugative DNA transfer. *J Biol Chem* **279**: 55324-55333.
- Lu J, Wong JJ, Edwards RA, Manchak J, Frost LS & Glover JN (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.
- Lu J, Manchak J, Klimke W, Davidson C, Firth N, Skurray RA & Frost LS (2002) Analysis and characterization of the IncFV plasmid pED208 transfer region. *Plasmid* **48**: 24-37.
- Miller DL & Schildbach JF (2003) Evidence for a monomeric intermediate in the reversible unfolding of F factor TraM. *J Biol Chem* **278**: 10400-10407.

- Penfold SS, Simon J & Frost LS (1996) Regulation of the expression of the traM gene of the F sex factor of Escherichia coli. *Mol Microbiol* **20**: 549-558.
- Sastre JI, Cabezon E & de la Cruz F (1998) The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J Bacteriol* **180**: 6039-6042.
- Schwab M, Reisenzein H & Hogenauer G (1993) TraM of plasmid R1 regulates its own expression. *Mol Microbiol* **7**: 795-803.
- Short JM, Fernandez JM, Sorge JA & Huse WD (1988) Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Res* **16**: 7583-7600.
- Tabor S & Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* **82**: 1074-1078.
- Trott O & Olson AJ (2009) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**: 455-461.
- Wong JJ, Lu J, Edwards RA, Frost LS & Glover JNM (2011) Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**: 6775-6788.

## Chapter 6 Concluding Remarks

### *Summary of findings*

Plasmids are able to offer benefits to the host, like antibiotic resistance and enhanced virulence, while maximizing their own propagation. One way in which plasmids ensure their own survival is through specific interactions between components of the conjugation machinery and their target plasmid DNAs. This is a mechanism that allows transfer of only the cognate plasmid DNA. The phenomenon of plasmid specificity among conjugative components has long been observed for many plasmids *in vivo*. Work by other groups has revealed the structural basis for specificity of relaxase binding to DNA from its cognate plasmid (Datta, *et al.*, 2003, Guasch, *et al.*, 2003). Our work provides insight into the structural basis of specificity in F-like plasmids at two additional levels of interaction -TraM-TraD interaction and TraM-*sbmA* interaction.

Crystallization of the C-terminal tetramer of TraM with the C-terminal peptide of TraD showed that contact between the two proteins involves specific interactions between their amino acid side-chains. The nature of the interaction is largely through electrostatic interactions, with additional important contributions from the C-terminal phenylalanine aromatic side chain and the free carboxylate. Amino acid substitutions in the pED208 TraM binding pocket and TraD tail allow pED208 TraM to distinguish its cognate TraD from that of F, which was demonstrated *in vivo*. The importance of the last 8 amino acids was shown directly to be important in TraD binding affinity to TraM using an *in vitro* pulldown assay. Most residues that show interaction with TraM through the crystal

structure have an *in vivo* effect on conjugation if mutated, indicating the interactions observed in the crystal are important for physiological function.

Some binding could still be observed in the pulldown assay for the C-terminal tail truncations of the full-length cytoplasmic domain. The exact regions of TraD outside of the C-terminal tail that bind TraM have not been determined. Additional evidence for the additional TraD interaction with TraM comes from demonstration that the TraD cytoplasmic domain is able to bind to TraM, and with only ~10-fold less binding affinity, with truncation of the last 38 amino acids. *In vivo* mating assays suggest that there are no further residues in the C-terminal extension beyond the ATPase homology domain (577-717) that contact TraM, as deletions in this region do not result in any additional decrease in conjugation efficiency (Sastre, *et al.*, 1998), making it more likely that the ATPase domain is involved. More needs to be done to elucidate the other regions of TraD that interact with TraM.

The crystal structure of pED208 TraM bound to DNA provides confirmation that the DNA-binding domain of TraM is an RHH fold, and reveals the basis of specificity in TraM-DNA interaction among F-like plasmids. Sequence analysis shows that the RHH  $\beta$ -sheet residues that contact the major groove are unique among 4 commonly studied F-like plasmids, with the exception of pED208 and R1. These two plasmids share the same  $\beta$ -sheet residues, which is reflected in the similarity of their DNA binding motifs. *In vitro* experiments show that F TraM binding is specific for the F *sbmA* binding motif, and greatly diminished for that of pED208 or R100. Further confirmation that the  $\beta$ -sheet residues impart specificity in DNA binding was shown with mutation of the  $\beta$ -

sheet residues in R100 TraM to that of F, which subsequently increased its ability to bind to F DNA. These results were further confirmed *in vivo* by complementation mating assays.

Our work also provides insight into the mechanism of cooperative TraM binding to *sbmA*. The cooperativity of *sbmA* binding between 2 TraM tetramers was shown with EMSA assays. The arrangement of 4 binding motifs in close proximity is unique to TraM and results in high affinity binding of both tetramers simultaneously, in contrast to the binding of a single tetramer. The crystal structure shows that strikingly, the cooperativity between the TraM tetramers occurs without any protein-protein interaction. Cooperativity is entirely mediated through DNA, via DNA kinking derived from electrostatic repulsion of the phosphate backbone from the alpha 1-2 loop acidic residues, and through underwinding that possibly minimizes the energetic unfavorability from the steric restrictions imposed by the limited linker length and unwinding of the  $\alpha 2$  helix end. An EMSA experiment using hybrid DNA containing binding sites for both F and pED208 binding sites confirms that the mechanism of cooperative binding to *sbmA* is conserved between F and pED208.

While a fair number of other proteins are also able to bind cooperatively to DNA without protein-protein interaction, TraM has the unique feature that all monomers required to form the cooperatively-binding oligomer interact to form a symmetrical, tetrameric helical bundle, instead of a dimer of dimers as is usually seen between cooperative pairs of RHH folds. The advantages or reasons for this are unknown at this time. It may be that availability of 4 faces of the tetramer may be needed for interacting with the TraD hexamer(s) in a complex oligomeric

arrangement. With the dimer-of-dimer arrangement, only similar faces are available.

It appears that there is more specificity with regards to TraM-*sbmA* binding than there is for TraM-TraD tail interaction, at least among the more well-studied F-like plasmids. In fact, the sequences of the R1, R100, and F TraD C-terminal tails are exactly the same, and F TraD can substitute for that of R1 in binding to R1 TraM (Beranek, *et al.*, 2004). This may reflect the function of the C-terminal tail interaction, which may be to provide only initial tethering of TraM to TraD. Other TraD regions, or interactions with other factors like that which is strongly suggested between TraD and Tral (Lang, *et al.*, 2010, Lang, *et al.*, 2011), may supply additional specificity. Another possibility is that the specificity between TraM and DNA is sufficient to prevent recruitment of incompatible relaxosomes to the conjugative pore.

#### *Future directions*

Many characteristics of the TraD interaction with TraM are still unknown. The stoichiometry of binding between TraM and TraD oligomers has yet to be clarified. The TraM tetramer is able, in theory, to bind up to 4 TraD C-terminal tails simultaneously, but whether it actually does, and even if all chains need to come from the same hexamer has yet to be shown. Which *sbm* sites are involved in TraM-TraD binding is also still unknown.

The regions of that TraD interact with TraM besides the TraD tail are another unknown factor. It appears likely that these other regions are within the ATPase domain, as further deletion of the TraD tail did not lead to additional loss

of conjugation frequency (Sastre, *et al.*, 1998). One possible approach to detect an interaction is to perform the same pulldown assay used to test for interaction with a TraD<sup>137-717</sup> and TraD<sup>645-717</sup> using different TraD constructs. However, from the crystal structure of TrwB, it appears unlikely that substantial portions of the TraD ATPase domain can be truncated without disrupting the structural integrity of the fold. Another option would be to use a homology modelling program like MODELLER (Eswar, *et al.*, 2008) to create a model for the structure of TraD, and choose individual residues predicted to be on surfaces which may interact with TraM for site-directed mutagenesis. These mutants could then be tested for potential TraM function quickly through the *in vivo* conjugation assay, where the mutants would be tested for their ability to complement a TraD knockout F-plasmid derivative. Mutant residues that cause conjugative defect may also result from disruption of interaction with Tral or other factors. Therefore, to determine if TraM binding is disrupted, site-directed mutants of the His<sub>6</sub>-TraD<sup>137-717</sup> would be made and tested for their ability to bind TraM in a pulldown assay followed by Western blot detection of TraM.

It remains a puzzle how F TraM interacts with its *sbmA* site. There is likely to be asymmetrical contributions from each B-sheet chain, due to the variation and lack of palindromic symmetry in the binding motif. The possibility of a role for the  $\alpha$ 1- $\alpha$ 2 loop in interacting with the phosphate backbone is intriguing and supported by the increased length of F *sbmA* required for maximal binding to TraM compared to pED208. Such questions would be answered directly with an atomic model of the F TraM N-terminal domain-*sbmA* complex. Work up to this point using the strategy of crystallizing TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] or TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] A18D K22I has been unsuccessful in obtaining diffracting crystals. I propose

that this work may be continued by cloning a version of the hybrid TraM with the wild-type linker length, as outlined in the discussion of Appendix 2.

While the cooperative mechanism and specificity of TraM binding to *sbmA* is much clearer, the role of TraM-*sbmA* binding in relaxosome structure and function is still largely unknown. *sbmA*-like sites likely have a key role in relaxosome function, as they are present across plasmids F, R1, R100, and pED208 (Chapter 5), but precisely what remains open to speculation. The crystal structure confirms earlier circular permutation assays showing that TraM does not impose an overall bend on the DNA (Fekete & Frost, 2002), so it does not contribute to relaxosome structure in that respect. The presence of TraM induces negative supercoiling of *oriT*-containing plasmids by whole linking numbers, much more than would be expected from the degree of unwinding seen in the crystal structure (Mihajlovic, *et al.*, 2009). However, TraM has been shown to aggregate non-specifically on DNA at high concentrations, which may yield additional underwinding. TraM may function in an analogous fashion to ParB of the P1 plasmid and its homologues. ParB has a specific DNA binding site, *parS*, and is capable of polymerizing along the DNA for long distances of up to several kilobases. Like TraM, ParB alters the topology of the plasmid DNA to which it binds by changing the degree of supercoiling (Lobocka & Yarmolinsky, 1996). Another possible outcome of TraM polymerization on DNA has been proposed, the induction of nucleosome-like structure similar to TraK of the plasmid RP4 (Di Lorenzo, *et al.*, 1991, Fekete & Frost, 2002). Electron microscopy of TraM on F DNA has indicated that TraM shortens the DNA but does not induce a significant bend, which is consistent with this idea (Di Lorenzo, *et al.*, 1992, Fekete &

Frost, 2002). More work is needed to elucidate the role of TraM in the bigger picture of relaxosome function.

## References

- Beranek A, Zettl M, Lorenzoni K, Schauer A, Manhart M & Koraimann G (2004) Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. *J Bacteriol* **186**: 6999-7006.
- Datta S, Larkin C & Schildbach JF (2003) Structural insights into single-stranded DNA binding and cleavage by F factor Tral. *Structure* **11**: 1369-1379.
- Eswar N, Eramian D, Webb B, Shen MY & Sali A (2008) Protein structure modeling with MODELLER. *Methods Mol Biol* **426**: 145-159.
- Fekete RA & Frost LS (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.
- Guasch A, Lucas M, Moncalian G, *et al.* (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Biol* **10**: 1002-1010.
- Lang S, Kirchberger PC, Gruber CJ, *et al.* (2011) An activation domain of plasmid R1 Tral protein delineates stages of gene transfer initiation. *Mol Microbiol* **82**: 1071-1085.
- Lang S, Gruber K, Mihajlovic S, *et al.* (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol Microbiol* **78**: 1539-1555.
- Lobocka M & Yarmolinsky M (1996) P1 plasmid partition: a mutational analysis of ParB. *J Mol Biol* **259**: 366-382.
- Mihajlovic S, Lang S, Sut MV, *et al.* (2009) Plasmid r1 conjugative DNA processing is regulated at the coupling protein interface. *J Bacteriol* **191**: 6877-6887.
- Narra HP & Ochman H (2006) Of what use is sex to bacteria? *Curr Biol* **16**: R705-710.
- Sastre JI, Cabezon E & de la Cruz F (1998) The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J Bacteriol* **180**: 6039-6042.

## Appendix A

### Investigation of pED208 TraM binding to longer *oriT* fragments

#### Introduction

The arrangement of relaxosome protein binding sites at *oriT* is very different for each of the more well-studied F-like plasmids. For example, R1 has two *sbm* sites at *oriT* (Schwab, *et al.*, 1991) , and R100 has four (Abo & Ohtsubo, 1995). Therefore, conservation of a mechanistic feature of TraM binding would be remarkable and speak to its importance in relaxosome function. To date, in the literature, TraM binding cooperativity between *sbm* sites has only been investigated in the F plasmid. In F, TraM binds cooperatively to its 3 binding sites at *oriT*, *sbmA*, *sbmB*, and *sbmC*. Cooperativity also exists between *sbmA* and *sbmB*, and within *sbmA* (Fekete & Frost, 2002).

The arrangement of relaxosome protein binding sites at *oriT* in pED208 is similar to that of F in that there are 3 *sbm* sites, *sbmA*, *sbmB*, and *sbmC*, where *sbmA* is the highest affinity site. However, the precise arrangement of the sites is different from F. In order from 5' to 3', the sites are arranged *sbmBAC* in pED208 (Figure A-1A), whereas in F they are arranged *sbmABC* (Figure 1-1A). The arrangement of IHF binding sites between the *sbm* sites is also different. Unlike the F plasmid, both IHF sites at *oriT* are interspersed between the *sbm* sites. Binding of pED208 TraM to the other *sbm* sites, as well as combinations of *sbm* sites on longer fragments of DNA was investigated by EMSA in order to see if like the F plasmid, cooperativity of binding within and between *sbm* sites (Fekete & Frost, 2002) exists in pED208.

## Results

Instead of 3 distinct shifted species corresponding to each *sbm* site in *sbmABC*, there is a gradual shifting of band migration with increasing TraM with no distinct species (Figure A-1B). The bands are also fairly smeared, which may indicate conformational instability. This result is similar to that obtained previously by EMSA titration of pED208 TraM onto an *sbmABC*-containing DNA fragment (Di Lorenzo, *et al.*, 1991). With *sbmAB*, there is a distinct shifted species with a  $K_d$  of approximately 40 nM TraM, presumably due to loading of TraM onto *sbmA*. However, the shift is not clean as the pED208 *sbmA* fragment (background smears can be seen in the lane between the unbound DNA and the first shift). A second shifted species, presumably due to loading of TraM into *sbmB*, is visible and begins to form at 4 nM TraM. The majority of the second shifted species occurs at  $\mu$ M TraM concentrations, and with highly smeared intermediate species (Figure A-1C). Determining of cooperativity between *sbmA* and *sbmB* is difficult because the mobility of the first and second shifted species is not constant, unlike the F plasmid, calling into question whether a Hill plot would be valid using the current data.

As a positive control to see if cooperativity of F TraM-*oriT* binding could be replicated under experimental conditions current at the time, I performed titration of F TraM onto *sbmAB* or *sbmABC* and subjected them to EMSA with the same protocol as the pED208 samples. Results from (Fekete & Frost, 2002) were replicated with both *sbmAB* and *sbmABC* (Figure A-2A and A-2B), indicating that cooperativity is observable under the experimental conditions at the time. All of the shifted species formed distinct, sharp bands of similar mobility, with the exception of the final shifted species of *sbmABC*. The first and

second shifts of *sbmAB* where 50% of the higher mobility *sbmAB*-containing species was bound occurred at ~5 nM and 15 nM TraM, comparable to ~10 nM and 20 nM observed in (Fekete & Frost, 2002). For *sbmABC*, these shifts occurred at ~7 nM and 15 nM, which is comparable to ~2 nM and 9 nM previously observed (Fekete & Frost, 2002). Therefore, the decreased binding affinity and instability of the shifted species in pED208 TraM-*oriT* binding relative to those of the F plasmid is valid.

Titration of pED208 TraM onto *sbmAC* (Figure A-1D) yielded similar results to titration onto *sbmAB*. In fact, the smearing of the intermediate species between shifts is greater, and the degree of migration of the second shifted species is even less uniform than that of *sbmAB*. The first shift occurs at approximately 40 nM, and no distinct second shifted species occurs at the highest protein concentration tested, 2  $\mu$ M.

Titration of pED208 TraM onto individual *sbmB* (Figure A-1E) and *sbmC* (Figure A-1F) sites failed to yield a clean, distinct shifted species. The lack of a distinct shifted species and presence of species of intermediate migration indicates that it is doubtful that both sites within the *sbmB* or *sbmC* species are being loaded simultaneously, unlike with *sbmA*. In conclusion, these results indicate that there is no observable cooperativity in TraM binding within other *sbm* sites besides *sbmA*, and between any *sbm* sites in pED208 *oriT*.

#### *Binding of TraM to IHF-ihfC-sbmA complexes*

The lack of stable, cooperative loading of TraM onto longer *oriT* fragments led us to hypothesize that IHF may be needed for cooperativity to

occur. To test this, preliminary studies were performed to see if a stable complex could be formed between a DNA fragment containing the *ihfC* and *sbmA* binding sites (Figure A-3A) as identified in (Di Lorenzo, *et al.*, 1995). In addition, if stable complexes could be formed, crystallization of the complexes would have been attempted to obtain clarification of the 3-dimensional arrangement of proteins at the relaxosome, for the purpose of crystallization. The role of IHF is particularly intriguing because it is known to impose a sharp, effectively 180° turn in the DNA, and in addition, there are poly-AT tracts on either side of the *sbmA* site which have a known tendency to form intrinsic bends. Binding of TraM and IHF to the *ihfC-sbmA* fragment was investigated by EMSA.

TraM bound to the 66bp *ihfC-sbmA* fragment in a manner much different from its binding to 24-30bp *sbmA* fragments. Instead of high-affinity binding in the nanomolar range, and cooperativity between the two tetramers as observed with a single shifted species, TraM shifted the unbound *ihfC-sbmA* at a much higher concentration and with broadly-smeared intermediate species over a wide concentration range (Figure A-3B). No distinct protein-DNA complex was formed until ~500 nM, whereas it forms as soon as unbound DNA begins to be shifted in titration of TraM onto the 24bp *sbmA* fragment (Figure 4-10A). Whereas differences in how the pED208 TraM binds to longer *oriT* fragments were already apparent in the *sbmAB* and *sbmABC* fragments, the differences became especially apparent with this experiment. In F *sbmA*, additional stretches of DNA 150 base pairs in length flanking the *sbmA* site did not have an effect on  $K_d$  (Fekete & Frost, 2002) compared to the 30 base pair minimal *sbmA* fragment (Figure 4-3C). This effect is not a result of unique characteristics

of the *sbmA* flanking sequences, as very similar effects are seen when the flanking sequences on either side of *sbmA* are mutated to random bases (Figure A-3B). Nevertheless, titration of IHF onto *sbmA* was performed. A concentration of 5 $\mu$ M IHF was chosen for IHF-*ihfC-sbmA* complex formation, as free DNA was completely bound. A distinct IHF-*ihfC-sbmA* complex forms prior to very slowly migrating ones at higher concentrations of a few micromolar (Figure A-3C). Titration of TraM onto those complexes resulted in a shifted species forming, but the stability and physiological relevance of the species is in doubt as there is a significant difference in the apparent  $K_d$  compared to that of minimal *sbmA*, and in the variable mobility of the shifted species. After 3 hours of electrophoresis, the IHF-*ihfC-sbmA* complexes themselves showed variable migration, in addition to the TraM- IHF-*ihfC-sbmA* complexes (Figure A-3D).

A possible explanation for these results is that the *ihfC* and *sbmA* sites cannot be occupied simultaneously. The TraM-IHF-*ihfC-sbmA* complex was modelled by manual placement of the DNA of the TraM-*sbmA* complex onto the DNA bound to the IHF-DNA complex crystal structure (Rice, *et al.*, 1996), while maintaining base-stacking and backbone geometry as well as possible given the amount of distortion in TraM-bound *sbmA*, and taking into account the number of base pairs between the DNA bound by IHF and the GANTC binding motifs in the pED208 *oriT* region. The center of the IHF binding site was presumed to be the center of the region protected by IHF, since the crystal structure of the IHF-DNA complex showed that bending of DNA around the protein is symmetrical. In the resulting model there is a steric clash between a TraM RHH domain and the IHF protein (Figure A-3E).

### *Effect of DNA length on binding affinity of pED208 TraM to sbmA*

To determine the effect of *sbmA*-flanking DNA length on pED208 TraM binding to *sbmA*, *sbmA* fragments were constructed that had varying lengths of flanking DNA on either side of the *sbmA* binding site (Figure A-4E). These fragments, ranging from 30, 40, 48, and 56 base pairs in length were compared in terms of TraM binding affinity and characteristics of the 66 base pair *ihfC-sbmA* fragment pED208 TraM. As the flanking sequences get longer, the binding affinity of TraM for the *sbmA*-containing fragment decreases (Figure A-4B). These results confirm the dependence of TraM binding affinity on the length of adjacent DNA. These results also show that the migration and stability of the pED208 TraM-*sbmA* complex is highly dependent on the conditions under which the gel was run. The gels for the length dependence experiment were run on 8% minigels for 45 min, while the preceding *ihfC-sbmA* gels were run on 6% full-sized gels for 3 hrs. The smaller gels show a distinct shifted species of constant migration with no smeared intermediates, even for for the same 66 bp *ihfC-sbmA* fragment. While F plasmid TraM-*sbmA* complexes are stable while run on full-sized gels, pED208 traM-*sbmA* complexes are not. The reasons for the relative instability of pED208 TraM-*sbmA* complexes while run on full-size gels are unknown at this time.

### **Discussion**

*Difference in stability and cooperativity of pED208 TraM on longer oriT fragments relative to the F plasmid.*

The inter-site cooperativity observed in the F *sbm* sites was not replicated in pED208. While the reasons for the reduced binding affinity and lack of cooperativity in pED208 TraM binding to longer *oriT* fragments are not

apparent, a possible explanation of why IHF did not assist in TraM binding may be illustrated through modeling of the IHF-TraM-*ihfC-sbmA* complex (Figure A-3E). There is a steric clash between the IHF protein and the ribbon-helix-helix domain of one of the TraM tetramers bound to *sbmA*, meaning that it is not possible for both of them to bind to their respective DNA binding motifs at the same time in their previously determined conformations. Upon closer inspection of the DNaseI protection analysis of pED208 TraM and IHF on full-length *oriT* DNA, it suggests that the *ihfC* and *sbmA* sites are in fact not occupied at the same time (Di Laurenzio, *et al.*, 1995). Similar regions of protection for *sbmA* and *sbmB* are shown in the presence of TraM alone or with TraM and IHF. However, the region protected by IHF only appears when IHF alone is incubated with the fragment, and is susceptible to DNaseI digestion when both TraM and IHF are present. It appears likely that TraM competes with IHF for access to adjoining *sbmA* binding sites, and that the *ihfC* site is either non-physiologic or occupied in a sequentially different stage of relaxosome formation than when TraM is present.

Competition of IHF off the *ihfC-sbmA* fragment by TraM may also explain the aberrant migration of the IHF-*ihfC-sbmA* complex during TraM titration. Instead of the mobility of the IHF-DNA complex remaining stable followed by shifting to a slower migrating complex with increasing TraM, the mobility of the IHF-DNA complex actually increases with increasing TraM until its concentration reaches 5 nM, then decreases as TraM increases, until most of the IHF-DNA is shifted to a defined complex at 50 nM TraM. Due to the sharp U-turn in the DNA imposed by IHF binding, the migration of the DNA may well increase as IHF is competed off by TraM, as bent DNA migrates slower than linear DNA. This may

occur as 2 bends on the same DNA fragment are bent in “cis” rather than in opposite directions “in trans” (Zinkel & Crothers, 1987), or if the bend is moved closer to the ends of the DNA from the middle (Zwieb & Adhya, 2009). Since TraM imposes no overall bend while bound to its *sbmA* site, as confirmed by our crystal structure of the complex, it appears more likely that TraM binding is preventing IHF from imposing any further bend in the DNA than what may already be intrinsic to the poly-AT tracts at and around its binding site.

## **Materials and Methods**

### *Expression and purification of IHF*

All expression and purification protocols and the IHF expression plasmid (IHF cloned into pET21a) were kindly provided by Phoebe Rice.

BL21-DE3 (Stratagene)  $RbCl_2$  competent cells were transformed with the IHF expression plasmid. These cells were grown in LB media containing 100ug/mL ampicillin at 37°C until the OD 600 reached ~0.7, followed by induction with 0.5 mM IPTG for 5 hours at 27°C.

Cells from 1 L of cell culture were suspended in 200 ml Lysis Buffer (100 mM Tris, 1 mM EDTA, 10% sucrose, 10% glycerol, 1 M NaCl) with 2 Complete EDTA protease inhibitor tablets (Roche). Lysozyme was added to 0.2 mg/mL and the cell suspension was stirred in ice for 15min. The cells were lysed by sonication for 20 seconds 6 times, followed by centrifugation at 39,000 x g for 40 minutes to remove cell debris. Ammonium sulfate was added to 50% saturation to the clarified cell lysate and stirred for 15 min on ice. Precipitated protein was pelleted by centrifugation at at 39,000 x g for 30 minutes. Ammonium sulfate

was added to 80% saturation to the supernatant, which was stirred for 15 min on ice, followed by centrifugation as in the previous step.

The 80% ammonium sulfate pellet was dissolved in 50 ml of Buffer A (20 mM MES-0.1 mM EDTA-5% Glycerol, pH 5.5) and loaded onto a 5 ml heparin column (5 mL Hi-Trap™ Heparin HP, GE Healthcare Life Sciences) in 20% Buffer B (20 mM MES-0.1mM EDTA-5% Glycerol-2 M NaCl, pH 5.5). IHF was then eluted by 20%-80% gradient of Buffer B over 20 column volumes. The peak was then concentrated and buffer exchanged in 15 mL concentrators (Amicon Ultracel 3K, Millipore) with 2.5% Buffer B. The heparin peak was then loaded onto a 20 mL cation exchange column (SP Sepharose Fast Flow™, GE Healthcare Life Sciences) in 2.5% Buffer B and eluted in a 2.5%-47.5% Buffer B gradient over 10 column volumes. TraM-containing fractions were concentrated in 15 mL concentrators (Amicon Ultracel 3K, Millipore) and exchanged into 0.5 M ammonium acetate. Protein concentration was determined using BCA protein assays (Pierce) following the manufacturer's instructions.

#### *Expression and purification of pED208 TraM*

TraM expression and expression was carried out as described in Chapter 4.

#### *Generation of pED208 sbmB, sbmC, sbmAB, sbmAC, and sbmABC oriT fragments for EMSA*

Forward and reverse oligos contained the *sbm* fragments highlighted in (Figure 4-11A) and one flanking base pair. Forward and reverse oligos were annealed at 0.2  $\mu$ M final concentration in 10 mM Tris pH 7.5 and 100 mM NaCl. Attempting to anneal at higher concentrations of DNA yielded a continuous smear upon acrylamide gel electrophoresis and not a single defined species.

Annealing was done in a PCR machine by heating a 100  $\mu$ L sample to 95°C for 15 minutes and cooled to 25°C at a rate of 0.01% change in temperature (approximately 2.5 hrs). Fragments were purified by electrophoresis on an 8% acrylamide gel in 1xTBE buffer, and extracted by crush-and-soak. Gel bands were cut out and crushed by centrifugation through a 0.6ml eppendorf tube with a hole pierced in the bottom into a 2ml eppendorf tube. The gel pieces were then incubated overnight at 4°C on a nutator in 0.9 mL of 10mM Tris pH 8.5 and 100  $\mu$ l of 3M sodium acetate. 2 volumes of 100% ethanol was added to each volume of supernatant and left on ice for 15min, followed by centrifugation at 4°C for 15 min at 16,100 x g. The 100% ethanol was removed and 500  $\mu$ l 70% ethanol was added to rinse the pellet, followed by centrifugation at 4°C for 2 min at 16,100 x g. The 70% ethanol was then removed and the pellet was dried at 37°C. The pellets were then suspended in T4 polynucleotide kinase (Invitrogen) for  $^{32}$ P labeling.

*sbmABC* fragments were generated by PCR from pED208 plasmid DNA. The forward primer had the sequence CTT GAA TTC CTC CTG GCT GAC, and the reverse had the sequence GGT AAC GAG ATC GGT GAT CTG.

*pED208 TraM binding to sbm fragments by electrophoretic mobility shift assay*

EMSA of TraM binding to *sbm* fragments was carried out as described in “ $K_d$  determination of pED208 TraM binding to *sbmA* by electrophoretic mobility shift assay” in Chapter 4

*Binding of IHF to ihf-sbmA by electrophoretic mobility shift assay*

IHF-*ihfC-sbmA* binding buffer was 50 mM Tris pH 7.5, 100 mM KCl, 10% glycerol, 30 ng/μL bovine serum albumin (Pierce), 20 ng/μL polydI·dC (Roche). All other steps were the same as described in Chapter 4 for TraM-*sbmA* EMSA.

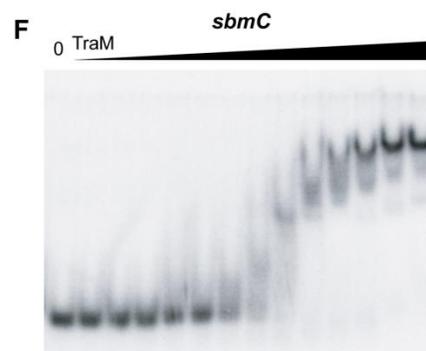
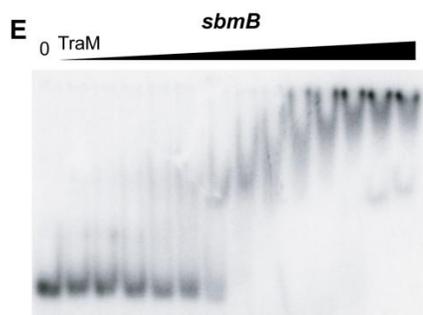
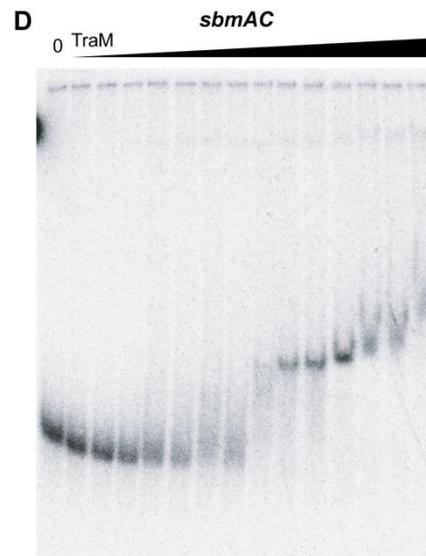
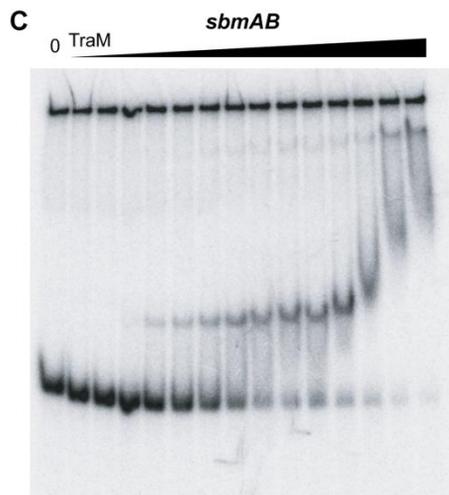
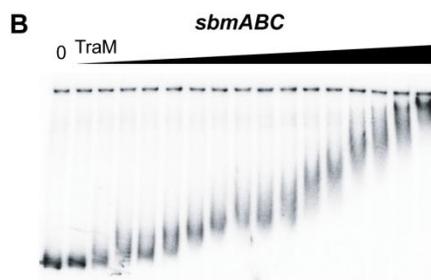
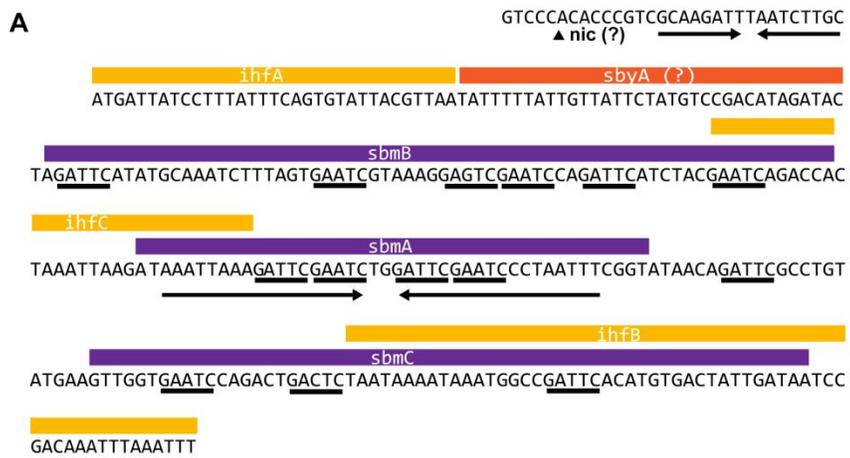


Figure A-1 EMSA of pED208 TraM binding to longer *oriT* fragments

a) pED208 *oriT* sequence. Inverted repeats are shown by arrows. GANTC TraM binding motifs are underlined.

b) Titration of TraM onto *sbmABC*.

c) Titration of TraM onto *sbmAB*

d) Titration of TraM onto *sbmAC*

e) Titration of TraM onto *sbmB*

f) Titration of TraM onto *sbmC*

For *sbmABC*, *sbmAB*, and *sbmAC*, TraM concentrations in each lane are 0, 0.1, 0.5, 2, 4, 7, 10, 20, 40, 70, 100, 200, 500, 1000, and 2000 nM.

For *sbmB* and *sbmC*, TraM concentrations in each lane are 0, 1, 2, 4, 7, 10, 20, 50, 100, 300, 1000, 3000, 10000, and 30000 nM.

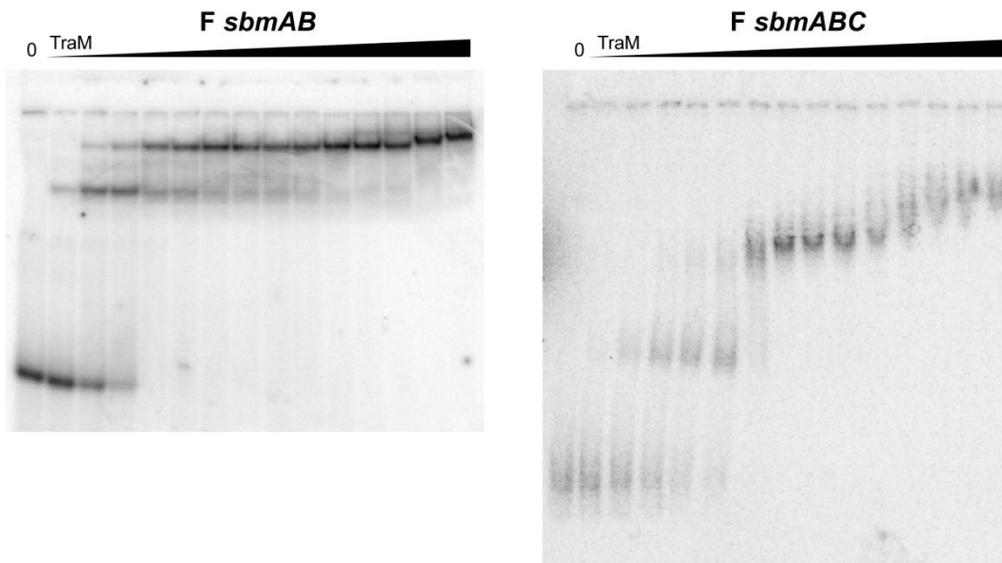


Figure A-2 EMSA of F TraM binding to longer *oriT* fragments

Left: Titration of TraM onto *sbmAB*. TraM concentrations in each lane are 0, 2, 5, 10, 20, 40, 70, 100, 150, 300, 500, 750, 1000, 2000, and 5000 nM.

Right: Titration of TraM onto *sbmABC*. TraM concentrations in each lane are 0, 1, 2, 4, 7, 10, 20, 40, 70, 100, 200, 500, 700, 1000, and 2000 nM.

**A**

	IHF	TraM
<i>ihfC-sbmA</i>	GAATCAGACCCTAAATTAAGATAAATTAAGATTCGAATCTGGATTCGAATCCTAATTCGGTA	
	CTTAGTCTGGTGATTAAATCTATTTAATTTCTAAGCAATGACCTAAGCTTAGGGATTAAAGCCAT	
<i>random-sbmA</i>	AGCTCGGTCACTCGTACGCTTCAGATGCCATGGATTCGAATCTGGATTCGAATCGCAGACTGTCTGA	
<i>-random</i>	TCGAGCCAGTAGCATGCGAAGTCTACGGTACCTAAGCTTACACCTAAGCTTAGCGTCTGACAGACT	

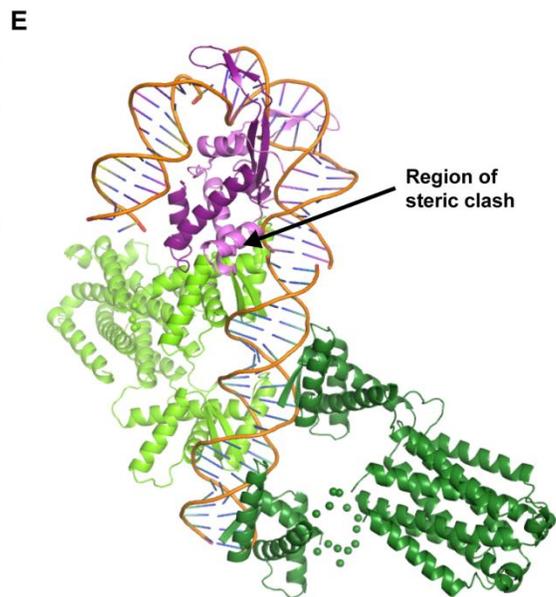
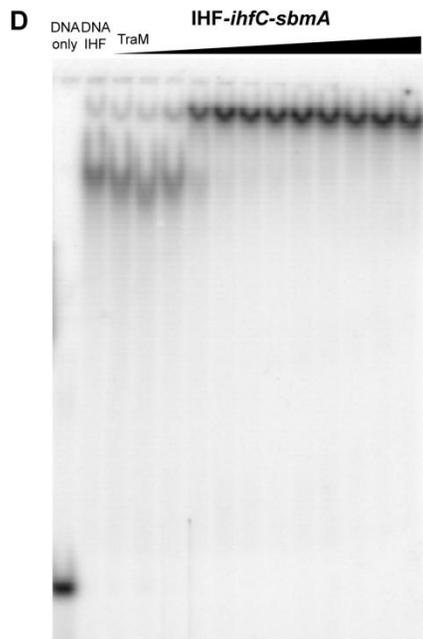
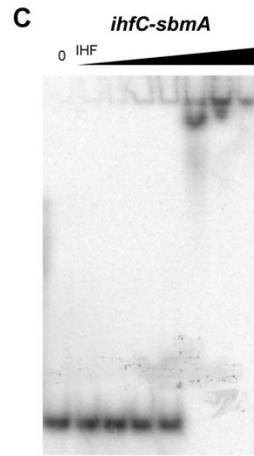
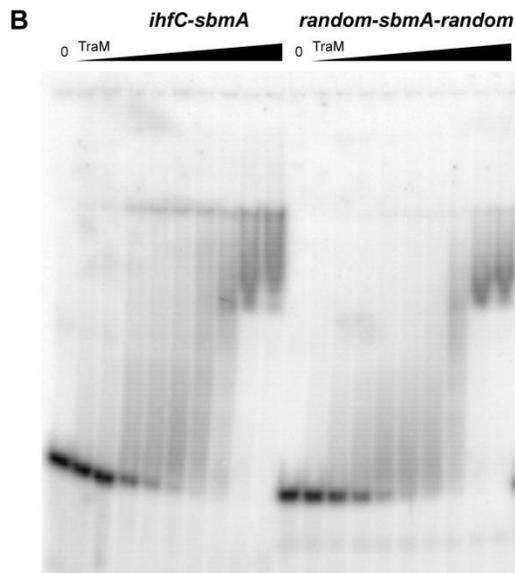


Figure A-3 EMSA of pED208 TraM and IHF binding to *ihfC-sbmA* DNA

- a) *ihfC-sbmA* fragment and randomized flanking sequence *sbmA* fragment used. The presumed IHF binding site, based on the number of DNA bases bound by IHF in the IHF-DNA crystal structure (Rice, *et al.*, 1996) and the center of the binding site shown by footprinting (Di Lorenzo, *et al.*, 1995). The TraM binding motifs are boxed in green.
- b) Titration of TraM onto *ihfC-sbmA* or *random-sbmA-random*. TraM concentrations in each lane are 0, 2, 5, 10, 20, 50, 100, 200, 500, 750, 1000 nM.
- c) Titration of IHF onto *ihfC-sbmA*. IHF concentrations in each lane are 0, 10, 50, 200, 1000, 5000, 15000, 50000 nM.
- d) Titration of TraM onto IHF-*ihfC-sbmA* complexes and electrophoresis. IHF concentration in each lane is 5 $\mu$ M. TraM concentrations in each lane are 0, 1, 5, 20, 100, 500, 1000 nM.
- e) Model of TraM-IHF-*ihfC-sbmA*, assuming *ihfC* is a physiologically relevant binding site and *ihfC* and *sbmA* are occupied simultaneously.

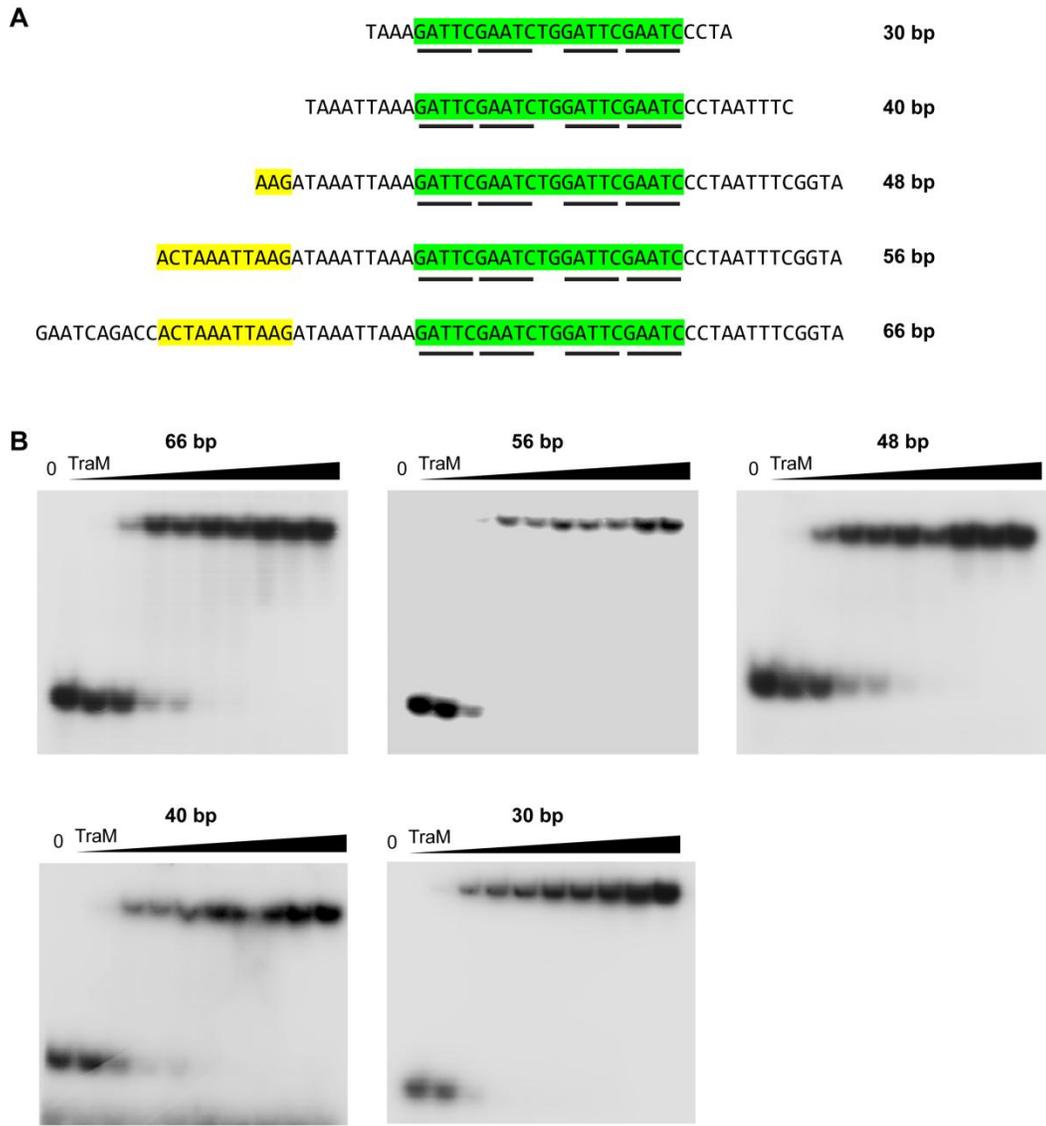


Figure A-4 Variation in binding affinity of pED208 TraM to *sbmA* with varying lengths of flanking DNA.

- sbmA* fragments used for EMSA. The presumed IHF binding site, based on the number of DNA bases bound by IHF in the IHF-DNA crystal structure (Rice, *et al.*, 1996) and the center of the binding site shown by footprinting (Di Lorenzo, *et al.*, 1995) is shown in yellow. The TraM binding motifs are boxed in green.
- Titration of TraM onto *sbmA* fragments of varying length up to that of *ihfC-sbmA*

## References

- Abo T & Ohtsubo E (1995) Characterization of the functional sites in the oriT region involved in DNA transfer promoted by sex factor plasmid R100. *J Bacteriol* **177**: 4350-4355.
- Di Laurenzio L, Frost LS, Finlay BB & Paranchych W (1991) Characterization of the oriT region of the IncFV plasmid pED208. *Mol Microbiol* **5**: 1779-1790.
- Di Laurenzio L, Scraba DG, Paranchych W & Frost LS (1995) Studies on the binding of integration host factor (IHF) and TraM to the origin of transfer of the IncFV plasmid pED208. *Mol Gen Genet* **247**: 726-734.
- Fekete RA & Frost LS (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.
- Rice PA, Yang S, Mizuuchi K & Nash HA (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**: 1295-1306.
- Schwab M, Gruber H & Hogenauer G (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol Microbiol* **5**: 439-446.
- Zinkel SS & Crothers DM (1987) DNA bend direction by phase sensitive detection. *Nature* **328**: 178-181.
- Zwieb C & Adhya S (2009) Plasmid vectors for the analysis of protein-induced DNA bending. *Methods Mol Biol* **543**: 547-562.

## Appendix B

### Co-crystallization of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] with F *sbmA*

#### Introduction

The crystal structure of pED208 TraM bound to *sbmA* answers many questions about TraM-*sbmA* interaction. It confirms sequence analysis, *in vivo* and *in vitro* evidence that the N-terminal DNA-recognition domain of TraM is a ribbon-helix-helix (RHH) domain. The RHH fold is a frequently-used protein motif in prokaryotic transcriptional regulators characterized by a short N-terminal  $\beta$ -sheet containing residues whose side chains form specific contacts with DNA bases in the major groove, followed by two  $\alpha$ -helices (Schreiter & Drennan, 2007). The residues predicted to be those of the  $\beta$ -sheet which contact DNA bases, shown to have an effect on conjugation *in vivo*, and shown to be important for DNA binding are the same as those which contact the DNA bases in the crystal structure. We can also identify non-specific interactions between TraM and *sbmA*, which includes interaction between the end of the second  $\alpha$ -helix and the DNA phosphate backbone, and the electrostatic repulsion between acidic TraM residues and the phosphate backbone. The sum of these interactions results in distortion of the DNA helical axis and unwinding of the DNA helix (Wong, *et al.*, 2011).

However, many aspects of F TraM-*sbmA* interactions are not understood. The overall mechanism for cooperative recognition of F *sbmA* by two TraM tetramers from the same plasmid is conserved (Wong, *et al.*, 2011), but many of the details about specific recognition of *sbmA* remain unknown. The TraM binding motif of pED208, GANTC, is perfectly palindromic, as is the  $\beta$ -sheet which binds to it. However, the sequence of the F plasmid binding motif,

A(G/C)CG(G/C)T, is much less strictly conserved, is longer by 1 base pair, and not as obviously palindromic. This makes prediction of how the F TraM RHH domain recognizes its binding motif difficult. How F TraM recognizes such a different binding motif with a conservative substitution, Asn 5 for Gln 5, in only 1 out of the 3  $\beta$ -sheet residues which contact DNA bases, is unknown.

It is worth noting that a significant number of RHH domains recognize non-palindromic DNA sequences using the perfectly palindromic  $\beta$ -sheet that results from homodimerization, including MetJ (Rafferty, *et al.*, 1989), Arc (Toro-Roman, *et al.*, 2005), and  $\omega$  (Weihofen, *et al.*, 2006). It is also notable that a number of RHH domains recognize perfectly or almost-perfectly palindromic sequences with an asymmetric contributions from each chain of the RHH domain, e.g. ParR (Schumacher, *et al.*, 2007), NikR (Schreiter, *et al.*, 2006). and CopG (Gomis-Ruth, *et al.*, 1998). These experimental observations add an additional layer of unpredictability in assigning TraM-*sbmA* contacts for the F plasmid.

The additional length of *sbmA*, required by the F TraM *sbmA* is another aspect of F TraM-*sbmA* that is not well-understood. The minimal length of F *sbmA* for maximal binding affinity of F TraM is 30 bp, whereas the minimal length of pED208 *sbmA* for pED208 TraM is 24 bp (Chapter 5). This results in 3 flanking base pairs in F *sbmA* outside the F binding motifs, and 1 flanking base pair outside those of pED208. The precise nature of the contributions of the additional base pairs to F TraM binding is unknown. However, based on the additional length of the  $\alpha$ 1- $\alpha$ 2 linker in F, the presence of 2 additional basic residues compared to pED208 (Chapter 5), and the protection from enzymatic cleavage at

the  $\alpha$ 1- $\alpha$ 2 upon DNA binding (Lu, J., unpublished data), it was hypothesized that the extra base pairs form contacts with the TraM  $\alpha$ 1- $\alpha$ 2 loop that are important for full binding affinity.

In light of these difficulties, obtaining the crystal structure of the F-plasmid TraM in complex with its cognate *sbmA* sequence was highly desirable for elucidating the nature of F RHH domain interaction with the F binding motif. Large, superficially good crystals of the F TraM-*sbmA* complex were obtained. However, despite extensive screening of crystallization conditions and DNA constructs, no crystals were obtained that yielded diffraction higher than 4.5 Å (Lu, J., unpublished data). Therefore, a new strategy for obtaining a crystal structure that would show F TraM RHH contacts with *sbmA* was needed.

Co-crystallization of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] with F *sbmA* was attempted as another means to determine how the RHH domain of F interacts with its cognate DNA. This strategy was chosen because the pED208 C-terminal domain forms the majority of the crystal contacts in the TraM-*sbmA* complex crystals (Figure B-1A). Some contacts arise from the N-terminal domain packing against the C-terminal domain, but no DNA-DNA or DNA-protein crystal contacts exist. It was hoped for that the presence of the pED208 C-terminal domain in the chimeric TraM would be sufficient to form crystal contacts similar to the wild-type pED208 TraM-*sbmA* structure, and the TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] –F *sbmA* complex could be crystallized in the same conditions as wild-type pED208 TraM-*sbmA*. If successful, we predicted that it would provide the information desired about F TraM-*sbmA* contacts as well as avoid the need for excessive crystallization screening.

## Results

### *Crystallization screening of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>]-F sbmA under pED208 TraM-sbmA crystallization conditions*

Grid screens were set up around the pED208 TraM-sbmA crystallization conditions (36% MPD, 4.5% PEG2000, 100mM cacodylic acid pH 6.5, at room temperature, 22°C). The range of conditions tested included were from 33% MPD/4.13% PEG2000 – 38% MPD/4.75% PEG2000, and pH 6.0-6.7 100mM cacodylic acid. No crystals were obtained. Additive screening also did not yield any crystals. The range of precipitant concentrations chosen was thought to be sufficient for reaching crystal nucleation, if any crystals would form under these conditions, since drops remained clear at 34% MPD/4.25% PEG2000 and lower, and formed heavier precipitate at 35% MPD/4.37% PEG2000 and higher without any crystals forming.

### *Crystallization of screening of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>]A18D+K22I - F sbmA under pED208 TraM-sbmA crystallization conditions*

A double mutant of the chimeric TraM protein, TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>]A18D+K22I, was cloned, expressed and purified. The rationale for generating this mutant was to make the crystal contacts between the pED208 C-terminal domain and the F N-terminal domain more like those found in pED208, so that it would have a better chance of crystallizing in the same space group and conditions as the wild-type pED208 TraM-sbmA complex. Upon inspection of the C-terminal domain/N-terminal domain interface (Figure B-1B), the amino acids forming crystal contacts are conserved in the proteins with the exception of those at positions 18 and 22. Therefore, these residues are mutated to those of pED208 to create TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>]A18D+K22I. A similar grid screen around

the wild-type pED208 TraM-*sbmA* crystallization conditions was set up, but no crystals were obtained.

*Crystallization screening of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>]-F sbmA using commercial sparse matrix screening kits*

The first round of crystallization screening was carried out using manual, 24-well crystallization tray (Qiagen) setups with 500 uL reservoir solution and 1+1  $\mu$ L drops. Kits screened were Wizard I and II (Molecular Dimensions) kits, and ProComplex (Qiagen). One possible crystallization hit was found, condition #4 of the Wizard I kit -100 mM imidazole pH 8.0, 30% PEG 8000, 0.2M NaCl (Figure B-2A). This condition was optimized to 100 mM imidazole pH 8.0, 25% PEG4000, 0.2 M NaCl, which yielded 0.1  $\mu$ m irregularly-shaped crystals (Figure B-2B). Other conditions to try and improve the crystals were varying NaCl (0.15-0.25 M), varying salt (0.2 M KCl, sodium acetate, calcium acetate), additives (1 mM spermine, 5% glycerol), and DNA construct (Table B-2). Free interface diffusion in capillaries was attempted, but no crystals resulted. No conditions were found that would yield at least superficially single, regular crystals. The largest irregular crystals were soaked in mother liquor+various cryoprotectants (10-20% glycerol, sucrose, PEG400, MPD, ethylene glycol) and screened for diffraction at the home x-ray source. No diffraction was obtained from any of the crystals.

The second round of crystallization screening used robotic (Art Robbins Gryphon) 96-well tray setups (Art Robbins Intelliplate) with 70  $\mu$ L reservoir volume and 0.2+0.2  $\mu$ L drops, and screens in the 96 deep well block format (Qiagen). Screens used were Classics I and II, JCSG+, Nucleix, MBCClass, PEGS (Qiagen). Two hits were obtained, Condition #48 of the Classics II screen (100 mM Bis-Tris

pH 5.5, 45% MPD, 0.2 M  $\text{CaCl}_2$ ) (Figure B-2C), and Condition # 52 of the MBClass screen (100 mM Tris pH 8.0, 20% PEG2000, 2% MPD, 0.3 M  $\text{Mg}(\text{NO}_3)_2$ ) (Figure B-2E). The Hit #2 was not reproducible with a relatively narrow grid screen around those conditions. Therefore, it was initially set aside for optimization of the third hit, which was easily reproducible.

Well conditions varied to optimize Hit #3 were PEG % (12% - 20%), PEG length (1000 and 3350), pH (6.5-9.0), PEG/MPD ratio (14/0, 14/2.8, 14/4.2),  $\text{Mg}(\text{NO}_3)_2$  concentration (0.2-1.0 M), salt type of salt ( $\text{NH}_4\text{NO}_3$ ,  $\text{MgCl}_2$ ,  $\text{MgSO}_4$ , MgAcetate,  $\text{CaCl}_2$ ), protein:DNA ratio (3:1, 2:1, 1:2), and replacement of MPD with glycerol, PEG400, or PEG 4000. Protein concentration (8, 25 mg/mL) and temperature (4°C and 16°C) were. The best condition was 100 mM Tris pH 8.0, 15% PEG2000, 1.5% MPD, 0.7 M  $\text{Mg}(\text{NO}_3)_2$ , but the best crystals were irregularly-shaped and very small (<0.5  $\mu\text{m}$ ) (Figure B-2F). These crystals were soaked in various cryoprotectants and screened for diffraction at a synchrotron x-ray source, and no diffraction was observed for any of them.

More screening was done to see if the crystals from Hit #2 could be reproduced. A wider range of MPD concentration was used, 20-45%, and the boundary between clear drops and precipitate formation was between 30-35% MPD. This range was used as a starting point for additive screening. 2.5-5% glycerol, 3.5% PEG400, 3.5% PEG4000 were tested as additives, and crystals were obtained with 3.5% PEG400 as an additive.

Extensive screening was carried out to improve the crystals. The MPD concentration varied from 26-32%, and different lengths of PEG were used as

additives (200, 1000, 4000). pH was varied from 4.5-7.0, CaCl<sub>2</sub> was varied from 0.1-0.5 M, and different salts were used instead of MgCl<sub>2</sub>, NaCl, calcium acetate. The MPD/PEG ratio was varied from 25/8 – 30/1, which resulted in marginally better crystals at 28/5. Higher and lower protein concentrations (8 and 25 mg/ml) were used, and larger but uglier crystals resulted from 25 mg/mL. A wide range of temperatures were tested (4, 16, 20, 25, 27, 30°C), and growth at 25-27°C resulted in marginally better crystals. The best crystals at this point grew at 28% MPD, 5% PEG400, 100 mM Bis-Tris pH 5.5, 0.2 M CaCl<sub>2</sub>, and 27°C, with a drop size and ratio of 0.5+0.5 µL (Figure B-2D). Drop sizes and ratios were varied, but none gave significantly better crystals (1+1, 2+2, 1.5+1, 3+2). Buffer type was also varied (ADA, Bis-Tris propane, L-histidine, and succinic acid). The A18D K22I mutant was also used in crystallization trials and crystals did form, but were no better than the original chimeric protein. A second additive, 2% propanol, was found to give more consistently reproducible crystals with an adjustment of precipitant concentration to 24 % MPD and 4.3% PEG400.

Other methods of growing crystals were also attempted. Crystals were obtainable under microbatch growth conditions (sitting drops were covered with 20µl of oil, at 23% MPD and 4.1% PEG400), but were no better, even after attempting different ratios of paraffin to silicone oil. Growth using the capillary counterdiffusion method (Ng, *et al.*, 2003) where the protein solution was mixed with 0.05% low-melting point agarose was attempted, but no crystals resulted.

## **Discussion**

No crystals resulted after exhaustive screening of crystallization conditions. This prompted additional consideration of an aspect of the TraM [F<sup>1-</sup>

<sup>55</sup>:pED<sup>56-127</sup>] design which became more apparent after routine sequencing of the TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] A18D K22I mutant during the cloning process. Two residues of the linker joining the N- and C-terminal domains were missing (Figure B-3A). It may not be immediately apparent from the name of the construct, unless one notes that the numbering of the linker residues in F and pED208 are off by 2 residues due to the additional linker length of the  $\alpha$ 1- $\alpha$ 2 loop. The F residues are numbered +2 relative to pED208 residues at the homologous location, therefore resumption of pED208 at the residue numbered +1 to the F numbering would result in 2 missing amino acids. The effect of these 2 missing amino acids, and whether these effects would influence crystallization of the Hybrid TraM-F *sbmA* complex is unknown.

The significance of these 2 missing amino acids is not immediately apparent, as TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] is able to bind F *sbmA* at the same affinity as wild-type (Chapter 5). The effects on protein conformation may be more subtle, but still enough to affect the ability of the complex to crystallize in the same space group using similar crystal contacts to that of wild-type pED208 TraM. On one level, defining the linker length in TraM appears to be arbitrary, since different linker lengths occur for each monomer of the tetramer due to the ability of the  $\alpha$ 2 helix to unwind to different degrees (Figure B-3B). Nevertheless, evidence for the linker length being critical for TraM function lies in its conservation among the F-like plasmid family (Figure B-3C). The length is perfectly conserved between pED208 and F TraM, and is exactly the same for the more well-known members of the F plasmid family, like R1 and R100. The linker length may affect the level of unwinding of *sbmA* DNA, which would in turn affect the angle of offset between the two RHH domains relative to each other

(Figure 6-3D). Any difference relative to the wild-type pED208 complex would affect whether TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] can form crystal contacts in exactly the same way.

Therefore, a future direction for continuation of this project would be to make a derivative of the chimeric TraM with wild-type linker length. I propose to generate this mutant is by site-directed mutagenesis of the pre-existing TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] and TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] A18D K22I clones using the primers in Figure 6-3E pJWM612F, pJWM612R and the overlap extension method of (Ho, *et al.*, 1989). The expressed and purified protein would be tested for TraM binding by EMSA, and subject to crystallization screens under pED208 TraM-*sbmA* conditions and if necessary, with commercial screening kits.

## Materials and Methods

### *Growth media and bacterial strains*

Media and antibiotics for bacterial growth and bacterial strains used are described in (Wong, *et al.*, 2011).

### *Primers, Plasmids, and cloning of TraM and TraD mutants*

Construction of plasmid pJLM407 used for expression of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] is described in (Wong, *et al.*, 2011) and Table B-1. Construction of plasmid pJLM407-A18DK22I used for expression of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] A18D K22I is described as follows. PCR primer pair JWM612F and JWFM-A18DK22I-R were used to amplify a fragment containing the N-terminal portion of F *traM* from pJLM407. Primer pair and JWFM-A18DK22I-F and JWM612R were used to amplify a fragment containing the C-terminal portion of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] from

pJLM407. The PCR products of the above 2 reactions and the primers JWM612F and JWM612R were used to amplify the full-length TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] A18D K22I fragment, which was further digested by *EcoRI* and *BamHI* and cloned into the *EcoRI*-*BamHI* sites of pJLM407.

*Expression and Purification of TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] and TraM [F<sup>1-55</sup>:pED<sup>56-127</sup>] A18D K22I*

Expression and purification of hybrid TraM is the same as that for pED208 TraM described in (Wong, *et al.*, 2011) and Chapter 4.

**Table B-1 Oligonucleotides for *oriT* and *sbmA* length variants**

	Description & references
pACM407	pACYC184 with a F TraM <sup>1-55</sup> :pED208 TraM <sup>56-127</sup> ; (Wong, <i>et al.</i> , 2011)
pACM407-A18DK22I	pACM407 with F TraM <sup>1-55</sup> :pED208 TraM <sup>56-127</sup> mutant A18D K22I
JWM612F	GAT AAG CTT GAT ATC GAA TTC
JWM612R	GTG GAT CCA CCA GAA CAT TC
JWFM-A18DK22I-F	GAA AAA ATA AAT GAT ATT ATT GAG ATT CG
JWFM-A18DK22I-R	CGA ATC TCA ATA ATA TCA TTT ATT TTT TC

**Table B-2 Variant *sbmA* oligos for crystallization optimization**

	Description
30F/31AR	GAT ACC GCT AGG GGC GCT GCT AGC GGT GCG CTA TGG CGA TCC CCG CGA CGA TCG CCA CGC A
31AF/30R	G GAT ACC GCT AGG GGC GCT GCT AGC GGT GCG CTA TGG CGA TCC CCG CGA CGA TCG CCA CGC
31AF/31AR	G GAT ACC GCT AGG GGC GCT GCT AGC GGT GCG CTA TGG CGA TCC CCG CGA CGA TCG CCA CGC A

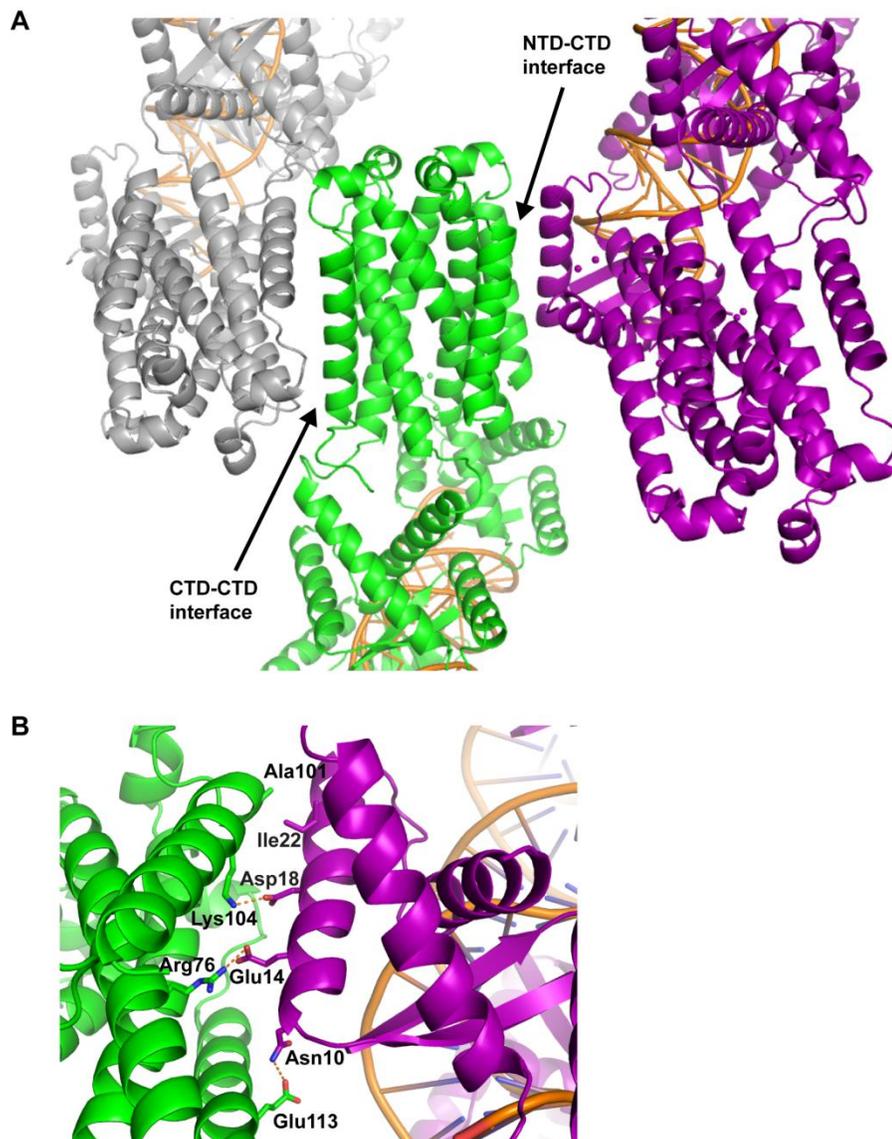


Figure B-1 TraM [ $F^{1-55}$ :pED $^{56-127}$ ] crystallization rationale

a) pED208 TraM-*sbmA* crystal contacts

b) Close-up view of the N-terminal domain-C-terminal domain crystal contacts in the pED208 TraM-*sbmA* complex

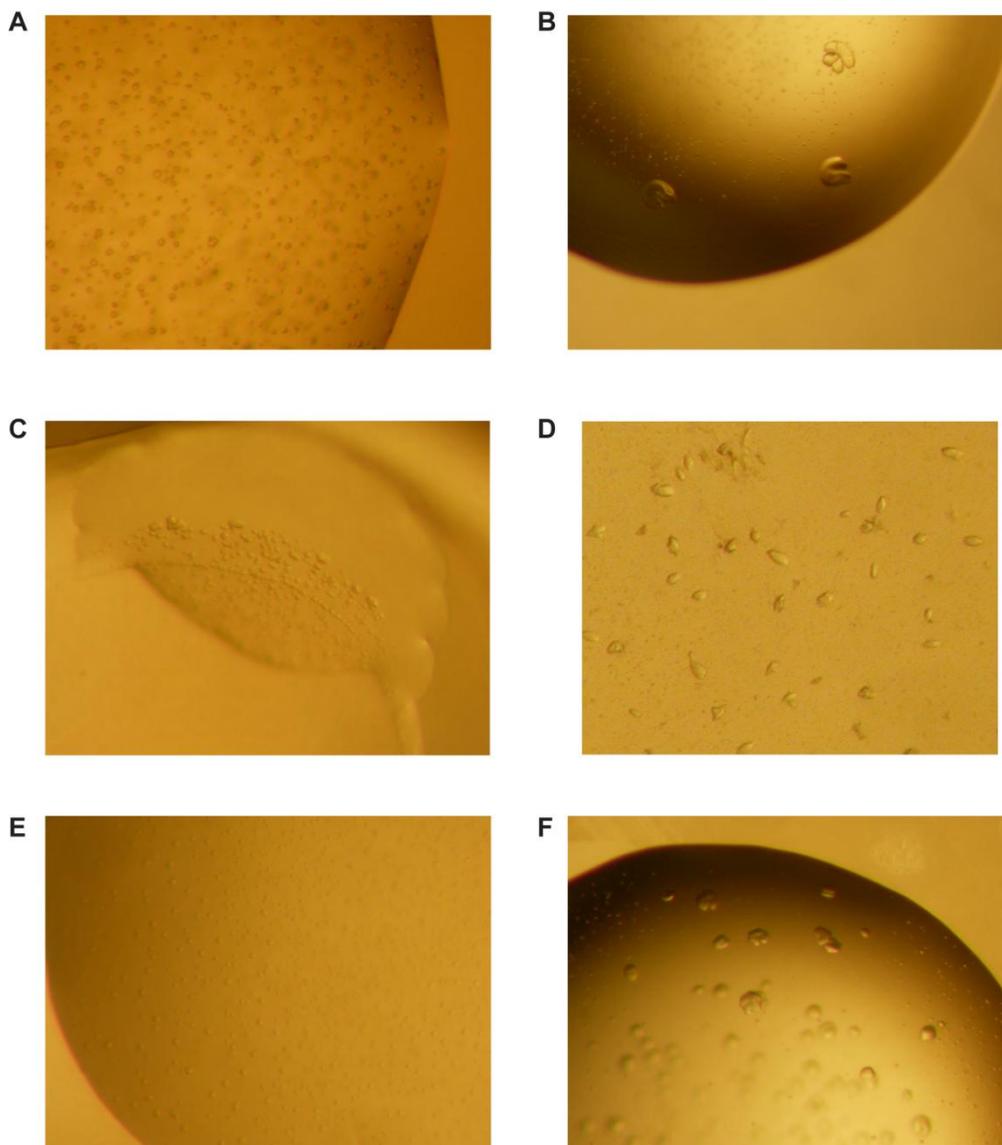


Figure B-2 Crystallization hits of TraM [F<sup>1-55</sup>;pED<sup>56-127</sup>] – F *sbmA* complexes  
a) Hit #1 -initial drop from Wizard I #4  
b) Optimized crystals of Hit #1  
c) Hit #2 -initial drop from Classics II #48  
d) Optimized crystals of Hit #2  
e) Hit #3 -initial drop from MBClass #53  
f) Optimized crystals of Hit #3

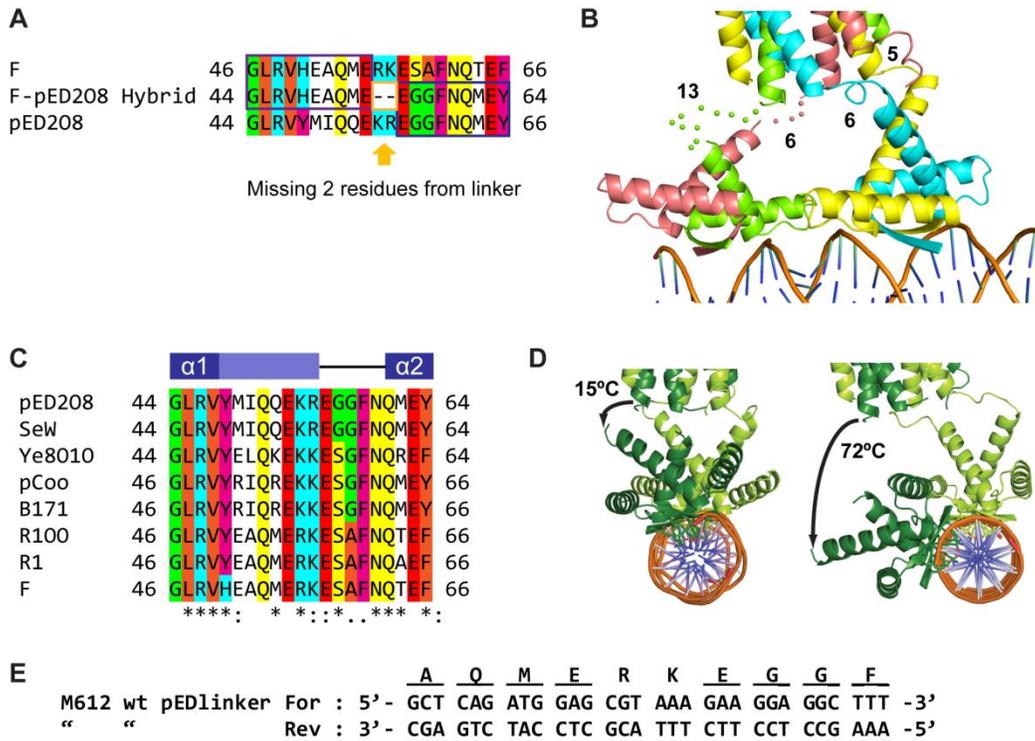


Figure B-3  $\alpha 1$ - $\alpha 2$  loop in F-like TraM

- Sequencing result of TraM [F<sup>1-55</sup>;pED<sup>56-127</sup>] A18D K22I
- Linker lengths of the chains in the pED208 TraM tetramer when bound to *sbmA*
- Sequence alignment of F-like plasmid TraM
- Angle of offset between the two RHH domains with DNA unwinding as shown in the crystal structure (left) or with no DNA unwinding in ideal B-DNA.
- Forward and Reverse site-directed mutagenesis primers for introducing the 2 missing residues into the  $\alpha 1$ - $\alpha 2$  loop in TraM [F<sup>1-55</sup>;pED<sup>56-127</sup>]

## References

- Gomis-Ruth FX, Sola M, Acebo P, *et al.* (1998) The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO J* **17**: 7404-7415.
- Ho SN, Hunt HD, Horton RM, Pullen JK & Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**: 51-59.
- Ng JD, Gavira JA & Garcia-Ruiz JM (2003) Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. *J Struct Biol* **142**: 218-231.
- Rafferty JB, Somers WS, Saint-Girons I & Phillips SE (1989) Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. *Nature* **341**: 705-710.
- Schreiter ER & Drennan CL (2007) Ribbon-helix-helix transcription factors: variations on a theme. *Nat Rev Microbiol* **5**: 710-720.
- Schreiter ER, Wang SC, Zamble DB & Drennan CL (2006) NikR-operator complex structure and the mechanism of repressor activation by metal ions. *Proc Natl Acad Sci U S A* **103**: 13676-13681.
- Schumacher MA, Glover TC, Brzoska AJ, Jensen SO, Dunham TD, Skurray RA & Firth N (2007) Segrosome structure revealed by a complex of ParR with centromere DNA. *Nature* **450**: 1268-1271.
- Toro-Roman A, Mack TR & Stock AM (2005) Structural analysis and solution studies of the activated regulatory domain of the response regulator ArcA: a symmetric dimer mediated by the alpha4-beta5-alpha5 face. *J Mol Biol* **349**: 11-26.
- Weihofen WA, Cicek A, Pratto F, Alonso JC & Saenger W (2006) Structures of omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Res* **34**: 1450-1458.
- Wong JJ, Lu J, Edwards RA, Frost LS & Glover JNM (2011) Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**: 6775-6788.