

The Structure of the Plasmid pCU1 TraI Relaxase and the Role of the pCU1 TraI Relaxase-Helicase during Conjugative Plasmid Transfer

Rebekah Potts Nash

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry.

Chapel Hill
2011

Approved by

Advisor: Matthew Redinbo, Ph.D.

Reader: Kevin Slep, Ph.D.

Reader: Linda Spremulli, Ph.D.

Steven Matson, Ph.D.

Gary Pielak, Ph.D.

Abstract

Rebekah Potts Nash: The Structure of the Plasmid pCU1 TraI Relaxase and the Role of the pCU1 TraI Relaxase-Helicase during Conjugative Plasmid Transfer
(Under the direction of Matthew R. Redinbo)

Bacteria disseminate genetic material to neighboring cells using conjugative plasmid transfer (CPT). During CPT, a donor bacterium transfers one strand of a double-stranded DNA plasmid to a recipient. Each conjugative plasmid encodes a complex of proteins necessary for its transfer. One of these proteins, the relaxase, initiates plasmid transfer by severing the *nic* site of the transferred strand (T-strand). A DNA helicase then separates the T-strand from the parent strand, starting at the *nic* site. The relaxase acts a second time to terminate transfer by resealing the nicked T-strand. The resistance plasmid pCU1 encodes a multi-domain protein, TraI, that supplies both the relaxase and helicase activities required for its transfer. We analyzed the structure and function of pCU1 TraI in order to compare it to similar plasmid-encoded proteins and to identify TraI-mediated activities that could be targeted by inhibitors.

Characterization of the pCU1 relaxase revealed unique structural and functional modifications that this enzyme has introduced into the traditional relaxase-mediated DNA nicking mechanism. First, while the overall fold of the pCU1 relaxase is similar to that of homologous relaxases, its conserved DNA nicking residues (Y18,19,26,27) are flipped up to 14 Å out of the relaxase active site. Second, the pCU1 relaxase preferentially utilizes Y26 or a combination of Y18+Y19 when nicking DNA. In contrast, homologous relaxases use the first tyrosine in amino acid sequence for DNA nicking. Third, the pCU1 relaxase lacks the sequence-specific DNA binding characteristic of homologous relaxase enzymes. However, it maintains highly sequence specific and metal-dependent DNA nicking.

Analysis of the pCU1 helicase established the extent of the minimal helicase domain, the location of the seven conserved helicase motifs, and the substrate requirements of the helicase ATPase activity. The pCU1 helicase harnesses the energy released during ATP hydrolysis to drive DNA strand separation. After optimizing an ATPase assay for use with pCU1 TraI, small molecule libraries were screened for their ability to inhibit pCU1 TraI, and several potential TraI ATPase inhibitors were identified.

In summary, this structural and functional characterization of pCU1 TraI identified features unique to this enzyme and revealed particular activities could be targeted by TraI-specific inhibitors.

Dedication

To my parents, Karl and Brenda Potts, and to my husband, Dave, your patience and unwavering support not only made this work possible, but also a joy to complete.

Acknowledgements

I cannot hope to adequately thank those who have wisely guided me through graduate school and who have made my time in graduate school such a pleasure. I will simply hope to convey here a small measure of my gratitude, and, then, to trust that those who read these words will be able to surmise the true magnitude of my appreciation.

I thank the many collaborators who have welcomed me as a passing member into their labs, including the Matson lab, the Macromolecular Interactions Facility, and the Center for Integrative Chemical Biology and Drug Discovery. I especially appreciate the guidance of Steve Matson and Kambiz Tahmaseb, Ashutosh Tripathy, Bill Janzen, Emily Hull-Ryde and Cathy Simpson.

I thank the MD/PhD program and in particular Dr. O, David Siderovski, Carol, Alison, and Liz for their support and wisdom as I transitioned to graduate school, as I ventured into the world of grant writing, and as I now prepare to return to medical school.

I thank the hard working, gracious, and thoughtful members of the Chemistry Department who have made my time as a chemistry graduate student such a pleasure, and in particular want to thank Karen Gilliam, Laura Condie, and Lars Sahl for the help they provided me.

I thank all the members of the Redinbo lab for their friendship and support. It has been a pleasure to work alongside each of you. I especially want to thank Monica for her friendship, great scientific questions, and editing prowess; Yuan for his friendship, amazing scientific insight, persistent ambivalence, and guidance with an over-abundance of binding assays; Joe for his friendship, his infectious love of science, and his guidance as I learned the techniques of CD and protein purification; Mike for his friendship, his mastery of music that he so graciously shared with me and Dave, and being an ESFP; I would like to thank my fellow Furmanite Bret for his friendship

and ever-patient help as I waded through crystallography; Laurie for her friendship and guidance as I collected and then refined crystallographic data; Bill for his friendship and energetic support of our lab's work; and Keith for his friendship, plasmid transfer prowess, and gracious help with all things science.

I thank all the former members of the Redinbo lab for the wisdom and knowledge they imparted to me during their time at UNC. I particularly want to thank Eric for his friendship, for taking out all the lion and bear traps in Carolina North, and for answering endless questions about computer scripts; Scott for his friendship, love of science and his Excel prowess; Laura for being a mentor extraordinaire during my rotation in the lab, for lab bike rides, and for introducing me to the joys of the synchrotron; Jill for her help with numerous and various crystallography programs and problems; Andy for his friendship, spinning ability, and amazing bike rides, and Mike Miley for his help both during his time here in the Redinbo lab and at the UNC crystallography core.

I thank the undergraduates and exchange students I have had the joy of working with. Franklin, Drew, Dan, and Johanne: I am so grateful for your hard work, but even more for your friendship. I wish you only the best.

I thank Matt for the opportunity to work in his laboratory. I am incredibly grateful for the support, advice, and encouragement he gave me as I worked and trained over these last five years. I greatly respect his love of science, of research, and his ability to share these traits with those who work for him.

I thank the loving, supporting, and ever-wise members of Christ Community Church whom I have the delight and honor of considering my friends, in particular Chris and Erin, Lauren, Nikki, Robin, Jessie, Rik and Bobbie, Kent and Jeline, Mike and Celette.

Finally, I thank my family for their patience and encouragement, and unconditional love and support. I thank my Mom and Dad, Karl and Brenda Potts, and my brother Stephen, and lastly, but most importantly, I thank my husband Dave. You guys are my heroes.

Table of Contents

List of Tables.....	x
List of Figures.....	xi
List of Abbreviations and Symbols.....	xiii
Chapter 1: Introduction to Conjugative Plasmid Transfer.....	1
1.1 An Overview of the Replication and Transfer of Genetic Material in Bacteria.....	1
1.2 The Classification of Transmissible Plasmids.....	2
1.3 The Mechanism of Conjugative Plasmid Transfer.....	4
1.4 Figures.....	6
Chapter 2: The Structure of the Plasmid pCU1 TraI Relaxase.....	11
2.1 Introduction to the Structure of the Relaxase.....	11
2.2 Construct Design and Cloning of the pCU1 TraI Relaxase.....	12
2.3 Expression and Purification of the pCU1 TraI Relaxase.....	13
2.4 Crystallization, Data Collection, and Data Processing of the pCU1 TraI Relaxase.....	15
2.5 Analysis of the pCU1 TraI Relaxase Structure.....	16
2.5.1 The Dimer in the Asymmetric Unit.....	16
2.5.2 The Overall Fold and the Active Site HUH Motif of the pCU1 TraI Relaxase.....	17
2.5.3 The N-terminus of the pCU1 TraI Relaxase.....	18
2.5.4 The Motif 3 Tyrosines of the pCU1 TraI Relaxase.....	19
2.6 Conclusions.....	20
2.7 Tables and Figures.....	20
Chapter 3: DNA Binding and DNA Nicking Activities of the Plasmid pCU1 TraI Relaxase.....	29

3.1 Introduction to Relaxase Function.....	29
3.2 Construct Design, Mutagenesis, Expression, and Purification of the pCU1 TraI Relaxase.....	30
3.3 Characterization of pCU1 TraI Relaxase Stability by Circular Dichroism (CD) Spectroscopy.....	31
3.4 Characterization of DNA Binding by the pCU1 TraI Relaxase.....	33
3.4.1 Introduction to DNA Binding by the pCU1 TraI Relaxase.....	33
3.4.2 Methodology.....	33
3.4.3 pCU1 TraI Relaxase DNA Binding Affinity and Specificity.....	34
3.5 Characterization of DNA Nicking by the pCU1 TraI Relaxase.....	36
3.5.1 Introduction to DNA Nicking by the pCU1 TraI Relaxase.....	36
3.5.2 Methodology.....	38
3.5.3 Metal-Dependent DNA Nicking by the pCU1 TraI Relaxase.....	40
3.5.4 Tyrosine-Dependent DNA Nicking by the pCU1 TraI Relaxase.....	42
3.5.5 Aspartic Acid-Dependent DNA Nicking by the pCU1 TraI Relaxase.....	44
3.5.6 Model of DNA Nicking by the pCU1 TraI Relaxase Active Site Tyrosines and Adjacent Aspartic Acid.....	46
3.5.7 DNA Structure- and Sequence-Dependent Nicking and Ligation by the pCU1 TraI Relaxase.....	48
3.6 Summary of Data.....	51
3.7 Tables and Figures.....	52
Chapter 4: DNA Binding and DNA-Dependent ATPase Activities of the pCU1 TraI Helicase.....	70
4.1 The Classification of Helicases.....	70
4.2 Introduction to Helicase Structure and Function.....	71
4.3 Analysis of the Sequence and Predicted Structure of the pCU1 TraI Helicase.....	72
4.4 Characterization of DNA Binding by the pCU1 TraI Helicase.....	74
4.4.1 Introduction to DNA Binding by the pCU1 TraI Helicase.....	74
4.4.2 Construct Design and Cloning of the pCU1 TraI Helicase.....	74

4.4.3 Expression and Purification of the pCU1 TraI Helicase.....	75
4.4.4 Effect of DNA on the Oligomeric State of pCU1 TraI.....	76
4.4.5 DNA Binding Stoichiometry and Affinity of pCU1 TraI and the pCU1 TraI Helicase.....	78
4.4.6 Summary of DNA Binding by pCU1 TraI and the pCU1 TraI Helicase.....	81
4.5 Characterization of pCU1 TraI Helicase ATPase Activity.....	83
4.5.1 Introduction to pCU1 TraI Helicase ATPase Activity.....	83
4.5.2 Design of an NADH-coupled ATPase Activity Assay.....	84
4.5.3 Optimization and Application of an NADH-coupled ATPase Activity Assay.....	85
4.6 Design, Optimization, and Validation of a High Throughput Assay of pCU1 TraI Helicase ATPase Activity.....	87
4.6.1 ADP Quest Assay Design.....	87
4.6.2 Results of the ADP Quest Assay Optimization and Validation.....	88
4.7 A High Throughput Screen for Inhibitors of the pCU1 TraI Helicase ATPase Activity.....	91
4.8 Future Steps for the Identification and Characterization of pCU1 TraI ATPase Inhibitors.....	92
4.9 Tables and Figures.....	93
Chapter 5: Characterization of Plasmid pCU1 Transfer.....	119
5.1 Introduction to Plasmid Transfer.....	119
5.2 Method Development for Characterization of Plasmid pCU1 Transfer.....	119
5.3 Analysis of Plasmid pCU1 Transfer.....	122
5.4 Tables and Figures.....	124
References.....	129

List of Tables

Chapter 1

No Tables

Chapter 2

Table 2.1 Crystallographic Statistics.....21

Chapter 3

Table 3.1 DNA Binding Assay Substrates.....53

Table 3.2 Thermal Stability of the pCU1 TraI Relaxase.....54

Table 3.3 DNA Nicking and DNA Cross-over Assay Substrates.....55

Table 3.4 Summary of Metal Ion Data.....57

Table 3.5 Contribution of Individual Tyrosine and Aspartic Acid Residues to DNA Nicking.....58

Chapter 4

Table 4.1 Conserved SF1/SF2 Helicase Motifs.....94

Table 4.2 pCU1 TraI Oligomerization and DNA Binding.....95

Table 4.3 DNA Substrates for DNA Binding and ATPase Activity Assays.....96

Table 4.4 Impact of NADH-coupled ATPase Activity Assay Conditions on
pCU1 TraI ATPase Activity.....98

Chapter 5

Table 5.1 Minimal Inhibitory Concentration of Antibiotics for Selected Cell Lines.....125

Table 5.2 Number of Transconjugates Generated during Plasmid pCU1 Transfer Assay125

List of Figures

Chapter 1

Figure 1.1 The Classification of Transmissible Plasmids.....	7
Figure 1.2 The Mechanism of Conjugative Plasmid Transfer.....	8
Figure 1.3 DNA Nicking Mechanism of the Relaxase Motif 3 Tyrosine.....	10

Chapter 2

Figure 2.1 The Structure of the Conjugative Relaxase in Complex with Plasmid <i>oriT</i> DNA.....	22
Figure 2.2 Structure of the pCU1 TraI Relaxase.....	23
Figure 2.3 Structural Comparison of Relaxase Enzymes.....	24
Figure 2.4 Amino Acid Sequence Alignment and Secondary Structure of the pCU1 TraI Relaxase.....	25
Figure 2.5 SDS-PAGE of Purified pCU1 TraI Relaxase.....	26
Figure 2.6 Size Exclusion Chromatography of the pCU1 TraI Relaxase.....	26
Figure 2.7 Structural Comparison of the Relaxase N-terminus.....	27

Chapter 3

Figure 3.1 Circular Dichroism Spectra of the pCU1 TraI Relaxase.....	59
Figure 3.2 DNA Binding by the pCU1 TraI Relaxase.....	60
Figure 3.3 Schematics of DNA Nicking and Cross-over Experiments.....	61
Figure 3.4 Effect of Metal on the DNA Nicking Activity of the pCU1 TraI Relaxase.....	62
Figure 3.5 DNA Nicking Activity of pCU1 TraI Relaxase Tyrosine Mutants.....	63
Figure 3.6 DNA Covalent Attachment Activity of pCU1 TraI Relaxase Tyrosine Mutants.....	64
Figure 3.7 DNA Nicking Activity of pCU1 TraI Relaxase Aspartic Acid Mutants.....	65
Figure 3.8 Model of DNA Nicking by the pCU1 TraI Relaxase Active Site.....	66
Figure 3.9 Representative Gels Illustrating the DNA Nicking and Ligation Activities of the pCU1 TraI Relaxase.....	67
Figure 3.10 Four Regions of the pCU1 <i>oriT</i> DNA Sequence.....	68
Figure 3.11 Importance of the pCU1 TraI Relaxase N-terminus during DNA Nicking.....	69

Chapter 4

Figure 4.1 Classification of Helicases.....	99
Figure 4.2 Four Models of Helicase DNA Unwinding Mechanisms (from Matson, 2003).....	100
Figure 4.3 Structures of Representative Helicases and the Conserved RecA Core (from Caruthers, 2002).....	101
Figure 4.4 Predicted Structure of the pCU1 TraI Helicase.....	102
Figure 4.5 DNA Binding Curves.....	106
Figure 4.6 Representative DNA Binding Density Curves.....	108
Figure 4.7 Average Degree of Binding as a Function of Fluorescence Anisotropy.....	109
Figure 4.8 Maximum Average Degree of Binding as a Function of Substrate Length, Structure.....	110
Figure 4.9 Design of the NADH-coupled ATPase Activity Assay.....	111
Figure 4.10 pCU1 TraI ATPase Activity as a Function of ssDNA Length, NTP Concentration.....	112
Figure 4.11 Impact of pCU1 TraI Construct Length and Mutagenesis of the Walker A and B Motifs on pCU1 TraI ATPase Activity.....	113
Figure 4.12 Design of the ADP Quest ATPase Activity Assay.....	115
Figure 4.13 Validation of the ADP Quest ATPase Activity Assay Design.....	116
Figure 4.14 High Throughput Screening of the LOPAC against pCU1 TraI ATPase Activity.....	117
Figure 4.15 Putative Inhibitors of the pCU1 TraI ATPase Identified through High Throughput Screening.....	118

Chapter 5

Figure 5.1 Oxygen Biosensor Technology.....	126
Figure 5.2 Detection of Cell Population Size using the Oxygen Biosensor Plate Method.....	127
Figure 5.3 Plasmid pCU1 Transfer Protocol.....	128

List of Abbreviations and Symbols

%	percent
°	degree
Σv_i	average degree of binding
6FAM	fluorescein
6-His	6xhistine affinity tag
Å	angstrom
ADPNP	5'-adenylyl- β , γ -imidodiphosphate
AI	arabinose inducible
Ala, A	alanine
Amp	ampicillin
App. K_D	apparent dissociation constant
APS	Advanced Photon Source
APS	ammonium persulfate
Asn, N	asparagine
Asp, D	aspartic acid
ATP	adenosine triphosphate
AU	asymmetric unit
BME	β -mercaptoethanol
bp	base pairs
BSA	bovine serum albumin
C	Celsius
^{44}Ca	isotope calcium-44
$\text{C}\alpha$	alpha carbon
CBD	chitin binding domain

CCD	charge-coupled device
CCP4	Collaborative Computing Project No. 4
CD	circular dichroism
CFU	colony forming unit
CICBDD	Center for Integrative Chemical Biology and Drug Discovery
cm	centimeter
CNS	Crystallography and NMR System
CPT	conjugative plasmid transfer
CSS	complexation significance score
C-terminus, C-term	carboxy terminus
CTP	cytidine triphosphate
Cu ²⁺	copper cation
ddH ₂ O	double distilled water
δN	delta nitrogen
dmol	decimole
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
Dtr	DNA-transfer replication
DTT	D,L-dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
EMSO	electrophoretic mobility shift assays
εN	epsilon nitrogen
F plasmid	fertility plasmid
FA	fluorescence anisotropy

Fe ²⁺	iron cation (ferrous)
Fe ³⁺	iron cation (ferric)
⁵⁷ Fe	isotope iron-57
FPLC	fast protein liquid chromatography
g	gram
<i>g</i>	gravitational acceleration
Gln, Q	glutamine
Glu, E	glutamic acid
GTP	guanosine triphosphate
HGT	horizontal gene transfer
HIC-Up	Hetero-compound Information Center Uppsala
His, H	histidine
HPLC	high performance liquid chromatography
h	hour
HTS	high throughput screening
HUH	histidine residue-hydrophobic residue-histidine residue
ICP-MS	inductively coupled plasma mass spectroscopy
IC50	concentration of compound required for 50% inhibition of enzyme activity
ID	insertion device
IDT	Integrated DNA Technologies
IMPACT	Intein Mediated Purification with an Affinity Chitin-binding Tag
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	inverted repeat
K	Kelvin
K _D	dissociation constant
kD	kilodalton

K_{path}	molar absorption coefficient
L	liter
LB	luria broth
LDH	lactate dehydrogenase
LIC	ligation independent cloning
LOPAC	library of pharmaceutically active compounds
M	molar
MBP	maltose binding protein
μ	micro
μg	microgram
μL	microliter
μM	micromole
μm	micron
mdeg	millidegrees
mg	milligram
Mg^{2+}	magnesium ion
^{24}Mg	isotope magnesium-24
MIC	minimum inhibitory concentration
min	minutes
mL	milliliters
mm	millimeter
mM	millimole
Mn^{2+}	manganese cation
^{55}Mn	isotope manganese-55
MOB	mobility
MobA	mobilization protein A

mol	mole
MOPS	3-(N-morpholino)propanesulfonic acid
MPF	mating pair formation
MRE	mean residue ellipticity
MW	molecular weight
NADH	nicotinamide adenine dinucleotide
NEB	New England Biolabs
ng	nanogram
Ni	nickel
Ni ²⁺	nickel cation
⁶⁰ Ni	isotope nickel
nL	nanoliter
nm	nanometer
nM	nanomole
N-terminus, N-term	amino terminus
NTP	nucleoside triphosphate
OD	optical density
O/N	overnight
<i>oriT</i>	origin of transfer
PAGE	polyacrylamide gel electrophoresis
PDB	Protein Data Bank
PEG	polyethyleneglycol
PEP	phosphoenolpyruvate
pH	negative log (base 10) of the molar concentration of hydronium ions
Phe, F	phenylalanine
PISA	Protein Interfaces, Surfaces and Assemblies

PK	pyruvate kinase
pKa	negative log (base 10) of the acid dissociation constant
pM	picomolar
ppb	parts per billion
R plasmid	resistance plasmid
R ²	coefficient of determination
Rep	replication
Rif	rifamycin
RMSD	root mean square deviation
RP4	replication protein 4
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
S	seconds
SEC	size exclusion chromatography
Ser, S	serine
SER-CAT	Southeast Regional Collaborative Access Team
SF	superfamily
SLS	static light scattering
S _N 2	bimolecular nucleophilic substitution
Spec	spectinomycin
ssDNA	single-stranded DNA
SSRL	Stanford Synchrotron Radiation Laboratory
Strep	streptomycin
T4CP	type 4 coupling protein
T4SS	type 4 secretion system
TAE	Tris-Acetate-EDTA

TAM	TAMRA
TBE	Tris-Borate-EDTA
TEMED	tetramethylethylenediamine
Tet	tetracyclin
TEV	tobacco etch virus
T _m	melting temperature
TM	template modeling
TraI	transfer initiation
T-strand	transferred strand
TTP	thymidine triphosphate
Tyr, Y	tyrosine
U	hydrophobic residue
UNC-CH	University of North Carolina at Chapel Hill
v	volume
w	weight
WT	wild type
XDS	X-ray Detector Software
Zn ²⁺	zinc cation
⁶⁶ Zn	isotope zinc

Chapter 1: Introduction to Conjugative Plasmid Transfer

1.1 An Overview of the Replication and Transfer of Genetic Material in Bacteria

Bacteria store genetic information on chromosomes and extra-chromosomal elements. Extra-chromosomal elements are often mobile, capable of moving between bacterial hosts and occasionally into and out of bacterial chromosomes; they include gene cassettes, bacteriophages, single-stranded and double-stranded plasmids, and transposons¹⁻³. Bacterial chromosomes are themselves mosaics of stable elements and mobile elements such as prophages, integrating conjugative elements (ICEs or conjugative transposons), and mobilizable genomic islands⁴.

Bacteria replicate and then disseminate genetic information using a variety of mechanisms, which are classified into two groups, based upon the direction of the transfer. During vertical gene transfer, a parent bacterium duplicates its genetic information and then divides into two daughter bacteria, donating a copy of its genetic material to each of its progeny. In this case, both extra-chromosomal elements and chromosomal material are duplicated and disseminated. During horizontal gene transfer (HGT), a parent bacterium replicates and transfers a mobile genetic element to another parent bacterium. HGT can occasionally result in the dissemination of chromosomal material, if this material is part of a mobile element, or if the material conscripts the transfer machinery of a mobile element. However, HGT is the primary route of transfer for extra-chromosomal elements¹⁻³.

An entire population of bacteria can quickly acquire new genetic elements through the process of HGT. As a result, HGT is the major route by which bacteria increase their genetic diversity. For example, the horizontal transfer of conjugative plasmids and transposons is implicated in the dissemination of antibiotic resistance genes and virulence factors among bacterial pathogens^{1,5-}

⁹. Due to the importance of HGT for the spread of antibiotic resistance genes and virulence factors,

we became interested in the underlying mechanism controlling this type of transfer between bacteria. In particular, we chose to focus on the horizontal transfer of double-stranded DNA (dsDNA) plasmids.

1.2 The Classification of Transmissible Plasmids

Horizontal gene transfer of double-stranded DNA (dsDNA) plasmids between bacterial cells is specifically referred to as conjugative plasmid transfer (CPT). Transmissible dsDNA plasmids are divided into two broad classes based upon their ability to move between bacterial hosts (Figure 1.1). The first class, referred to as conjugative plasmids, encode all the factors necessary for their own transfer, while the second class, the mobilizable plasmids, require the presence of a second conjugative element for transfer. Both conjugative and mobilizable plasmids contain genes encoding the mobility (MOB) set of enzymes, which are required for plasmid replication. In addition to these MOB elements, conjugative plasmids contain genes encoding a Type 4 secretion system (T4SS). The T4SS forms the pore through which DNA can move from the donor bacterium to the recipient. A type 4 coupling protein (T4CP) links the MOB replication apparatus to the T4SS transfer apparatus. All conjugative, and a limited number of mobilizable, plasmids encode their own T4CP. Even though the nomenclature would suggest otherwise, the physical transfer of both conjugative and mobilizable plasmids is referred to as conjugation^{2, 3, 8, 10, 11}.

The relaxase enzyme is the only highly conserved member of the MOB family of proteins, and it is necessary for the transfer of both mobilizable and conjugative plasmids. Therefore, the presence of a relaxase allows for the identification and classification of bacterial DNA plasmids that undergo CPT. The relaxase enzyme is defined by the presence of a limited number of amino acid motifs. The HUH motif, also referred to as motif 2, consists of a histidine-hydrophobic residue-histidine and is the most highly conserved relaxase motif. Often a third histidine, found slightly upstream of motif 2 in amino acid sequence, forms a triad with the two histidines of the HUH motif. Together, the three histidines are referred to as the Histidine Triad or HUH(+H) motif. The HUH(+H) motif is responsible for coordinating a metal cation in the relaxase active site. Motif 3,

found upstream of motif 2 in amino acid sequence, consists of one or more tyrosines and is responsible for relaxase-mediated DNA plasmid nicking and ligation. Of note, some papers refer to this tyrosine-dominated motif as motif 1; in this manuscript we will use the original nomenclature which refers to the motif as motif 3^{8, 10, 12, 13}. Both of these motifs were first identified in a related class of enzymes, the Rep class of enzymes. In Rep enzymes, the motifs are found in numerical order in amino acid sequence (the HUH motif 2 is followed by the tyrosine-dominated motif 3) and a highly variable 3rd motif, motif 1 (futLt, U[VL][VL]YP, FLTY(P), or FLTLT) is found N-terminal of both motifs 2 and 3^{12, 14, 15}.

Based upon the sequence of their respective relaxase enzymes, and in particular the sequence surrounding motifs 2 and 3, mobilizable and conjugative plasmids have been classified into six families. These families are MOB_F, MOB_H, MOB_C, MOB_Q, MOB_P, and MOB_V (Figure 1.1). The MOB_Q, MOB_P, and MOB_V plasmid families encode relaxases containing only one motif 3 tyrosine. The well characterized plasmid R1162 MobA relaxase (MOB_Q) and plasmid RP4 TraI relaxase (MOB_P) are representative single tyrosine relaxases¹⁶⁻²². Relaxases encoded by the MOB_H and MOB_C plasmids are not well characterized, and it is still unclear how many tyrosines these enzymes typically include within motif 3. The MOB_F family is the best characterized of the six plasmid families and has been subdivided into two clades (MOB_{F1}, MOB_{F2}) and numerous subclades. For all six MOB families of plasmids, the plasmid-encoded relaxase enzyme is often found at the N-terminus of a multidomain transfer-initiation protein. A helicase or primase domain can be located downstream of the relaxase, and in some cases additional protein-protein interaction domains are present^{3, 8, 10, 23-25}.

In 1947, Lederberg and Tatum reported the transfer of a MOB_{F1} plasmid, the F plasmid, between bacterial cells; their work represented the first description of the horizontal transfer of genetic material between bacteria. Since then, the F plasmid and its relaxase have been extensively studied, and the F plasmid has served as the best characterized model of CPT^{5, 13, 26-35}. Recently, additional members of clade MOB_{F1} and their respective relaxases have been analyzed, including the

plasmid R388 TrwC and its relaxase³⁶⁻⁴⁴ and, as reported here, the plasmid pCU1 and its TraI relaxase^{45, 46}. By characterizing these and other conjugative plasmids, the general mechanism of plasmid transfer can be determined, as well as the role that the relaxase plays during transfer.

1.3 The Mechanism of Conjugative Plasmid Transfer

During CPT, a donor bacterium transfers one strand of a double-stranded DNA plasmid to a neighboring recipient bacterium⁴⁷. As detailed above, each conjugative plasmid contains genes for a MOB protein complex, a T4SS protein complex, and a T4CP, all of which are necessary for its transfer (Figure 1.1). The MOB, or Dtr (DNA-transfer replication), set of proteins forms a large complex called the relaxosome in concert with the host-encoded Integration Host Factor (IHF). The relaxosome includes the relaxase and is the replication machinery of the plasmid. The T4SS complex, which forms a secretion system through which DNA travels to the recipient, can also be referred to as the mating pair formation (MPF) complex. The coupling protein that links these two complexes is the type 4 coupling protein (T4CP) and is considered by some to be one of the MOB proteins, though this designation is not universally accepted^{3, 11, 48}.

Within the relaxosome, the relaxase initiates plasmid transfer by creating a single-stranded break at a *nic* site in the transferred strand (T-strand) of the dsDNA plasmid (Figure 1.2). This *nic* site is found within a larger segment of the T-strand called the origin of transfer, or *oriT*. The *oriT* serves as the binding site for the relaxase and other members of the relaxosome. The relaxase cleaves the T-strand at the *nic* site with one of its motif 3 tyrosines. In particular, the tyrosine initiates a bimolecular nucleophilic substitution-type (S_N2) attack on the scissile phosphate linking the two nucleotides of the T-strand *nic* site, thus generating a free 3' hydroxyl and a covalent phosphotyrosine bond (Figure 1.2, Figure 1.3). A DNA helicase then separates the T-strand from the parent strand, beginning at the *nic* site. In many cases, this helicase is found C-terminal of the relaxase on a multidomain protein. As the T-strand is freed from the parent strand, it begins to travel into the recipient cell, most likely driven by the T4CP. Some data indicate that the covalently bound relaxase also travels into the recipient at this time, serving to pilot the T-strand through the T4SS pore (Figure

1.2, right-sided path). In this case, a second relaxase-helicase protein in the donor is required to unwind the T-strand and terminate transfer. A DNA polymerase within the donor bacterium replaces the T-strand as it is transferred to the recipient, starting at the free 3' hydroxyl at the *nic* site, thus generating the T'-strand. As a result, a hybrid T/T' strand *nic* site is created, and no genetic information is lost by the donor^{2, 11, 23, 28, 47-49}.

Once the T-strand is completely separated from the parent strand, it must be severed from the newly synthesized T'-strand at the hybrid *nic* site in order to release it to the recipient (Figure 1.2). This second DNA-nicking step is accomplished by a nucleophilic group capable of initiating an S_N2 attack on the hybrid *nic* site's scissile phosphate. If the relaxase has traveled into the recipient, or if the relaxase contains only one motif 3 tyrosine, then a second relaxase enzyme in the donor could provide a nucleophilic DNA nicking tyrosine. If the relaxase remains in the donor, and it contains multiple tyrosines within motif 3, then it could provide the required nucleophile itself. Alternatively, transfer could stall until the host polymerase completes synthesis of the T'-strand, and the free 3' hydroxyl at the end of the T'-strand could act as the attacking nucleophile. Most conjugative systems characterized to date utilize either multi-tyrosine relaxases or synthesize multiple relaxases during CPT. Regardless of the mechanism employed, following the second nick, a series of ligation reactions re-circularizes the original T-strand and the new T' strand, and the relaxase(s) is(are) released from the plasmids^{2, 23, 28, 47, 49, 50}.

Since CPT is responsible for the spread of bacterial virulence and antibiotic resistance throughout populations of bacterial pathogens, an understanding of this process could lead to the development of novel antibiotics targeting resistance-spreading bacterial infections. The resistance-encoding conjugative plasmid pCU1 encodes a multidomain TraI protein that supplies the relaxase and helicase activities necessary for pCU1 transfer. Therefore, the goals of the work presented here are to first characterize the structure and function of the relaxase and helicase domains of the pCU1 TraI protein, and to then compare these activities to those of other homologous conjugative enzymes. Finally, we hope to use this information to identify potential routes by which the relaxase or helicase

activity of the TraI protein could be inhibited. By blocking either the relaxase or helicase activity of pCU1 TraI, spread of this antibiotic resistance-encoding plasmid would be prevented.

1.4 Figures

Figure 1.1 The Classification of Transmissible Plasmids

Figure 1.2 The Mechanism of Conjugative Plasmid Transfer

Figure 1.3 DNA Nicking Mechanism of the Relaxase Motif 3 Tyrosine

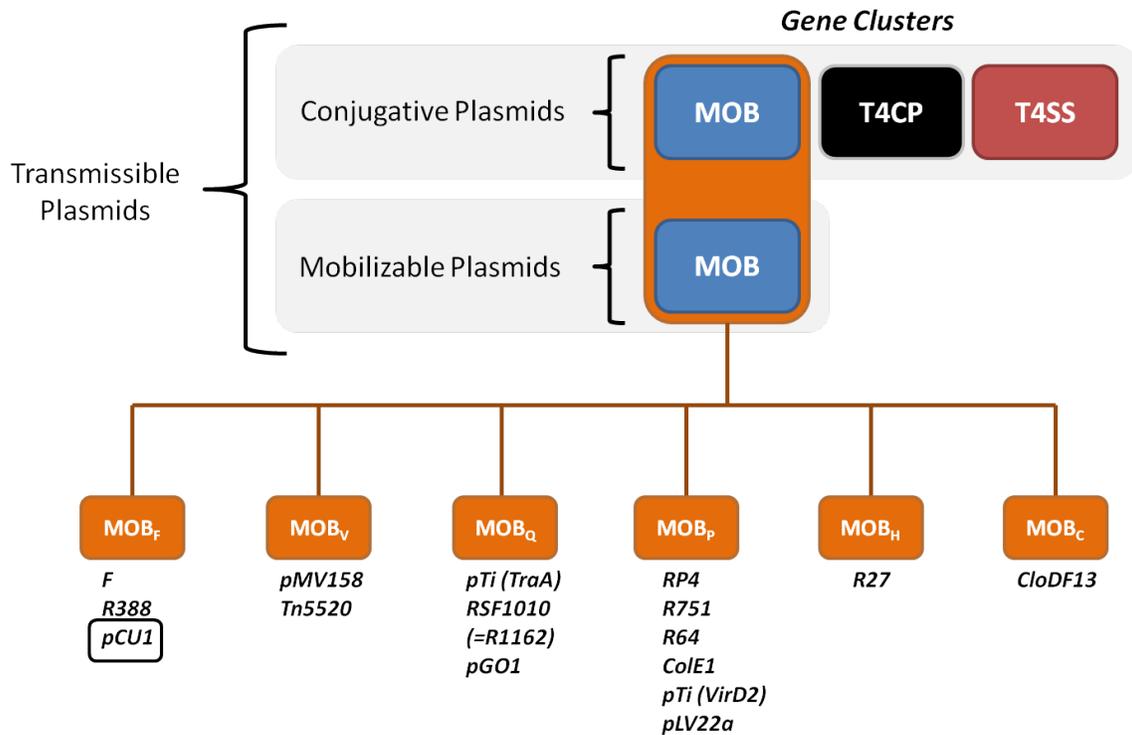


Figure 1.1 The Classification of Transmissible Plasmids

Transmissible plasmids are classified as either conjugative or mobilizable based upon the number of gene clusters they contain and the method by which they are transferred. Mobilizable plasmids contain the MOB gene cluster, while conjugative plasmids contain the MOB gene cluster as well as the Type 4 coupling protein (T4CP) and the type 3 secretion system (T4SS) gene clusters. Transmissible and conjugative plasmids have been classified into 6 families (MOB_F, MOB_H, MOB_Q, MOB_P, MOB_V, MOB_C) based upon the sequence of their relaxase enzyme, a conserved member of the MOB family of proteins. Representative members of each family are provided. Plasmid RSF1010 is also referred to as plasmid R1162. Plasmid pTi encodes two relaxase enzymes that are classified as either MOB_Q (TraA) or MOB_P (VirD2)^{3, 8, 10}.

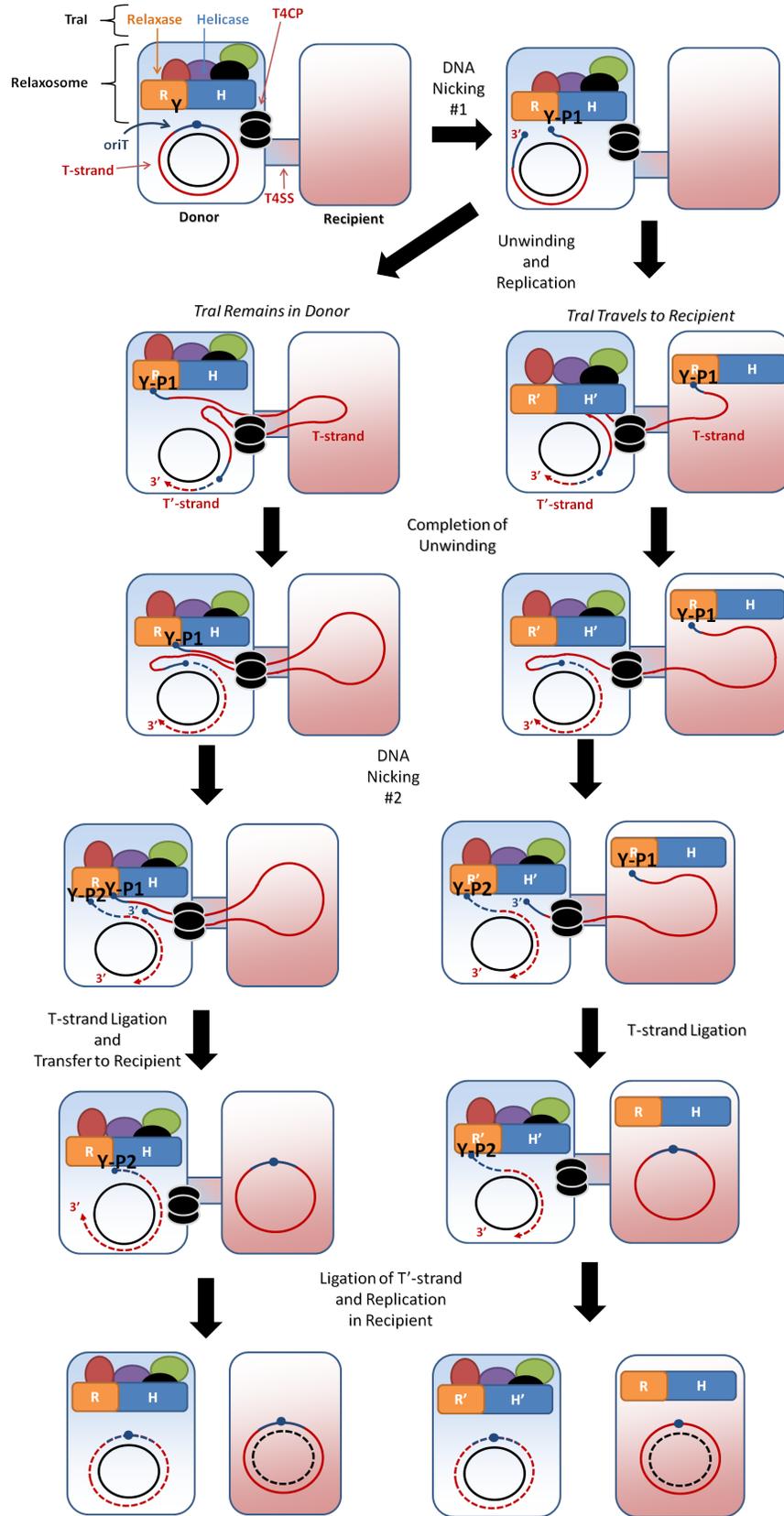


Figure 1.2 The Mechanism of Conjugative Plasmid Transfer

Plasmid transfer is dependent on the activities of a conjugative relaxase and helicase. Conjugative relaxases contain one to four DNA nicking tyrosines and are often located at the N-terminus of a multidomain protein with a C-terminal helicase or primase domain. Plasmid pCU1 TraI is a multidomain protein consisting of an N-terminal multityrosine relaxase (orange rectangle, labeled R) and a C-terminal helicase (blue rectangle, labeled H). The pCU1 TraI relaxase domain and other members of the relaxosome protein complex bind the plasmid *oriT* (colored blue) of the transferred (T) strand (colored red). To initiate transfer, the relaxase nicks the T strand *nic* site (blue dot). This generates a phosphotyrosine bond between the relaxase and the 5' end of the T-strand (labeled Y-P1), as well as a free 3' hydroxyl (red 3'). The TraI helicase domain then unwinds the T-strand from the parent strand (colored black) starting at the *nic* site, while a host polymerase replaces the displaced T-strand starting at the free 3' hydroxyl. The newly generated DNA strand is dashed and labeled "T'-strand". The unwound T-strand passes through a pore created by the Type 4 secretion system (T4SS), driven by the pumping activity of the type 4 coupling protein (T4CP), and enters the recipient. Data indicate that for some systems the TraI relaxase likely travels into the recipient cell, acting as a pilot protein for the T-strand. Therefore two possible routes for plasmid transfer are illustrated. If TraI enters the recipient cell (right path), a second TraI enzyme (labeled R' and H') is responsible for providing the helicase activity in the host cell. After the T-strand is completely separated from the parent strand, the relaxase (either R or R') creates a second nick at the hybrid *nic* site within the hybrid *oriT* (intersection of the solid and dashed lines), generating a second phosphotyrosine linkage (labeled Y-P2) and a free 3' hydroxyl on the T-strand (blue 3'). To complete transfer, the unwound T-strand is recircularized when its freed 3' hydroxyl attacks Y-P1. Once replication is complete, the newly generated T'-strand is recircularized when the 3' end (red 3') attacks Y-P2.

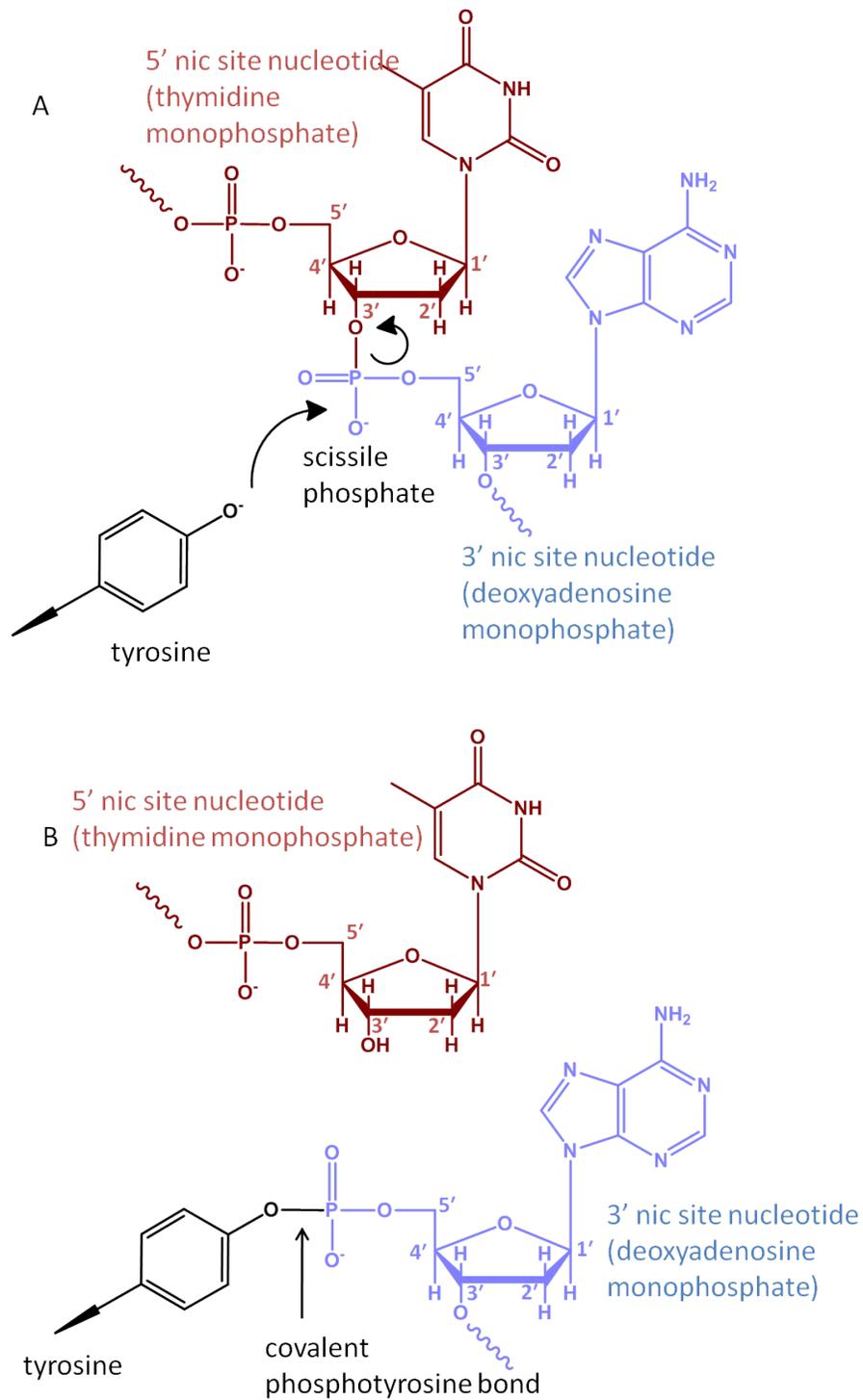


Figure 1.3 DNA Nicking Mechanism of the Relaxase Motif 3 Tyrosine

The tyrosine residue of the relaxase initiates an S_N2 attack on the scissile phosphate at the plasmid *nic* site, breaking the bond between the 3' sugar hydroxyl and the 5' phosphate and becoming covalently attached to the 5' phosphate. A) Reactants. B) Products. Sugar carbons are numbered.

Chapter 2: The Structure of the Plasmid pCU1 TraI Relaxase

2.1 Introduction to the Structure of the Relaxase

As discussed in Chapter 1, a relaxase is found encoded on all transmissible plasmids and is required for conjugative plasmid transfer (CPT)^{8,10}. Analysis of relaxase structure and function will provide insight into the crucial steps this protein accomplishes during CPT. To date, structures of the F plasmid TraI relaxase (subclade MOB_{F12}), the plasmid R388 TrwC relaxase (subclade MOB_{F11}), and the plasmid R1162 MobA relaxase (MOB_Q) have been described^{5,16,26,36,39,51}. In addition, the F and R388 relaxase structures have been solved in complex with a portion of their respective plasmid's *oriT* sequence. These relaxase-DNA complexes reveal that the relaxase binds the *oriT* as a partial DNA hairpin (Figure 2.1)^{36,39}. The 5' end of the bound *oriT* forms a hairpin, while the 3' end extends as a single strand and forms a U-turn around the protein N-terminus, before entering the active site of the enzyme^{5,29,36,39}. As seen in the structures of these three relaxases, the enzyme active site consists of the HUH motif's triad of histidines, which are located on two strands of a central β -sheet. As described in Chapter 1, this motif coordinates a metal cation, the identity of which is plasmid-specific^{5,17,36,51}. In DNA-bound structures of the relaxase, the DNA scissile phosphate is positioned above the coordinated metal. As a result, the metal cation withdraws electron density from the phosphate center to promote nucleophilic attack on the scissile phosphate by one of the enzyme's conserved motif 3 tyrosine residues^{5,36,51}.

We have determined the atomic structure of the relaxase enzyme responsible for the transfer of conjugative plasmid pCU1 (subclade MOB_{F11}) (Figure 2.2). Our goal is to correlate the pCU1 relaxase structure with the mechanism of the enzyme and to compare the pCU1 structure with those described above. pCU1 is an antibiotic resistance-encoding plasmid originally isolated from *Salmonella typhimurium*. It is a derivative of resistance plasmid R46 and confers on its host

resistance to the antibiotics ampicillin, streptomycin and spectinomycin^{45, 46, 52-54}. The relaxase activity encoded by pCU1 is located within the N-terminal 299 residues of a multifunctional TraI enzyme, and productive, plasmid-specific transfer is dependent upon TraI^{45, 46}. Briefly, our analysis of the pCU1 TraI relaxase structure identified features unique to the pCU1 relaxase (Figure 2.3) that allowed us to advance our understanding of the general mechanism of relaxase-mediated DNA binding and cleavage, as well as to highlight distinctive features of the pCU1 plasmid system.

2.2 Construct Design and Cloning of the pCU1 TraI Relaxase

As stated above, the relaxase of plasmid pCU1 is located within the plasmid-encoded TraI protein^{45, 46}. To determine the extent of the relaxase domain within pCU1 TraI, we aligned the pCU1 TraI amino acid sequence (GenBank: AAD27542) with that of R388 TrwC (GenBank: CAA44853) and F TraI (GenBank: BAA97974) (Figure 2.4). Sequence alignments were performed using AlignX or ContigExpress, both of which are components of Vector NTI Express 10.0.1 (Invitrogen, 2005), or using ClustalW within the BioEdit Sequence Alignment Editor (version 7.0.9.0, Tom Hall, 2007)^{55, 56}. When using ClustalW, default parameters were used, except for the gap open penalty which was adjusted to 5.0 for both pairwise and multiple alignment steps. The Blossum62 similarity matrix was used during alignments, when calculating percent sequence identity, and when annotating the alignment. The sequence identity between the full length pCU1 TraI and R388 TrwC was 42%, and the sequence identity between the full length pCU1 TraI and F TraI, 16%; these values increased to 48% and 33%, respectively, when restricting evaluation of the sequences to the N-terminal 330 residues of each enzyme. Residues 1 to ~300 of pCU1 TraI were found to encapsulate the regions of highest sequence homology among these three sequences, and contained all conserved, relaxase-specific motifs. Therefore, we determined that residues 1 to ~300 of pCU1 TraI corresponded to the relaxase domains of F TraI and R388 TrwC. Figure 2.4 (sequence alignment) was created in BioEdit.

Secondary structure predictors (Jpred3, version 2.2, <http://www.compbio.dundee.ac.uk/www-jpred/>) identified loop regions near residue 300, so the pCU1 TraI relaxase construct extending from

residue 1 to 299 was chosen as the final construct (Figure 2.4). Primers were designed to isolate residues 1-299 of pCU1 TraI using the Vector NTI Express 10.0.1 suite of programs (Invitrogen, 2005) and were then commercially synthesized (Integrated DNA Technologies (IDT)). All cloning and mutagenesis was verified by sequencing at the UNC-CH Genome Analysis Facility. Residues 1-299 of pCU1 TraI were cloned into the vector pTYB2 (IMPACT system, New England Biolabs), between *NdeI* and *SmaI* restriction sites in order to C-terminally fuse the relaxase to an intein and a chitin binding domain (CBD) affinity tag and thus generate the construct WT_299. Using ligation independent cloning, residues 1-299 of the relaxase domain were inserted into vector pMCG9⁵⁷ to create the mutant construct Nterm_299. In vector pMCG9, the relaxase was N-terminally fused to a maltose binding protein (MBP) affinity tag and a 6-His (MBP-HIS) affinity tag with a tobacco etch virus (TEV) protease cleavage site located between the protein sequence and the affinity tags⁵⁷. During purification, these affinity tags were removed by the TEV protease, leaving three non-native residues (serine-asparagine-alanine) at the N-terminus of Nterm_299.

2.3 Expression and Purification of the pCU1 TraI Relaxase

All proteins were over-expressed in BL21(DE3) *Escherichia coli*. One liter of LB broth was inoculated at a ratio of 1:50 or 1:100 with a saturated overnight culture of BL21(DE3) *Escherichia coli* containing the expression plasmid. Cells (1 L) were grown under antibiotic selection (ampicillin, 100 µg/mL) at 37°C with vigorous shaking until the cell density reached an OD₆₀₀ of 0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM and the temperature was dropped to 18°C. The protein was overexpressed at 18°C for 16 h, after which the cells were harvested by centrifugation (15 min at 4,500 rpm at 4°C) and resuspended in 10mL Buffer C (WT_299) (500 mM NaCl, 20 mM Tris-HCl pH 7.5, 10% glycerol, 0-5 mM EDTA, 0.01% azide) or Buffer A (Nterm_299) (500 mM NaCl, 20 mM dibasic Na₂(PO₄), 10% glycerol, 5 mM imidazole, 0-1 mM EDTA, 0.01% azide, pH 8.0) per 1 L growth. Resuspended cells were stored at -80°C.

Prior to purification of both WT_299 and Nterm_299, 50mL of resuspended cells were thawed in the presence of lysozyme, DNase (50 µg/mL) and a cocktail of protease inhibitors (Roche,

1 table per 50 mL resuspended cells), lysed on ice using a Sonic Dismembrator, Model 500 (Fisher Scientific) (2 min of 0.5 s pulses at 60% intensity), and centrifuged (70 min at 17,000 rpm at 4°C) to isolate the soluble fraction. All purification steps were performed on ice or at 4°C.

WT_299 was purified on chitin resin (NEB) using a batch bind method followed by an extended wash step in Buffer C, as described by the manufacturer. Cleavage of the intein and CBD tags, and subsequent release of WT_299 from the chitin resin, was induced by incubating the chitin resin with 50 mM D,L-dithiothreitol in Buffer C for 16 h. WT_299 was then eluted from the chitin resin in Buffer C by gravity column chromatography. Purification of one liter growth resulted in approximately 5 mg purified protein. WT_299 samples were separated by SDS-PAGE, stained with Coomassie Brilliant Blue, de-stained, and found to be >95% pure (Figure 2.5A, Figure 2.6).

Nterm_299 was purified over a 5 mL HisTrap HP column (GE Healthcare) on an AKTAexpress FPLC (GE Healthcare) pre-equilibrated in Buffer A. Fusion protein was eluted with a high imidazole Buffer B (500 mM NaCl, 20 mM dibasic Na₂(PO₄) pH 8.0, 10% glycerol, 500 mM imidazole, 0-1 mM EDTA, 0.01% azide) after washing the column to baseline, as monitored by A₂₈₀, with Buffer A. The affinity tags were removed by cleavage with 4% (w/w) TEV protease during dialysis overnight into Buffer A at 4°C. The protein was then concentrated to ~2 mg/mL and further purified over two sequential 5 mL HisTrap HP gravity columns. Columns were pre-equilibrated in Buffer A. Protein was loaded on to the first column and eluted in Buffer A. Eluent was monitored for the presence of protein using a Bradford reagent. The initial flow-through was collected and then immediately loaded onto a second column. Again, the initial flow-through was collected and concentrated. Purification of one liter growth resulted in approximately 5 mg purified protein. Nterm_299 samples were separated by SDS-PAGE, stained with Coomassie Brilliant Blue, de-stained, and found to be >95% pure (Figure 2.5C).

For crystallization, WT_299 and Nterm_299 were dialyzed into 250 mM ammonium acetate while concentrating to 5-6 mg/mL (Nterm_299) or ~14 mg/mL (WT_299), then used immediately for

crystallization. All water was obtained from the laboratory Barnstead E-pure water filtration system, at > 17 megaohm-cm.

2.4 Crystallization, Data Collection, and Data Processing of the pCU1 TraI Relaxase

Crystals of WT_299 and Nterm_299 were grown at 20°C by the hanging-drop vapor diffusion method. Equal volumes of purified protein (WT_299 or Nterm_299) and well solution (250 mM triNa citrate, 22% PEG 3350 (untitrated)) were mixed and crystals grew over a course of 4-5 weeks. Crystals were initially small and hexagonal, but following seeding and the addition of 5 mM MgCl₂ to the purified protein sample, larger tetragonal crystals were obtained. The crystals were cryoprotected in 250 mM triNa citrate, 26% PEG 3350, and 10% ethylene glycol and flash frozen in liquid nitrogen for data collection at 100 K.

Crystals were initially screened at the UNC Biomolecular X-ray Crystallography Facility with a Rigaku R-Axis IV++ and CCD, while full data sets were collected at two different synchrotron sources. For crystals of WT_299, x-ray diffraction data were collected at Southeast Regional Collaborative Access Team (SER-CAT) Sector 22-ID beamline at the Advanced Photon Source (APS), Argonne National Laboratory. Two second exposures, with a 1.0 degree rotation between exposures were used to collect 180 degrees of data for each of six crystals. Data from the best diffracting crystals were then indexed and scaled with X-ray Detector Software (XDS, MPI for Medical Research, Heidelberg, 2009, <http://xds.mpimf-heidelberg.mpg.de/>). For crystals of Nterm_299, x-ray diffraction data were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) using remote access of beamline 9-2 ID. Ten second exposures, with a 1.0 degree rotation between exposures, were used to collect 360 degrees of data for each of four crystals. Data from the best diffracting crystals were then indexed and scaled with HKL2000 (HKL Research, Inc., 2005, <http://www.hkl-xray.com>). Initial phases of data from both constructs were determined by molecular replacement in Phaser (<http://www-structmed.cimr.cam.ac.uk/phaser>), a component of the Collaborative Computing Project No. 4 (CCP4) (<http://www.ccp4.ac.uk>, <http://ccp4wiki.org>) using the relaxase domain of TrwC as the search model (PDB code 1OMH). Prior to phasing, the bound

DNA substrate and TrwC protein residues 234-271 were removed from the PDB file (1OMH), and the side chains of nonconserved residues were mutated to alanines using Chainsaw^{58,59}, a component of CCP4. Data were refined using refmac (<http://www.yasbl.york.ac.uk/~garib/refmac/>), a component of CCP4, CNS (version 1.2, <http://cns.csb.yale.edu/v1.2>), and Phenix (version 1.6, <http://www.phenix-online.org>), and model building was performed in Coot (<http://biop.ox.ac.uk/cool>). Ligands were found in the standard monomer library in CCP4, on the HIC-Up server (release 12.1, <http://xray.bmc.uu.se/hicup/>), or were constructed using the Dundee PRODRG2 server (<http://davapc1.bioch.dundee.ac.uk/prodrg/>). Data were verified using MolProbity (<http://molprobity.biochem.duke.edu>). Figures were constructed in PyMOL (DeLano Scientific LLC, San Carlos, CA, USA, <http://www.pymol.org>, 2009). AreaMol, a component of CCP4, was used to calculate the solvent exposed surface area of each monomer within the asymmetric unit of the crystal. The character of the dimer interface was analyzed using the Protein Interfaces, Surfaces and Assemblies service (PISA) (version 1.18, http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html). To calculate the average thermal displacement parameters (B values), Baverage, a component of the CCP4, was used. Structural alignments were performed with TM-align (<http://zhanglab.ccmb.med.umich.edu/TM-align>), using the ProCKSI-Server (version procksi-8.7, www.procksi.net), to generate the root mean square deviation (RMSD) of Ca positions of the proteins analyzed. Within Coot, the OptAlign command, using the Kabsch algorithm (<http://pymolwiki.org/index.php/Kabsch>), was used to generate the RMSD of atoms within the histidine triad of the proteins analyzed. The coordinates of WT_299 and Nterm_299 have been deposited in the PDB (accession codes 3L57, 3L6T).

2.5 Analysis of the pCU1 TraI Relaxase Structure

2.5.1 The Dimer in the Asymmetric Unit

As detailed above, structures of the pCU1 TraI relaxase were determined by x-ray crystallography. The resulting data were determined to 2.3 Å (WT_299) and 1.9 Å (Nterm_299) (Table 2.1). The WT_299 and Nterm_299 crystallized as dimers in the asymmetric unit (asu) in the

space group P2₁. Upon analysis of the structural data describing WT_299, the two monomers were determined to have a combined 20,600 Å² of solvent accessible surface area of which 1,670 Å² was buried upon dimerization. While this analysis could suggest a physiologically relevant dimer⁶⁰, only 18 potential hydrogens bonds and 4 potential salt bridges could be formed at the dimer interface, and the structure's Complexation Significance Score (CSS) of 0.000 indicated that the dimer was most likely to be a result of crystal packing⁶¹. An analysis of Nterm_299 yielded similar results, with only 12 potential hydrogen bonds, 4 potential salt bridges and an equally weak CSS score. In addition, the WT_299 and Nterm_299 proteins were determined to be monomers by size exclusion chromatography (SEC) (Figure 2.5B, Figure 2.6; see Section 3.2 for details concerning purification by SEC; Nterm_299 data not shown). For both WT_299 and Nterm_299, the electron density of monomer A was consistently higher in quality than that of monomer B. As a result, 223 (WT_299) or 226 (Nterm_299) of the 299 residues were confidently modeled into the electron density of monomer A, while only 203 (WT_299 and Nterm_299) were placed into the electron density of the B monomer (Table 2.1).

2.5.2 The Overall Fold and the Active Site HUH Motif of the pCU1 TraI Relaxase

The overall fold of the pCU1 TraI relaxase domain resembles that of a human hand (Figure 2.2) as previously described for homologous relaxase structures (Figure 2.1)^{5, 16, 26, 36, 39, 51}. The palm consists of five central anti-parallel β-strands; the surrounding α-helices and β-sheets form the fingers, back of the hand and bottom edge of the palm. The N-terminus is located at the beginning of the first β-strand (β1), and the remaining strands that compose the palm are in the order β1-β3-β7-β6-β2, when looking at the “palm side” of the protein. The C-terminal 44 (monomer A) and 46 (monomer B) residues of WT_299 are disordered, but are predicted to be α-helical (Figure 2.4). Many of these residues have been placed in structures of homologous proteins, but the entire region has only been observed when the relaxase is in complex with its DNA substrate^{5, 36}.

The active site of WT_299 and Nterm_299 is found within the palm of the hand, and consists of three histidines, two on β7 (H160, H162) and one on β6 (H149) (Figure 2.2). Together, they form

the HUH(+H) motif and coordinate a divalent metal cation, occupying three of the bound metal's coordination sites. The remaining coordination sites surrounding the divalent cation are satisfied by a variety of compounds, such as water molecules or components from the cryoprotectant or well solution. For example, in the structure of WT_299, these remaining sites were occupied by the small molecule citrate (monomer A), or a water molecule and ethylene glycol (monomer B). Initially, several divalent metals were sequentially modeled into the active site of WT_299 to fulfill the electron density bound by the histidines. Mn^{2+} , Ni^{2+} , Zn^{2+} , Fe^{2+} , and Cu^{2+} all satisfied the electron density as assessed by difference density and R-factors. Mn^{2+} was later determined to be the most likely metal in the active site based on a series of activity assays and analysis by inductively coupled plasma mass spectrometry (ICP-MS), as described in Chapter 3. In both monomers of Nterm_299, Ni^{2+} was placed within the enzyme active site due to its ability to satisfy the electron density as assessed using difference density maps and R-factors, the use of a Ni-containing column during the purification of Nterm_299, and corroborating results from ICP-MS analysis of the protein. Other than a difference in the identity of the active site metal and the presence of three non-native residues at the N-terminus of Nterm_299 (serine-2, asparagine-1, and alanine0), the overall structures of the two constructs were highly similar with an 0.33 Å RMSD over 220 equivalent $C\alpha$ positions and a TM score of 0.97454^{62, 63}.

2.5.3 The N-terminus of the pCU1 TraI Relaxase

To consider the impact the three non-native N-terminal residues could have on the function of Nterm_299, the structures of WT_299, Nterm_299, and the DNA-bound relaxase domain of plasmid R388 TrwC³⁶ (PDB code 2CDM) were compared (Figure 2.7). The overall folds of the three proteins are similar (maximum RMSD over equivalent $C\alpha$ positions = 2.05 Å, Figure 2.7a,b), and the active site HUH motifs are practically identical in orientation (average RMSD of the three residues = 0.076 Å). However, the three non-native N-terminal residues of Nterm_299 clearly clash with the bound DNA substrate of TrwC (Figure 2.7c,d). The DNA substrate of the TrwC relaxase is a 27mer oligonucleotide encoding a portion of the plasmid R388 *oriT* sequence, and it is bound by the

relaxase as a partial hairpin. The DNA hairpin covers the surface of the of the relaxase above the active site, while a downstream single-stranded segment then wraps around the N-terminus of TrwC, making a hard U-turn into the active site (Figure 2.1). In Figure 2.7c and d, serine-2 of Nterm_299 clashes sterically with thymidine25 of the bound TrwC substrate, aspartic acid-1 clashes with the base and sugar of cytosine24, and alanine0 clashes with the base of guanine22 (all nucleotide base numbering reflects that of the TrwC structure, PDB code 2CDM). These clashes predict that Nterm_299 should be incapable of binding its plasmid's *oriT* without either physically moving its own N-terminus out of the way, or by binding the DNA in a novel orientation. Both options would disrupt sequence-specific interactions formed between the enzyme and the *oriT* DNA, in particular those between the N-terminus of the protein and the U-turn of the bound DNA. A series of DNA nicking assays, as described in Chapter 3, validated these structural implications. The importance of such DNA-protein contacts has also been illustrated during DNA binding and nicking assays involving the F TraI relaxase²⁹.

2.5.4 The Motif 3 Tyrosines of the pCU1 TraI Relaxase

The four tyrosines of WT_299 and Nterm_299 that are implicated in DNA nicking (Y18, Y19, Y26, and Y27) are located on Loop A, adjacent to the active site of the enzyme (Figure 2.2). In monomer B, only the first few residues of the loop were sufficiently ordered to be modeled into the structure, but in monomer A the entire loop was observed. On the loop, Y18 and Y19 were found displaced from the active site, flipped out and pointing away from the bound metal cation (Figure 2.2, 2.3). In the F TraI relaxase structure, the equivalent tyrosine residues (Y16, Y17) are oriented towards the active site, with Y16 (equivalent to pCU1 TraI relaxase Y18) positioned to attack the scissile phosphate of the DNA substrate^{5, 26, 51}. The equivalent tyrosines (Y18, Y19) in the R388 TrwC relaxase are also directed towards the active site, with Y18 shifted slightly behind the scissile phosphate, relative to Y16 of the F TraI relaxase^{26, 36, 39}. When the structure of WT_299 monomer A was overlaid with the R388 relaxase (PDB ID 2CDM) and F relaxase (PDB ID 1P4D) (Figure 2.3),

Y18 of WT_299 was displaced 14.2 Å away from the histidine triad as compared to Y16 of F relaxase and 12.4 Å as compared to Y18 of R388 relaxase.

2.6 Conclusions

Four conclusions are drawn from a structural analysis of the pCU1 TraI relaxase. First, the pCU1 TraI relaxase maintains the conserved fold of the MOB family of relaxase enzymes. Second, Loop A of the pCU1 TraI relaxase, and its accompanying four catalytic tyrosines, assume a unique orientation as compared to that of homologous relaxases. In particular, the DNA nicking tyrosines of Loop A are rotated away from the active site and are displaced from the metal-bound histidine triad (Figure 2.3). Third, the extended N-terminus of Nterm_299 is predicted to clash with DNA bound by the enzyme (Figure 2.7). In fact, as will be seen in Chapter 3, extension of the N-terminus inhibits nicking and elicits mis-nicking by the mutant enzyme. Finally, this structural data, in combination with ICP-MS analysis of WT_299 (see Chapter 3, Table 3.4), and DNA nicking data describing WT_299 (see Chapter 3, Figure 3.4), identifies Mn^{2+} as the metal most likely bound by the HUH motif of the pCU1 TraI relaxase.

2.7 Tables and Figures

Table 2.1 Crystallographic Statistics

Figure 2.1 The Structure of the Conjugative Relaxase in Complex with Plasmid *oriT* DNA

Figure 2.2 Structure of the pCU1 TraI Relaxase

Figure 2.3 Structural Comparison of Relaxase Enzymes

Figure 2.4 Amino Acid Sequence Alignment and Secondary Structure of the pCU1 TraI Relaxase

Figure 2.5 SDS-PAGE of Purified pCU1 TraI Relaxase

Figure 2.6 Size Exclusion Chromatography of the pCU1 TraI Relaxase

Figure 2.7 Structural Comparison of the Relaxase N-terminus

Parameter	WT_299 (PDB 3L57)	Nterm_299 (PDB 3L6T)
Data Collection		
X-ray source	APS Beam-line 22-ID	SSRL Beam-line BL9-2 (ID)
Wavelength (Å)	0.9999	0.9795
Temperature (K)	100	100
Space Group	P2 ₁	P2 ₁
Unit Cell Parameters [(a (Å), b (Å), c (Å); α (°), β (°), γ (°))]	50.28 x 58.52 x 87.12 90.00 x 83.88 x 90.00	50.39 x 58.66 x 86.53 90.00 x 95.01 x 90.00
Resolution (Å) (highest shell)	50.00-2.29 (2.43-2.29)	86.066-1.93 (2.00-1.93)
No. of unique reflections, observed (highest shell)	22081 (3362)	36385 (2836)
Percent Completeness (highest shell)	96.60 (91.80)	95.50 (76.00)
R _{sym} ^a (highest shell)	0.054 (0.401)	0.076 (0.480)
I/σ (highest shell)	18.25 (4.50)	31.59 (3.03)
B (isotropic) from Wilson Plot	45.74	44.33
Refinement		
Resolution (Å) (highest shell)	48.49-2.29 (2.40-2.29)	31.38-1.93 (1.98-1.93)
No. of unique reflections, observed (highest shell)	22077 (2499)	36342 (1954)
Percent of reflections observed (highest shell)	96.97 (93.00)	95.5 (68.00)
R _{work} ^b (highest shell)	0.173 (0.203)	0.168 (0.235)
R _{free} ^c (highest shell)	0.226 (0.244)	0.204 (0.250)
Molecules per asymmetric unit (AU)	2	2
No. of amino acids per AU	426	429
No. of water molecules per AU	219	366
B-factors		
Protein	47.07	46.79
Waters	50.16	57.32
RMSDs		
Bond Lengths (Å)	0.003	0.006
Bond Angles (°)	0.724	0.900
Ramachandran (%) Plot		
Favored	96.6	98.1
Allowed	2.9	1.4
Outliers	0.5	0.5

Table 2.1 Crystallographic Statistics

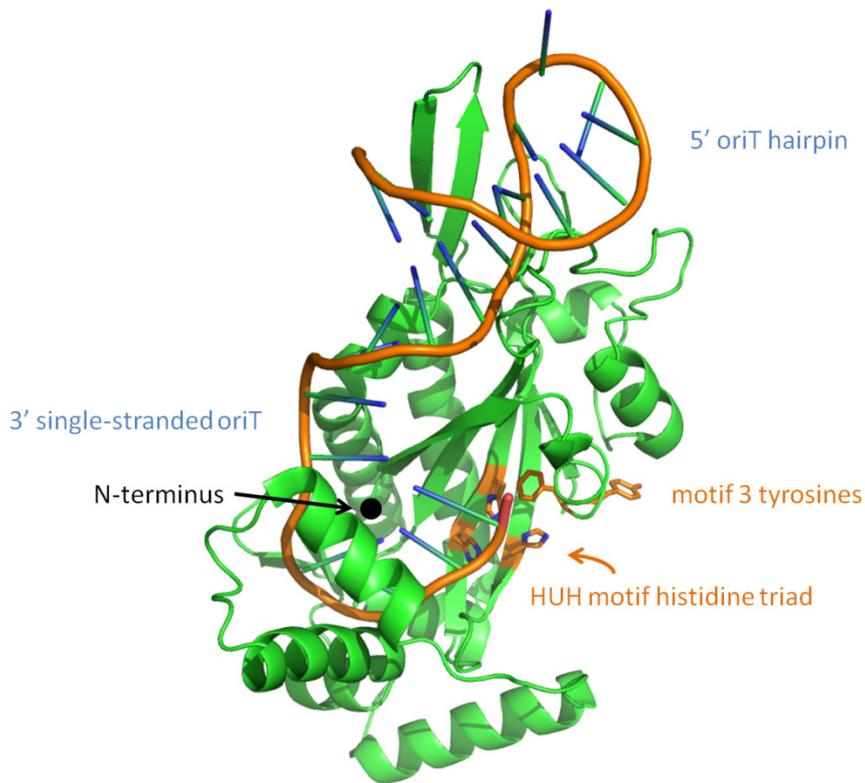


Figure 2.1 The Structure of the Conjugative Relaxase in Complex with Plasmid *oriT* DNA

The structure of the plasmid R388 TrwC relaxase was solved in complex with 27 nucleotides of the R388 *oriT* (PDB code 2CDM³⁶). The relaxase is shown in green, the DNA in orange and blue. The DNA is bound as a partial hairpin, with the 3' end of the partial hairpin extending as single-stranded DNA (ssDNA). The ssDNA wraps around the N-terminus of the relaxase (shown as a black sphere) and then enters the active site. The active site contains the HUH motif. The three histidines composing the motif are shown as orange and blue sticks. Two of the four motif 3 tyrosines of the R388 relaxase are shown as orange and red sticks.

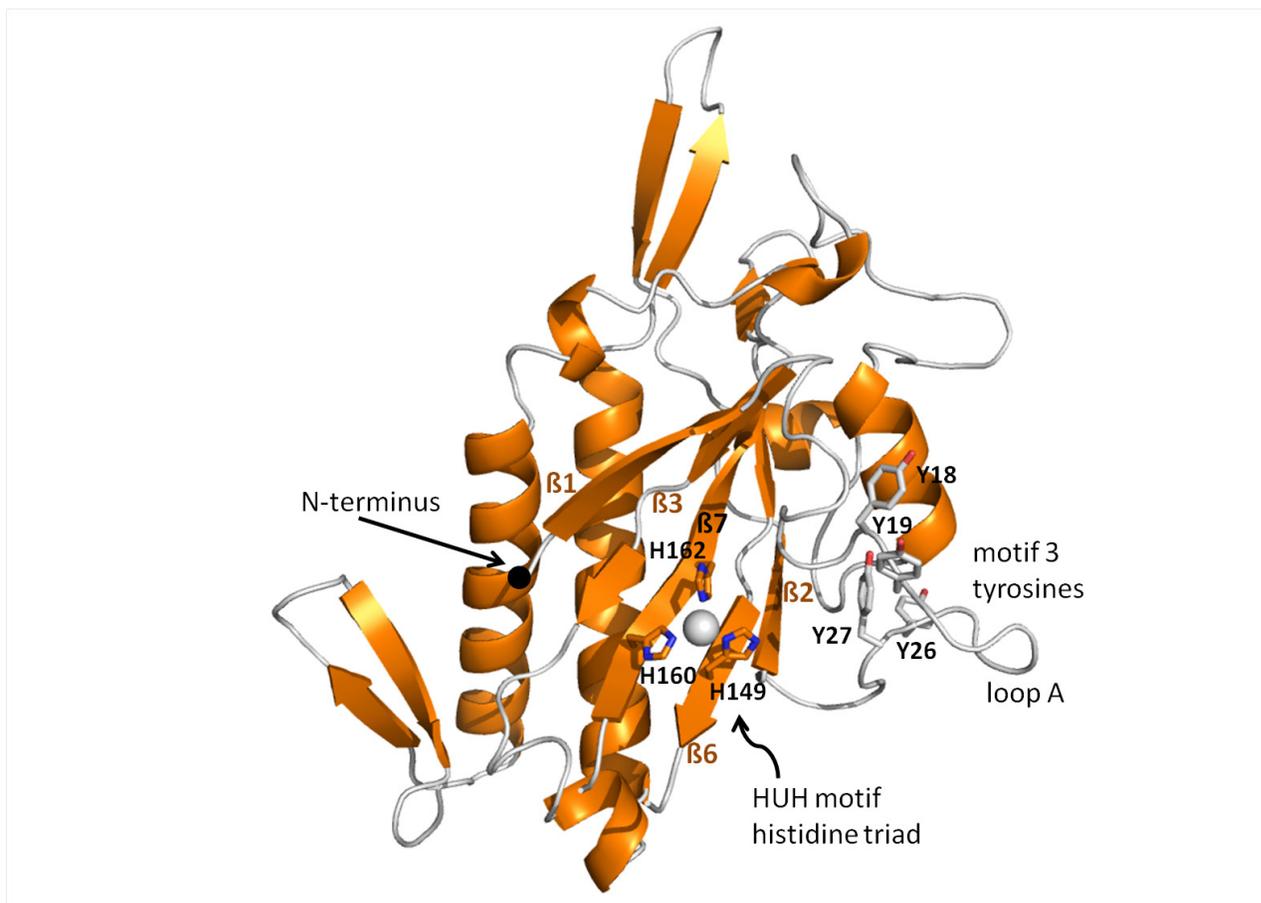


Figure 2.2 Structure of the pCU1 TraI Relaxase

The structure of the plasmid pCU1 TraI relaxase was solved to 2.3 Å resolution (PDB code 3L57). The relaxase is shown in orange (helices and sheets) and gray (loops). The N-terminus of the relaxase is shown as a black sphere. The active site contains the HUH motif. The three histidines composing the motif are shown as orange and blue sticks and are labeled according to their residue number. Each of the four motif 3 tyrosines is shown as gray and orange sticks and is labeled according to its residue number. The bound manganese ion is shown as a gray sphere. The five β -sheets composing the central palm of the relaxase are labeled and comprise β 1, β 3, β 7, β 6, and β 2 of the structure.

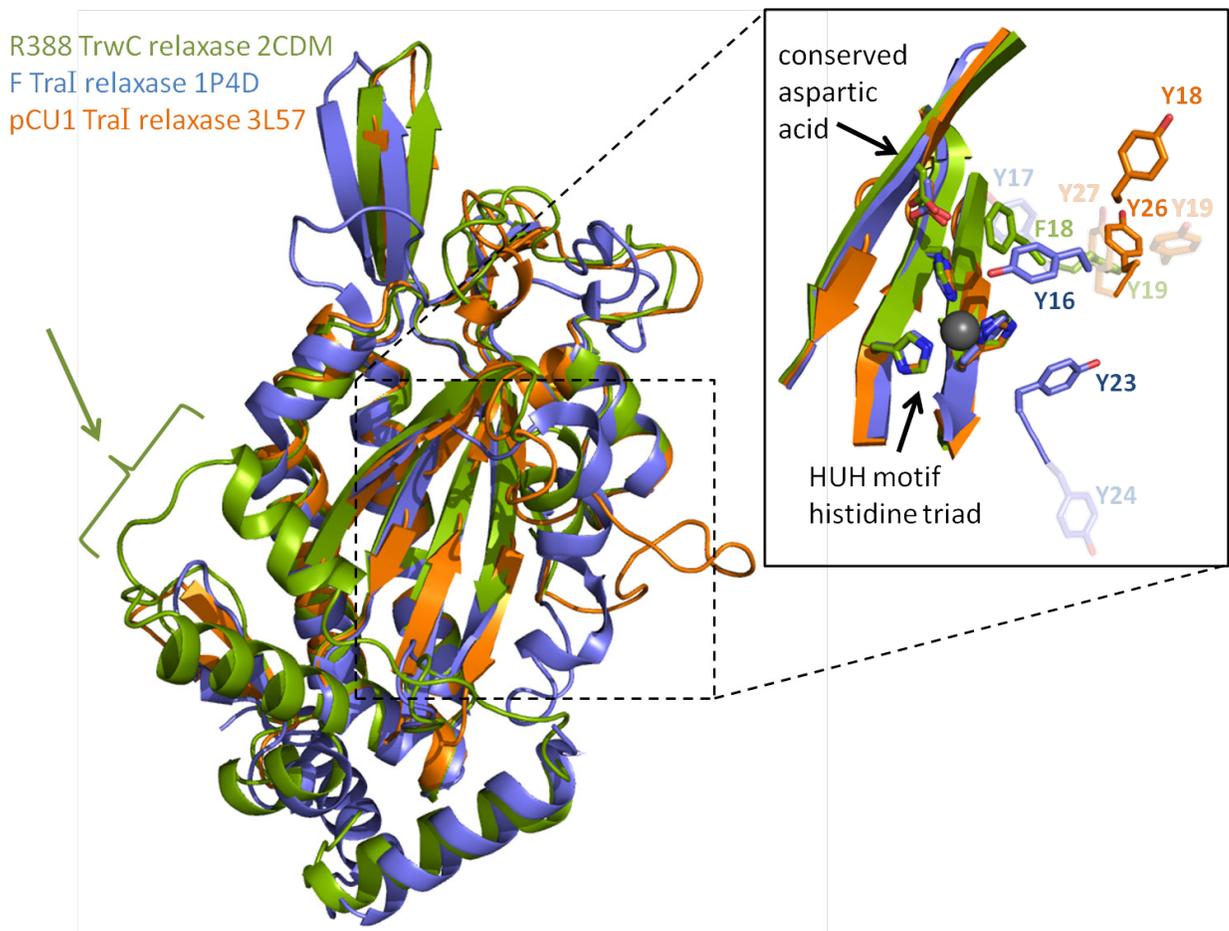


Figure 2.3 Structural Comparison of Relaxase Enzymes

The structures of the plasmid pCU1 TraI relaxase (PDB code 3L57, shown in orange), plasmid R388 TrwC relaxase (PDB code 2CDM, shown in green), and F plasmid TraI relaxase (PDB code 1P4D, shown in blue) are compared. The overall folds of the three proteins are similar. The R388 TrwC relaxase was solved in complex with DNA; therefore, the C-terminus of this protein was ordered (green arrow and bracket), while it remained disordered in the structures of the two other relaxases. The active sites and motif 3 tyrosines of each relaxase are shown in detail in the inset box. Unique aspects of each structure can be seen in the inset. The active site of each relaxase contains the HUH motif, and there is little to no variation between structures within the HUH motif. The three histidines composing the motif are shown as sticks. Each of the motif 3 tyrosines is shown as sticks and is labeled according to its residue number. The bound manganese ion from the pCU1 relaxase structure is shown as a gray sphere. As can be seen in the inset, the tyrosines of the pCU1 relaxase are flipped away from the active site. In comparison, the primary DNA nicking tyrosine of the R388 and F relaxases is directed towards the active site. A conserved aspartic acid serves to deprotonate the DNA nicking tyrosine for attack on the scissile phosphate.

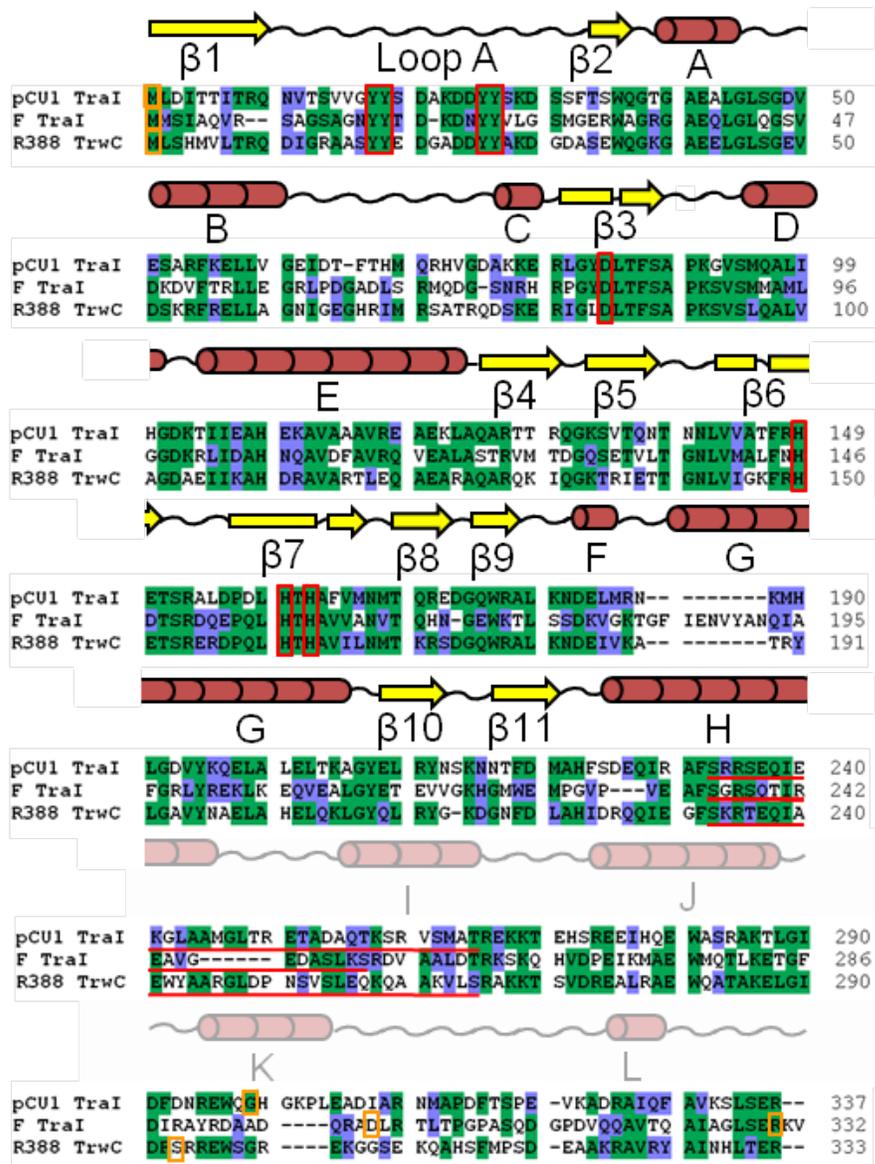


Figure 2.4 Amino Acid Sequence Alignment and Secondary Structure of the pCU1 TraI Relaxase

The initial 337 amino acids of pCU1 TraI (GenBank ID AA27542) were aligned with R388 TrwC (GenBank ID CAA44853) and F TraI (GenBank ID BAA97974) using the Clustal X program in BioEdit. Identical residues are shaded green, similar residues are shaded blue. Orange boxes indicate the first and last residues of the relaxase domains of each protein. Two constructs of the F TraI relaxase domain have been crystallized^{5, 51}, therefore the terminal residues of both are boxed. Red boxes indicate the location of the conserved catalytic tyrosine residues, aspartic acid residues, and active site HUH(+H) motifs. Red lines underline the residues forming the “DNA binding thumb” regions of each protein. Since this region is disordered in the pCU1 TraI relaxase crystal, the extent of this region is estimated. Above the three sequences is the secondary structure of the pCU1 relaxase. β -sheets are represented by yellow arrows, α -helices by maroon cylinders, and loop regions by black wavy lines. The secondary structure of residues 1-225 reflects that observed in the crystal structure of WT_299; the secondary structure of residues 226-330 reflects that predicted by Jpred 3.

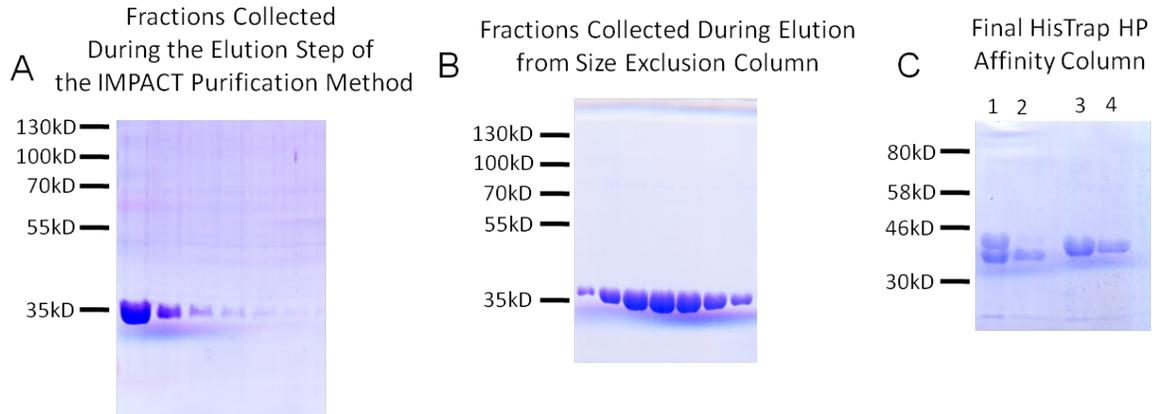


Figure 2.5 SDS-PAGE of Purified pCU1 TraI Relaxase

Molecular weight of WT_299 and Nterm_299 = 34 kD.

A) Lanes represent fractions collected during elution of WT_299 with Buffer C from chitin resin.

B) Lanes represent 2 mL fractions collected during elution of WT_299 with Buffer S from a HiLoad 16/60 Superdex 200 column (GE Healthcare).

C) Lanes represent samples of Nterm_299 prior to and following the final affinity purification step using HisTrap HP column. Lane 1 – column load (bands represent cleaved protein and Nterm_299-TEV fusion); Lane 2 – purified Nterm_299 eluting in the flow-through with Buffer A and subsequently collected; Lanes 3,4 – uncleaved Nterm_299-TEV fusion eluting in Buffer B and subsequently discarded.

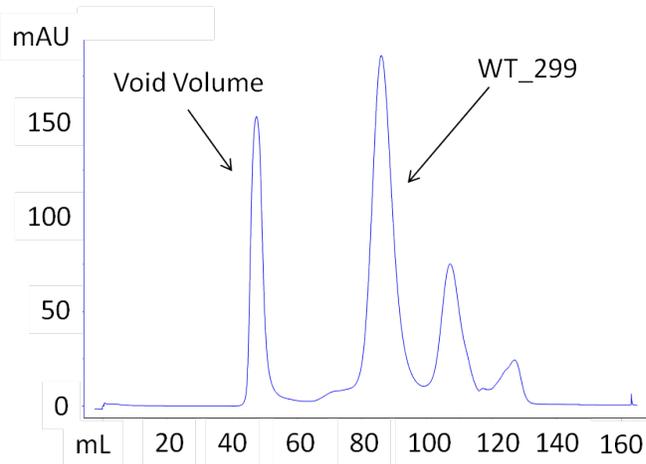


Figure 2.6 Size Exclusion Chromatography of the pCU1 TraI Relaxase

WT_299 was eluted with Buffer S from a HiLoad 16/60 Superdex 200 column (GE Healthcare) during the final purification step. MW of WT_299 is 33.9 kD; Void Volume (V_0) = 47.5 mL; Column Volume (V_T) = 118.5 mL; Elution Volume (V_e) = 86 mL. Calculated MW from elution volume = 21.4 kD where $\text{Log MW} = y = -3.1x + 6.0051$ and $x = K_{av} = (V_e - V_0)/(V_T - V_0)$.

Discrepancy in calculated and actual MW is likely due to both the high (500 mM) concentration of NaCl present in the elution buffer used during purification of WT_299 and the non-spherical shape of the relaxase protein.

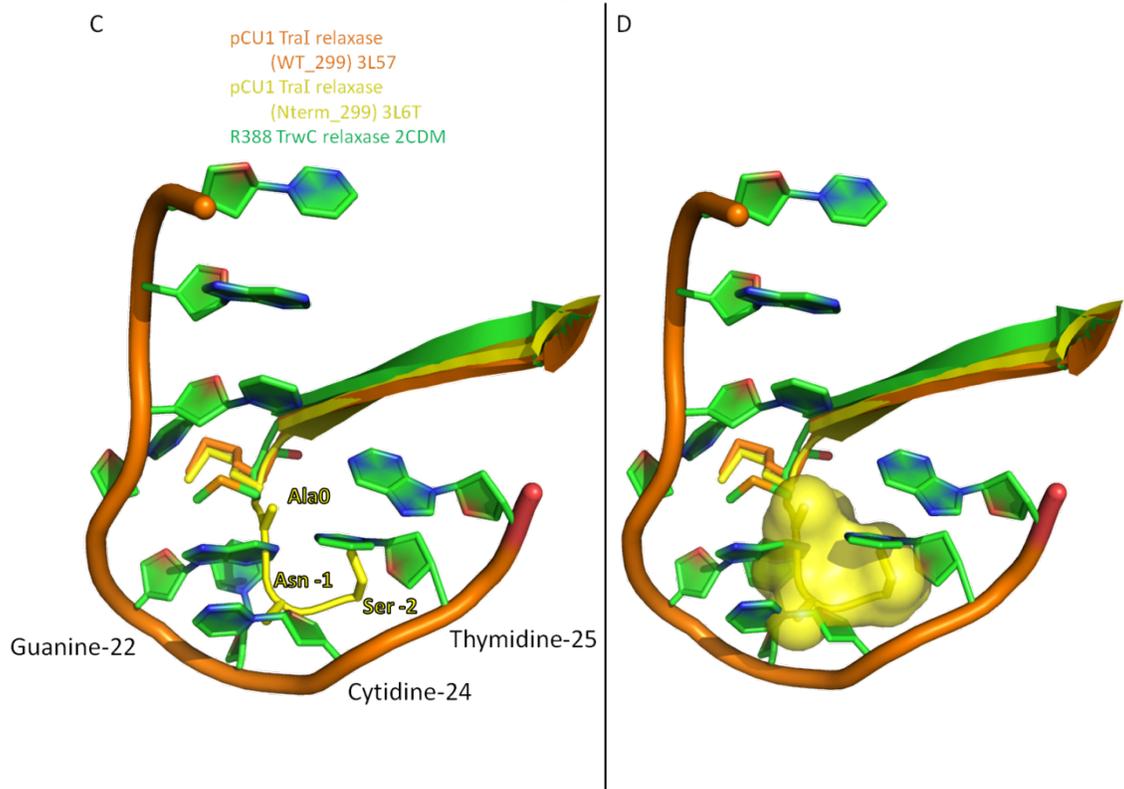
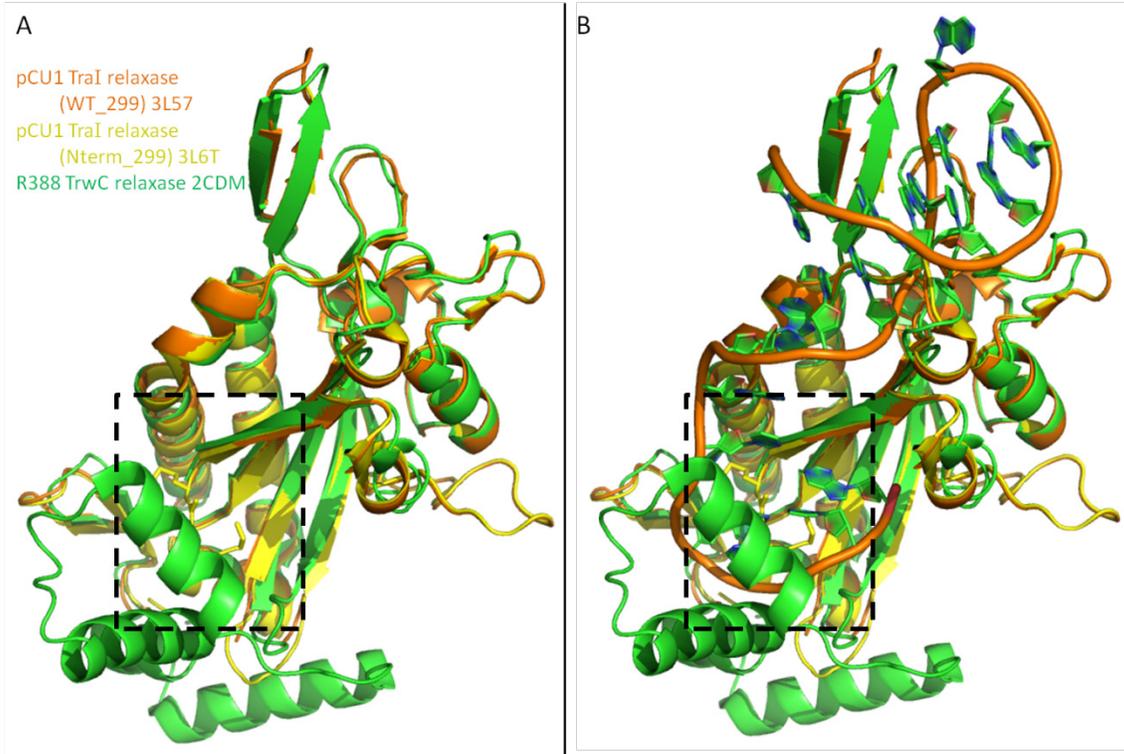


Figure 2.7 Structural Comparison of the Relaxase N-terminus

The structures of the wild type pCU1 TraI relaxase (WT_299, PDB code 3L57, shown in orange), N-terminal mutant pCU1 TraI relaxase (Nterm_299, PDB code 3L6T, shown in yellow), and plasmid R388 TrwC relaxase (PDB code 2CDM, shown in green) are compared.

A, B) The overall folds of the three proteins are similar. The R388 TrwC relaxase was solved in complex with 27nts of the R388 *oriT* DNA (DNA is seen in tiles B,C, and D with an orange sugar-phosphate backbone and green and blue bases); therefore, the C-terminus of this protein was ordered, while it remained disordered in the structures of the two other relaxases. The N-terminal bases of each relaxase and the bound *oriT* DNA of the R388 TrwC relaxase are shown in detail in C, and D.

C) The N-terminal methionines of each protein, as well as the N-terminal non-native residues Ser-2, Asn-1, and Ala0 of Nterm_299 are shown as sticks.

D) The N-terminal methionines of each protein are shown as sticks, the N-terminal non-native residues of Nterm_299 are shown in space filling mode. As can be seen in both C and D, the N-terminal 3 residues of Nterm_299 sterically clash with the bound DNA bases. In comparison, the N-terminal methionine of WT_299 and the R388 relaxase fits into the U-turn formed by the *oriT* DNA.

Chapter 3: DNA Binding and DNA Nicking Activities of the Plasmid pCU1 TraI Relaxase

3.1 Introduction to Relaxase Function

As detailed in Chapter 1, the relaxase enzyme functions within a multi-protein complex called the relaxosome to initiate and terminate conjugative plasmid transfer (CPT) (Figure 1.2). Despite the wide variety of conjugative plasmids and relaxase enzymes, all relaxases contain a conserved HUH motif and tyrosine-dominated motif 3. Therefore, it is predicted that all relaxases will incorporate a series of conserved basic mechanistic steps during initiation and termination of plasmid transfer. Structural and functional analyses of a limited number of relaxases have illustrated how these enzymes process the *oriT* and *nic* site of their respective plasmid^{2, 23, 49}. For example, the DNA-bound structure of the TrwC relaxase reveals that this relaxase binds its plasmid's *oriT* as a partial hairpin (Figure 2.1)^{36, 39}. DNA binding and nicking assays performed with the F TraI relaxase have determined that the most important protein-DNA interactions for this enzyme are formed between the relaxase and 8-10 nucleotides located immediately upstream of the *nic* site. All the relaxases characterized to date bind and nick their respective DNA substrate with high affinity and sequence specificity. In fact, DNA nicking by the F TraI relaxase can be eliminated upon mutation of a single base upstream of the *nic* site^{19, 32, 35, 39}.

We present here a summary of the DNA binding and DNA nicking activities of the pCU1 TraI relaxase. This analysis has identified four key functional and mechanistic characteristics of this relaxase, two of which are unique to the enzyme as compared to those previously studied. First, the pCU1 TraI relaxase binds all DNA substrates in a uniquely weak, non-specific manner. Second, however, the relaxase nicks DNA with high sequence specificity. Third, the relaxase can nick DNA in the presence of a variety of metals, but appears to preferentially use manganese. Finally, the relaxase preferentially nicks DNA with a unique combination of tyrosines, as compared to other

conjugative relaxases. In particular, the third tyrosine, Y26, and a combination of the first and second, Y18+Y19, are both able to nick DNA at wild type levels.

3.2 Construct Design, Mutagenesis, Expression, and Purification of the pCU1 TraI Relaxase

As described in Chapter 2, the pCU1 TraI relaxase domain is located at the N-terminus of the TraI protein, extending from residues 1 to 299. The wild type relaxase (WT_299) was cloned into the pTYB2 vector of the IMPACT system (New England Biolabs, NEB), and an N-terminal mutant relaxase construct (Nterm_299) was cloned into the vector pMCG9 using ligation independent cloning⁵⁷ (see Section 2.2 for details). Tyrosine and aspartic acid mutations were made within the WT_299 construct in the IMPACT system pTYB2 vector using QuikChange site directed mutagenesis (Stratagene). All cloning and mutagenesis was verified by sequencing at the UNC-CH Genome Analysis Facility.

Both wild type and mutant proteins cloned into the IMPACT system were expressed in *Escherichia coli* BL21(DE3), and each protein was then purified on chitin resin (NEB) (see Section 2.2 for details). A final size exclusion chromatography purification step was then performed at 4°C. After concentrating each protein to ~7 mg/mL, 5 mL of each was loaded onto a HiLoad 16/60 Superdex 200 column (GE Healthcare) pre-equilibrated with Buffer S (500 mM NaCl, 20 mM Tris-HCl pH 7.5, 5% glycerol, 5 mM EDTA, 0.01% azide) on an ATKExpress FPLC (GE Healthcare). TraI was eluted from the column in Buffer S and collected in 2 mL fractions at a 1.2 mL/min flow rate. N-terminal mutant protein (Nterm_299) was expressed in *Escherichia coli* BL21(DE3) and then purified using a combination of affinity chromatography and affinity tag cleavage by TEV protease, as detailed in Section 2.2.

The final purification step for each construct varied depending on the assays to be performed with the protein. For subsequent analysis by circular dichroism (CD) spectroscopy, protein was dialyzed into Buffer CD (100 mM NaF, 20 mM Tris-HCl pH 7.5, 5% glycerol, 0.01% azide). For DNA binding, nicking, and cross-over assays, protein was dialyzed into Buffer D (100 mM NaCl, 20 mM Tris-HCl pH 7.5, 5% glycerol, 0.01% azide) to lower the salt concentration and remove any

EDTA present. For inductively coupled plasma mass spectrometry (ICP-MS) analysis, a separate expression and purification of WT_299 and Nterm_299 was performed in order to prevent contaminating the protein sample with metals from outside sources. During this separate purification, the use of all metal-containing devices was avoided. Following purification, the samples were then dialyzed into Buffer ICP (20 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.01% azide). For all applications, samples were flash frozen in liquid nitrogen and stored in 60 μ L aliquots at -80°C ; water was obtained from the laboratory Barnstead E-pure water filtration system, at > 17 megaohm-cm (ddH₂O). All protein was separated by SDS-PAGE, stained with Coomassie Brilliant Blue, de-stained, and found to be $>95\%$ pure (Figure 2.5).

3.3 Characterization of pCU1 TraI Relaxase Stability by Circular Dichroism Spectroscopy

The secondary structure and stability of the wild type and N-terminal mutant relaxase domain of pCU1 TraI, both in the presence and absence of a saturating (6-fold molar excess) concentration of a DNA substrate, were determined by circular dichroism (CD) spectroscopy. The DNA substrate, 35oriT-hairpin, was commercially synthesized and HPLC purified (IDT) (see Table 3.1 for sequence). It consisted of the first 35 nt of the pCU1 *oriT* found upstream of the plasmid's *nic* site and was predicted to form a partial hairpin. DNA secondary structure predictions were performed using the M-fold server (<http://mfold.bioinfo.rpi.edu/cgi-bin/dna-form1.cgi>). The DNA was resuspended in Buffer R (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.05 mM EDTA, 0.01% azide), heated to 95°C for 10 min, and then allowed to slow cool to room temperature. Each experimental sample (total volume 400 μ L) contained a final protein concentration of 6 μ M; a final DNA concentration of 0 μ M, 24 μ M or 36 μ M; and 100 mM NaF, 21 mM Tris-HCl pH 7.5, 4.65-5.8 mM NaCl, and 5% glycerol. For each experimental sample, the corresponding blank sample was identical to the experimental sample, minus the protein component.

To examine the secondary structure of the pCU1 TraI relaxase, CD spectra of experimental and control samples were collected in triplicate from 190 nm to 260 nm using a Circular Dichroism Spectrometer, Model 62DS (Aviv), with a 3 s averaging time, while the temperature was held at

25°C. Due to significant noise at lower wavelengths, only data between 200 nm and 260 nm were processed using Excel 2007 (Microsoft, 2006) and SigmaPlot 8.02a (Systat, 2004). Data were initially collected as ellipticity (mdeg) vs. wavelength (nm), and were then converted to MRE (mdeg*cm²/dmol) vs. wavelength (nm) using Equation 3.1:

$$\text{Equation 3.1: } \text{MRE} = \frac{E * MW * 0.1}{C * l * \#}$$

where: *E*, ellipticity; *MW*, molecular weight of protein construct (g/mol); *C*, concentration of protein construct in g/cm³; *l*, path length in cm; and #, number of residues in construct. The same formula was then applied to the error associated with each data point. CD spectra were then plotted as mean residue ellipticity (MRE, mdeg*cm²/dmol) vs. wavelength (nm). Each data point is the average of three background subtracted (experimental sample signal – control sample signal) MRE values, and each error bar represents the standard error of the variance of three background subtracted MRE data points.

The resulting spectra were indicative of well folded proteins, with minima at 208 nm and 220 nm for both WT_299 and Nterm_299, though the average signal of the Nterm_299 spectrum was weaker than that of WT_299 (Figure 3.1). When comparing CD spectra of WT_299 collected before and after addition of 35oriT-hairpin, the observed minimum at 220 nm increased in intensity, while the observed minimum at 208 nm weakened in intensity. These shifts may reflect the ordering of the alpha-helical “thumb” region of the protein, which is only seen fully ordered in structures of the relaxase solved in complex with DNA (Figure 2.1).

The sensitivity of the pCU1 TraI relaxase to thermal denaturation was determined by following the change in ellipticity of each experimental sample at 222 nm as the temperature of the sample was increased from 10°C to 75°C. A 12 s averaging time, 1 degree step, and 2°C/min rate of change were used. The resulting data were plotted as fraction unfolded vs. temperature (°C) and fit using nonlinear regression assuming a two state model (folded monomer to unfolded monomer). The melting temperature (*T_m*) was calculated as described in Greenfield, et al. (2006)⁶⁴ (Table 3.2). The

error reported is the standard error as calculated when fitting the data. Upon the addition of the DNA substrate 35oriT-hairpin, the estimated melting temperature (T_m) of WT_299 and Nterm_299 shifted from 42.7°C to 61°C and from 44.6°C to 54.6° (Table 3.2), respectively, and the transition between folded and unfolded states became significantly sharper. This shift reflected an increased stability of the protein when bound to DNA.

3.4 Characterization of DNA Binding by the pCU1 TraI Relaxase

3.4.1 Introduction to DNA Binding by the pCU1 TraI Relaxase

The relaxase binds a portion of its plasmid's T-strand *oriT* prior to cleaving the T-strand *nic* site^{32, 35, 39}. We determined the binding affinity and specificity of the pCU1 TraI relaxase for a panel of DNA substrates (Table 3.1). The effect of DNA substrate length, secondary structure, and sequence on relaxase binding affinity was investigated. In contrast to the high affinity, sequence-specific DNA binding by the F TraI and R388 TrwC relaxase enzymes^{32, 35, 39}, the pCU1 TraI relaxase bound DNA in a nonspecific manner and with low affinity (Figure 3.2, Table 3.1).

3.4.2 Methodology

The affinity of WT_299 and Nterm_299 for a panel of DNA substrates (Table 3.1) was measured by fluorescence anisotropy-based DNA binding assays. 5' fluorescein (FAM)-labeled DNA substrates were commercially synthesized and HPLC purified (IDT). Each substrate was resuspended in Buffer R (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.05 mM EDTA, 0.01% azide) heated to 95°C for 10 min, and then allowed to slow cool to room temperature. Each stock was then diluted to 0.1 μ M in Buffer R prior to use. For each assay, purified protein was diluted to 2X assay concentration in a final buffer of 50 mM NaCl, 20 mM Tris-HCl pH 7.5, 5% glycerol, 5 mM EDTA 0.01% azide (Buffer B) and then serially diluted, in Buffer B, in a 96- or 384-well black assay plate (Costar) to generate 8 or more unique concentrations of protein, in triplicate. An equal volume of 0.1 μ M fluorescently-labeled DNA substrate was then mixed with each concentration of protein, resulting in a final concentration of 50 nM DNA and 1X protein and a total volume of 50 μ L. To generate a no-protein control, equal volumes of 0.1 μ M fluorescently-labeled DNA substrate and Buffer B were

mixed in three wells. The final buffer concentration during each assay was: 50 mM NaCl, 15 mM Tris-HCl pH 7.5, 2.5 mM EDTA, 2.5% glycerol. To determine the role of metal during DNA binding, the final assay buffer contained either 2.5 mM EDTA or 4 mM metal.

To detect changes in DNA binding as a function of protein concentration, the fluorescence anisotropy (FA) of the fluorescently-labeled substrate at 520 nm was monitored following excitation at 485 nm, using a PHERAstar fluorescence plate reader (BMG Labtech) in a T format, as the protein concentration increased. To generate binding curves and calculate the dissociation constant (K_D) for each experiment, normalized data were plotted as average FA vs. total protein concentration and fit to Equation 3.2:

$$\text{Equation 3.2 } f = \min + (\max - \min) \frac{\left\{ (T+x+K) - [(-T-x-K)^2 - 4Tx]^{\frac{1}{2}} \right\}}{2T}$$

using nonlinear regression in Graphpad PRISM v5.03 (Graphpad, 2010), where f , average FA signal detected; T , total DNA conc. (set to 50 nM); x , total protein concentration; K , K_D ; \min , average FA signal of no protein control; and \max , average FA signal of sample at a saturating concentration of protein. Each data point is the average of at least three replicates, with error bars representing the standard error of these replicates. For each substrate, the binding affinity and associated standard error, as calculated by Equation 3.2, are provided in Figure 3.2 and Table 3.1.

3.4.3 pCU1 TraI Relaxase DNA Binding Affinity and Specificity

By calculating the affinity of pCU1 relaxase for a panel of DNA substrate (Table 3.1), we determined that both WT_299 and Nterm_299 bound DNA substrates in an *oriT*- and hairpin-independent manner. The binding affinities of both WT_299 and Nterm_299 were similar for all substrates investigated, and varied between 300 nM and 1.2 μ M (Table 3.1, Figure 3.2, Nterm_299 data not shown).

This low affinity, sequence-independent DNA binding was unexpected, given the highly sequence-specific and high affinity DNA binding reported for similar relaxase enzymes^{19, 32, 35, 39}. However, as discussed in Section 3.3, both WT_299 and Nterm_299 appeared well folded by CD

spectroscopy (Figure 3.1, Table 3.2), and the design of WT_299 and Nterm_299 incorporated the full extent of the predicted relaxase domain of pCU1 TraI (Figure 2.4). To determine, however, if a longer protein construct would bind DNA with higher affinity or greater sequence specificity, the pCU1 TraI construct WT_411 was cloned, expressed, and purified, as described for WT_299. WT_411 consists of the first 411 residues of pCU1 TraI; its C-terminus was chosen based upon sequence alignments and secondary structure predictions. However, WT_411 bound DNA with an affinity and lack of sequence selectivity similar to that of WT_299.

Therefore, the design of the DNA binding assay was next modified to determine if the experimental conditions could be further optimized for DNA binding. First, the design of the DNA substrate was modified to include the *oriT* sequence downstream of the *nic* site (35/7oriT-hairpin), to determine if this sequence was important for pCU1 TraI relaxase binding. However, no improvement in DNA binding affinity or selectivity was observed when determining the binding affinity of Y18,19,26,27F_299 for this substrate (Figure 3.2a). Second, binding experiments were supplemented with a final concentration of 4 mM MnCl₂. This change yielded a modest 2-3 fold increase in DNA binding affinity (Figure 3.2b). Third, DNA binding was measured over a range of temperatures (22 to 37°C), but this temperature change had no effect on binding affinity. Finally, the total concentration of chloride ions was limited to 45 mM, taking into account contributions from the buffer components Tris-HCl and MnCl₂, but still no enhancement in binding affinity was observed.

To determine if the best equation had been used to fit the experimental data, data were also fit by two additional equations using the Graphpad PRISM software (two-sites specific binding, and one-site specific binding with Hill slope). The data were ambiguously fit when assuming two binding sites. When fitting the data with the Hill parameter, the calculated K_D did not improve, while the goodness-of-fit parameters (dependency and R²) showed slight or no improvement. These results did not justify the use of an additional parameter.

Therefore, these data indicate that WT_299 alone binds DNA in a weak, sequence-independent manner. We propose that the pCU1 TraI relaxase relies upon a second DNA-binding

protein to bind selectively to the pCU1 *oriT*. In fact, a pCU1-encoded protein, TraK, has been identified, for which the efficiency of pCU1 plasmid transfer decreases 100-fold upon deletion⁴⁶. Homologues of TraK have been identified within the F, R388 and RP4 plasmid sequences; all are known to be DNA-binding proteins that enhance relaxase function^{23, 50, 65, 66}. In particular, studies of RP4 have shown that its TraK homologue, TraJ, enhances sequence-specific binding and nicking by the RP4 TraI enzyme, though it is not required for TraI function^{23, 50}. During transfer of RP4, TraJ binds first to the proximal arm of an inverted repeat upstream of the *nic* site. It then aligns TraI for the initial nicking reaction⁹. In a similar manner, the F plasmid-encoded TraK homologue, TraY, promotes relaxase binding and nicking *in vivo*²³. Since the pCU1 TraI relaxase appears incapable of sequence-specific DNA binding when alone, it is possible that the role of pCU1 TraK is broader than that of its homologs. Indeed, pCU1 TraK may not merely enhance the function of pCU1 TraI, but may be required for pCU1 TraI DNA binding. If this is true, association of pCU1 TraI with the pCU1 *oriT* could be accomplished by a two-component system. TraK could first bind selectively to the pCU1 *oriT* to then promote sequence specific binding and nicking by TraI.

3.5 Characterization of DNA Nicking by the pCU1 TraI Relaxase

3.5.1 Introduction to DNA Nicking by the pCU1 TraI Relaxase

After binding its respective plasmid's *oriT*, each relaxase must cleave its plasmid's *nic* site in a sequence-specific manner in order to initiate T-strand transfer. Three conditions must be met for this DNA nicking reaction to occur efficiently. First, the DNA substrate must be properly positioned within the relaxase active site such that the scissile phosphate of the *nic* site can be attacked by the DNA nicking tyrosines^{5, 36, 39}. Second, the scissile phosphate must act as an electrophile. The negatively charged scissile phosphate is positioned within the relaxase active site adjacent to a divalent cation, which is itself bound by the conserved relaxase HUH motif, and a third coordinating residue, often another histidine. As a result of this arrangement, the negative charge on the phosphate is neutralized by the positively charged metal center^{12, 13, 15, 36, 39}. Finally, the tyrosine attacking the scissile phosphate must act as a nucleophile (Figure 1.3). The pK_a of tyrosine is 10; thus at

physiologic pH, the tyrosine hydroxyl is protonated. To deprotonate the hydroxyl, a neighboring residue must act as a base to extract the proton and activate the tyrosine, or the local pH of the active site must be higher than that of physiological pH (7 – 7.5).

In order to meet condition one, the *oriT* DNA bound by the TrwC and F relaxases forms a U-turn around the protein N-terminus prior to entering the active site. The protein-DNA and DNA-DNA contacts formed by the DNA U-turn, and the DNA sequence lying between the U-turn and *nic* site, are essential for properly aligning the scissile phosphate in the active site of these two relaxases^{5, 36, 39}. Analysis of the HUH motif of the F TraI has confirmed that this relaxase satisfies condition two by coordinating a divalent cation using the HUH motif histidines; upon mutation of these histidines, metal binding and plasmid transfer are eliminated¹³. Structures of the F plasmid and plasmid R388 relaxases reveal that while four tyrosines surround their active sites, the first in amino acid sequence is both properly aligned for attack on the DNA substrate's scissile phosphate and is most likely activated for attack by an adjacent aspartic acid, thus satisfying condition three (Figure 2.3)^{5, 26, 36, 39, 51}. In the case of the related R1162 MobA relaxase, which contains only one tyrosine near the active site, structural data again reveal that this tyrosine is directed towards the HUH motif of the active site, though no active site base has been identified¹⁶. The role of these tyrosines has been verified by functional assays in which Y16 of the F relaxase, Y18 of the R388 relaxase, and Y25 of the R1162 relaxase were shown to be the primary DNA nicking residues of each enzyme^{28, 33, 38, 67}. In the case of R388, the third tyrosine in primary sequence, Y26, also nicked DNA efficiently if the relaxase was presented with substrates forming a hairpin. Hairpin-forming substrates are predicted to mimic the secondary structure of the *oriT* during termination of plasmid transfer^{37, 38}.

Here we characterize protein and DNA features essential for DNA nicking by the pCU1 TraI relaxase. In particular, we identify the tyrosine and aspartic acid residues required for DNA nicking, the portion of the pCU1 *oriT* required for sequence-specific DNA nicking, and the metal most likely to support nicking by the pCU1 relaxase. The pCU1 TraI relaxase contains four possible catalytic tyrosines, as determined by sequence alignment with homologous enzymes (Figure 2.4). However,

as seen in Figure 2.3, the tyrosines of the pCU1 relaxase are located on a loop at the periphery of the active site, in an orientation unique from that of other relaxases. For any one of these tyrosines to attack the scissile phosphate of a bound DNA substrate, the loop on which the residues are located must undergo a conformational change in order to direct the nicking tyrosine towards the scissile phosphate. Therefore, it is unclear from the present structural data which tyrosine would then be oriented for attack. To determine the tyrosine(s) responsible for DNA nicking, we performed a series of DNA nicking experiments involving tyrosine mutants of the pCU1 TraI relaxase. We then complemented these data with DNA nicking experiments involving aspartic acid mutants to demonstrate that interactions between the four tyrosines and this aspartic acid determine the nicking activity of the pCU1 relaxase.

In addition to identifying the residues responsible for pCU1 TraI relaxase-mediated DNA nicking, we also investigated the sequence of the pCU1 *oriT* to determine which bases affected DNA nicking by the pCU1 TraI relaxase. Since DNA binding assays indicated WT_299 bound DNA with no sequence specificity, we were particularly interested to see if WT_299 also nicked and ligated DNA in a sequence-independent manner. The DNA nicking and ligation activities of WT_299 were tested using a series of DNA nicking and cross-over assays. Finally, we determined the importance of metals for DNA nicking by measuring the magnitude of DNA nicking by the pCU1 TraI relaxase in the presence of increasing concentrations of several different metal cations. In summary, we present in this section the DNA nicking and ligating activities of the relaxase of the resistance plasmid pCU1 and highlight the unique aspects of the mechanism utilized by this relaxase relative to those previously characterized.

3.5.2 Methodology

The DNA nicking and ligating activities of the pCU1 TraI relaxase for a panel of DNA substrates were determined by DNA nicking and DNA cross-over assays. 5' fluorescein-labeled ("FAM"), 3' TAMRA-labeled ("TAM") and unlabeled DNA substrates were commercially synthesized (IDT); all labeled DNA substrates were HPLC purified. Each substrate was resuspended

in Buffer R, heated to 95°C for 10 min, and then allowed to cool passively to room temperature. Each stock was then diluted to 10 μM in Buffer R prior to use. Table 3.3 provides the nucleic acid sequence of each substrate investigated. Where applicable, the substrate *nic* site is indicated by a vertical double line and bases forming an inverted repeat are boxed in black. DNA secondary structure predictions were performed using the M-fold server (<http://mfold.bioinfo.rpi.edu/cgi-bin/dna-form1.cgi>). A schematic of each DNA nicking and cross-over experiment is provided in Figure 3.3.

In general, each 10 μL assay contained 5 μM purified TraI relaxase, 1 μM 5' fluorescein-labeled DNA substrate, 4 mM MnCl₂, 50 mM NaCl, 18 mM Tris-HCl pH 7.5, 4.5% glycerol and, for cross-over assays, 1 μM unlabeled DNA substrate or 1 μM 3' TAMRA-labeled DNA substrate. The reaction was initiated upon addition of the enzyme. Each reaction was incubated at 37°C for 1 h, quenched by the addition of 10 μL 2X quenching solution (0.01% xylene cyanol, 85% formamide, 20 mM EDTA, 2X TAE, 0.2% SDS), and then subjected to electrophoresis over a denaturing 16% polyacrylamide gel (40 mL 16% acrylamide gel stock (8M urea, 16% polyacrylamide/bisacrylamide, 1X TBE), 400 μL 10% ammonium persulfate (APS), 40 μL tetramethylethylenediamine (TEMED)) in 1X TBE running buffer. Gels were visualized using a VersaDoc Imaging System, 4400 MP (BioRad) and the accompanying Quantity One software (BioRad). Band intensities were quantified using ImageQuant 5.2 (Molecular Dynamics 1999) and ImageJ 1.42 (Rasband, W.S., NIH 2008). Prior to quantification, standard background subtraction was performed on all gels.

DNA nicking activity was reported as percent of substrate nicked, where activity equals intensity of the product band divided by the sum of product and substrates band intensities, all multiplied by 100%. DNA cross-over activity was reported as percent product generated. If only one of the two cross-over substrates were fluorescently-labeled, the percent product was calculated as intensity of the product band divided by the sum of product and substrates band intensities, all multiplied by 100%. If both cross-over substrates were fluorescently-labeled, the percent product was calculated as intensity of the doubly-labeled (FAM+TAM) product divided by the sum of the FAM-

labeled substrate band, the FAM-labeled nicked product band and the doubly-labeled product band, all multiplied by 100%. All data processing was performed in Excel 2007 and plots were generated in SigmaPlot 8.02a (Systat, 2004) or Graphpad PRISM v5.03 (Graphpad, 2010).

DNA covalent attachment assays were prepared as described for the DNA nicking assays, with the following modifications: in the place of 1 μM 5' fluorescein-labeled DNA substrate, 1 μM 3' TAMRA-labeled DNA substrate was used. Assays were subjected to electrophoresis over a 16% denaturing bis-tris polyacrylamide gel (resolving gel: 1X bis-Tris gel buffer (0.41 M bis-Tris pH 6.5-6.8), 16% acrylamide, 0.1% APS, 0.1% TEMED; stacking gel: 1X bis-Tris gel buffer, 4% acrylamide, 0.1% APS, 0.1% TEMED) using a high molecular weight running buffer (50 mM MOPS, 50 mM Tris base, 1 mM EDTA, 0.1% SDS, pH 7.7, Boston Bioproducts). DNA nicking and resultant covalent attachment was reported as "Percent Substrate Nicked" and represents the intensity of the product band (fluorescently labeled TraI protein) divided by the sum of product and substrates band intensities, all multiplied by 100%.

To test the metal ion dependence of the pCU1 TraI relaxase, WT_299 (10 μM) was first treated with 10 mM EDTA for 24 h to chelate all metal ions initially present in the sample. DNA nicking by WT_299 was eliminated in the presence of 10 mM EDTA (Figure 3.4). The protein was then diluted either 2-fold or 10-fold upon the addition of 1 μM 35/7oriT-hairpin and 0-15 mM metal, to generate an excess of 0-10 mM metal ion. 10X stocks of all metals investigated were made by dissolving the corresponding metal ion salt (MgCl_2 , CaCl_2 , MnCl_2 , $\text{Ni}(\text{SO}_4)$, ZnCl_2 , $\text{Zn}(\text{acetate})_2$, and CuCl_2) in 1 M Tris-HCl pH 7.5.

WT_299 and Nterm_299 were analyzed by ICP-MS to determine the identity and concentration of metals present in each sample. Data were collected by the UNC-CH Chemistry Department Mass Spectrometry Core Facility on a Varian 820-MS (Palo Alto, CA). Protein was first digested in 70% HNO_3 . Each sample was then diluted in 10 mL of 1.4% HNO_3 . A control consisting of each buffer used during protein purification was analyzed to determine the concentration of metal ion in the buffers used. Each experimental sample, as well as the buffer control, was monitored for

^{24}Mg , ^{44}Ca , ^{55}Mn , ^{57}Fe , ^{60}Ni , and ^{66}Zn . Data were collected as part per billion (ppb) and converted to μM . The original concentration of each metal ion in each sample was then determined in Excel 2007 (Microsoft, 2006) and was normalized for protein concentration. Data are presented as fold increase in metal concentration over background levels, where background values represent the signal due to 1.4% HNO_3 alone.

3.5.3 Metal-Dependent DNA Nicking by the pCU1 TraI Relaxase

Both the F TraI and R388 TrwC relaxase enzymes exhibit metal-dependent DNA nicking, though the optimal metal for each is under debate^{5, 13, 17, 36, 51}. To determine the metal ion(s) most likely to be used by WT_299 during DNA nicking, the activity of the enzyme was measured in the presence of EDTA and several metal ions (Mg^{2+} , Mn^{2+} , Ca^{2+} , Ni^{2+} , Cu^{2+} , and Zn^{2+}) (Figure 3.4). In the presence of EDTA, the DNA nicking activity of WT_299 was inhibited, but this activity could be rescued by supplementing the reaction with an increasing concentration of metal ion. Mg^{2+} provided the most significant enhancement in nicking activity, as compared to the EDTA-treated sample, but Mg^{2+} , Mn^{2+} , Ca^{2+} , and Ni^{2+} all supported DNA nicking in a concentration dependent manner. As was also observed by Boer et al.³⁶, the relaxase maintained residual DNA nicking activity in the presence of EDTA. Despite controlling the pH of solutions of CuCl_2 , ZnCl_2 and $\text{Zn}(\text{acetate})_2$, relaxase activity was inhibited as the concentration of these metals increased, and relaxase activity was limited even at low concentrations of both Zn-containing compounds. We were unable to determine relaxase nicking activity in the presence of iron due to the metal's poor solubility within the pH range of the assay.

Using inductively coupled plasma mass spectrometry (ICP-MS), we then determined the concentrations of ^{24}Mg , ^{44}Ca , ^{55}Mn , ^{57}Fe , ^{60}Ni , and ^{66}Zn in samples of WT_299 (Table 3.4). When ranked according to fold increase over background levels, ^{66}Zn was by far the most prevalent metal in the WT_299 sample, followed by ^{60}Ni and ^{55}Mn . Based on these experimental data, as well as the low concentration of Ni^{2+} within the bacterial cytoplasm^{68, 69}, and the inability of Ca^{2+} and Mg^{2+} to satisfy the electron density within the active site of the crystal structure of WT_299, we concluded

that the pCU1 TraI relaxase most likely employed Mn^{2+} during DNA nicking and ligation reactions (Table 3.4).

Our results mirror those reported by Xia et al., who found that the DNA nicking activity of the R1162 MobA relaxase was supported by Mn^{2+} but inhibited by Zn^{2+} . In addition, Mn^{2+} was shown to stabilize MobA and was bound by the active site of the crystallized MobA relaxase domain^{16, 17}. In a similar manner, the MobM relaxase of plasmid pMV158 can cleave its *oriT* in the presence of Mg^{2+} and Mn^{2+} , but it tightly binds Mn^{2+} and is only stabilized against thermal denaturation by Mn^{2+} ⁷⁰. Larkin et al.⁵ demonstrated by isothermal titration calorimetry (ITC) that the affinity of F TraI relaxase was approximately 25 times stronger for Mn^{2+} than for Ca^{2+} and Mg^{2+} , and while the relaxase domain of R388 TrwC has been crystallized in complex with Zn^{2+} , the DNA nicking activity of this enzyme was also inhibited at high Zn^{2+} concentrations³⁶.

A comparison of the chemical properties of these metals may provide insight into the inhibitory nature of Cu^{2+} and Zn^{2+} . While Mg^{2+} , Mn^{2+} , Ca^{2+} , and Ni^{2+} are usually surrounded by five to six ligands in an trigonal bipyramidal or octahedral arrangement, Zn^{2+} and Cu^{2+} prefer three or four coordinating ligands⁶⁹. Within the relaxase active site, an octahedral or trigonal bipyramidal arrangement of the three histidine side chains, the bound DNA ligand, and the attacking tyrosine hydroxyl about the metal ion could ideally position the DNA's scissile phosphate for attack by the catalytic tyrosine. Lujan et al. and Datta et al. both observed an octahedral arrangement of ligands surrounding the bound Mg^{2+} in the active site of the F TraI relaxase^{26, 51}, and Lujan and colleagues noted that this octahedral arrangement would allow for the "flip-flop" mechanism proposed for the relaxase⁵¹. In summary, while the pCU1 TraI relaxase can utilize a variety of metal catalysts, Mn^{2+} appears to be the most likely candidate to occupy the active site during productive DNA nicking due to its availability and preferred coordination number.

3.5.4 Tyrosine-Dependent DNA Nicking by the pCU1 TraI Relaxase

The DNA nicking activity of the wild type pCU1 TraI relaxase domain (WT_299) was compared to that of phenylalanine mutants of the four potential DNA nicking tyrosines (Y18, Y19,

Y26, and Y27). DNA nicking was measured in the presence of two substrates, FAM-35/7oriT-hairpin and FAM-20/7oriT-half_hairpin (Table 3.3). FAM-35/7oriT-hairpin is a pCU1 *oriT*-encoding DNA substrate that includes the intact hairpin upstream of the *nic* site, and FAM-20/7oriT-half_hairpin is a pCU1 *oriT*-encoding DNA substrate that incorporates only the proximal arm of the hairpin sequence.

As shown in Figure 3.5, DNA nicking was inhibited by mutation of all four tyrosines to phenylalanines. Residual activity exhibited by the quadruple mutant in Figure 3.5 may be due to misincorporation of a tyrosine for a phenylalanine at one of the four positions in a small percentage of the protein expressed. Triple mutants that preserved only Y19 or Y27 were also incapable of nicking both substrates. Further, single mutants Y19F_299 or Y27F_299 nicked DNA at wild type levels. Unexpectedly, only limited DNA nicking occurred in the presence of Y18 alone, while, in contrast, nicking occurred at nearly wild type levels in the presence of Y26 alone. However, DNA nicking by single mutants Y18F_299 and Y26F_299 were both inhibited relative to wild type. These data indicate that as single residues, neither Y19 nor Y27 is sufficient, or required, for DNA nicking by pCU1 TraI relaxase, while both Y18 and Y26 are capable of nicking DNA. Therefore, Y19 and Y27 will be termed “partner tyrosines” and Y18 and Y26 will be termed “primary tyrosines”.

After establishing the identity of the primary and partner tyrosines, we first considered the impact the two partner tyrosines might have on the two primary tyrosines during DNA nicking by the relaxase. First, we examined the potential for interaction between pairs of adjacent primary and partner tyrosines (Y18 + Y19 or Y26 + Y27). As seen in Figure 3.5, the activity of Y18 was enhanced to that of wild type by the presence of its partner tyrosine Y19, but Y27 had no statistically significant impact on DNA nicking by the primary tyrosine Y26. Therefore, efficient DNA nicking by the relaxase was achieved by either the primary tyrosine Y26 alone or by the combination of primary and partner tyrosines Y18 + Y19. Second, we examined the potential for interaction between pairs of non-adjacent primary and partner tyrosines (Y18 + Y27 or Y26 + Y19), where Y27 and Y19 will be termed the “alternate partners” of Y18 and Y26, respectively. Both primary tyrosines (Y18

and Y26) actually suffered a reduction in nicking activity when their alternate partner was also present; for example, the activity of Y18 + Y19 + Y27 was reduced relative to Y18 + Y19; similarly, Y19 + Y26 + Y27 activity was reduced relative to Y26 + 27 (Figure 3.5). Therefore, alternate partner tyrosines had a dominant negative effect on the pCU1 TraI relaxase catalytic function. We also considered the role of substrate length on the relative activities of each mutant relaxase construct. A detailed analysis of the effect of substrate length on relaxase activity is provided in Section 3.5.7.

Taken together, these data support four conclusions concerning the DNA nicking activity of the pCU1 TraI relaxase. First, Y26 alone nicks DNA at wild type levels. Second, Y18 requires the presence of its partner Y19 to nick a statistically significant percent of the two DNA substrates investigated. Third, DNA nicking by both Y18 and Y26 is affected by the presence of their alternate partner, Y27 or Y19, respectively. Fourth, the combination of Y19 + Y27 exhibits limited DNA nicking in a substrate-dependent manner. These findings contrast those describing previously characterized conjugative relaxases^{28, 33, 37, 38}.

A complementary assay was performed to validate these results. When nicking DNA, the relaxase becomes covalently attached to the downstream portion of the DNA substrate via a phosphotyrosine bond (Figure 1.3). Therefore, WT_299 and each tyrosine to phenylalanine mutant were analyzed by an experiment capable of detecting the covalent attachment of a fluorescent 3' TAMRA-labeled DNA substrate to one or more of the active site tyrosines of pCU1 TraI (Figure 3.3C). The ability of each tyrosine to become covalently attached to the labeled DNA substrate (Table 3.3) was determined by comparing the percent substrate nicked by wild type enzymes to that of mutant enzymes. As seen in Figure 3.6, the activity of each pCU1 TraI relaxase construct, as determined by this assay, correlated well with the results in Figure 3.5. Three relaxase constructs (Y18F_299, Y26F_299, Y18,26F_299) exhibited a slightly higher activity when nicking the substrate 20/7oriT-half_hairpin-TAM, as compared to the substrate FAM-20/7oriT-half_hairpin.

3.5.5 Aspartic Acid-Dependent DNA Nicking by the pCU1 TraI Relaxase

Multi-tyrosine relaxases homologous to the pCU1 TraI relaxase primarily utilize the first tyrosine in amino acid sequence to nick DNA^{28, 33, 37, 38}. Therefore, we sought to understand the reason for the enhanced activity of Y26 alone relative to Y18 alone, as well as the role of the partner tyrosines, Y19 and Y27, during DNA nicking by pCU1 TraI relaxase (Figure 3.5, Figure 3.6). We hypothesized that a neighboring basic residue capable of activating a subset of the four tyrosines could mediate the observed pattern in DNA nicking activity.

Aspartic acid residues near the active site of the F plasmid TraI relaxase (D81) and the plasmid R388 TrwC relaxase (D85) are known to affect the efficiency of DNA nicking and plasmid transfer for these two systems. The role of these aspartic acids is to both improve metal binding by the HUH motif, as well as to activate the primary nicking tyrosine for attack by deprotonating the ring hydroxyl^{13, 36, 39}. The pCU1 TraI relaxase contains an aspartic acid at residue 84 that is properly positioned to interact with H162 of the HUH motif. D84 is also the primary candidate in the pCU1 TraI relaxase to act as a general base to deprotonate the nicking tyrosine's ring hydroxyl and activate it for catalytic attack (Figure 2.3, Figure 2.4).

Therefore, to investigate the role of D84 during DNA nicking by the pCU1 TraI relaxase, D84A mutants were created in the background of the wild type (WT) relaxase, as well as Y19,26,27F, Y26,27F, Y18,19,27F, and Y18,19F mutants of the relaxase, and the DNA nicking activity of each protein variant was examined using two DNA substrates (Figure 3.7). By mutating D84 in the background of each of these tyrosine mutants, the impact of D84 on specific tyrosine residues could be specified. The DNA nicking activities of WT_{_299}, Y26,27F_{_299}, Y18,19F_{_299}, and Y18,19,27F_{_299} were all reduced upon mutation of D84 (Figure 3.7). Interestingly, though, Y19,26,27F_{_299} exhibited an increase in DNA nicking activity in the presence of both substrates upon D84A mutation (Figure 3.7). Therefore, mutation of D84 revealed yet another distinction in nicking activity between Y18 and Y26. As discussed below, we propose that these data indicate D84 activates Y18 for catalysis indirectly through Y19, but activates Y26 directly (Figure 3.8). These data

support the hypothesis that the pCU1 TraI relaxase has modified the catalytic relaxase mechanism previously described^{28, 33, 37, 38} in order to nick its DNA substrate.

3.5.6 Model of DNA Nicking by the pCU1 TraI Relaxase Active Site Tyrosines and Adjacent Aspartic Acid

Using the DNA nicking data detailed in Sections 3.5.4 and 3.5.5, we generated a model in which a series of interactions formed between D84 and the four catalytic tyrosines of pCU1 TraI mediate cleavage of the scissile phosphate of the bound DNA substrate (Figure 3.8, Table 3.5). In this model, D84 serves as a base to deprotonate the ring hydroxyl of Y19 or Y26, thus generating a nucleophile at this residue. Either Y18 or Y26 can assume an orientation from which it performs an S_N2 attack on the scissile phosphate of a bound DNA substrate. Therefore, even in the triple mutant Y18,19,27F_299, Y26 can nick DNA with an activated, nucleophilic ring hydroxyl. However, when Y18 is in position to nick DNA, it still requires the presence of its partner tyrosine Y19 to serve as a bridge between itself and the activating effect of D84. Y27 appears to be a poor candidate for both acid-base interaction with D84 and scissile phosphate attack, except in the extreme case of the mutant Y18,26F_299. Here, Y27 appears capable of limited attack on the scissile phosphate when Y19 is present to form a bridge between Y27 and D84 (Figure 3.5).

While this model illustrates how the partner tyrosine Y19 enhances the DNA nicking activity of the primary tyrosine Y18, the presence of an alternate partner tyrosine was inhibitory under certain circumstances. In particular, the presence of the alternate partner Y19 reduced nicking by primary tyrosine Y26, and the alternate partner Y27 reduced nicking by primary tyrosine Y18(+Y19) (Figure 3.5). We propose that the role of D84 to polarize the tyrosine hydroxyls for attack actually results in this observed alternate partner effect. For example, in the case of Y26 or Y26+Y27, the addition of Y19 forces D84 to interact with either Y26 as before, leading to a productive nicking reaction, or to interact with Y19, which would lead to a nonproductive interaction. In the case of Y18+Y19, the addition of Y27 forces Y19, which is interacting with D84, to either maintain its interaction with Y18,

leading to a productive nicking reaction, or to interact with Y27, which results in limited DNA nicking.

Finally, the nicking activity of WT_{_299}, Y26,27F_{_299}, Y18,18F_{_299}, and Y18,19,27F_{_299} are all inhibited to ~50% of wild type upon mutation of D84 to alanine. In contrast, mutation of D84 increased the nicking activity of Y18 alone from ~10% of wild type to ~50% of wild type (Figure 3.7). Two conclusions are drawn from these data. First, since mutation of D84A did not eliminate the activity of wild type relaxase, we conclude that the nicking tyrosine is deprotonated for attack via a second, though less effective method. Two scenarios could produce this partially deprotonated tyrosine. If the local pH of the relaxase active site is higher than that of bulk solution, the tyrosine hydroxyl will exist in a protonated-deprotonated equilibrium. Alternately, a second basic residue could weakly activate the tyrosine. Analysis of the structure of the pCU1 TraI relaxase active site fails to reveal an obvious candidate for a secondary base; however, the surface of the active site cavity exhibits an overall negative potential which could increase the local pH. Therefore, it is most likely the tyrosine is weakly deprotonated due to a high local pH within the active site.

Second, since Y19,26,27F_{_299} was the only mutant to increase its DNA nicking activity in the presence of D84A, we conclude the increased activity is likely a result of a change in the net charge on the scissile phosphate in the relaxase active site upon mutation of D84. D84 is positioned to polarize one of the HUH motif histidines of pCU1 TraI relaxase (H162) by shifting the side chain proton on its ϵN to its δN , thus generating a lone pair of electrons and a partial negative charge at the ϵN , which faces the bound cation in the relaxase active site. As a result, the net charge on the active site decreases. However, mutation of D84 to alanine would increase the net active site charge by removing the polarizing influence of D84. Since Y18, when present alone, does not benefit from activation by D84 (Figure 3.8), this increase in net charge on the scissile phosphate is expected to enhance Y18's ability to attack the scissile phosphate within the active site.

In summary, these data describe the mechanism utilized by the pCU1 TraI relaxase to accomplish DNA nicking. This analysis provides another distinct example of the variety of

approaches conjugative relaxases employ to achieve plasmid transfer^{10, 16, 28, 37, 38, 51}. Previous data have shown that, during transfer of the conjugative F plasmid, its multi-tyrosine relaxase only uses its first tyrosine in primary sequence to nick DNA²⁸; in contrast, the closely related R388 TrwC multi-tyrosine relaxase uses the first tyrosine during initiation of plasmid transfer, but the third during termination^{37, 38}. The structurally homologous R1162 Mob A relaxase accomplishes plasmid transfer with only one functional DNA nicking tyrosine¹⁶. As we have outlined here, for pCU1 TraI, either the third tyrosine or a combination of the first and second tyrosines perform the critical first DNA cleavage step by this relaxase. It is possible that as additional plasmid systems and their respective relaxase enzymes are investigated, further mechanistic variations may be observed.

3.5.7 DNA Structure- and Sequence-Dependent Nicking and Ligation by the pCU1 TraI Relaxase

As detailed in Section 3.4, the pCU1 TraI relaxase constructs WT_299 and Nterm_299 bound DNA in a weak, non-specific manner. To determine if this relaxase also nicked and ligated DNA in a sequence-independent manner, the DNA nicking and ligation activities of WT_299 were tested using a series of DNA nicking and cross-over assays (Figure 3.3).

First, to examine the effect of DNA secondary structure on pCU1 TraI relaxase DNA nicking activity, 5' fluorescein-labeled substrates containing the pCU1 *oriT* sequence, *nic* site and either a full hairpin-forming inverted repeat (FAM-35/7oriT-hairpin), a mutated inverted repeat unable to form a hairpin (FAM-35/7-no_hairpin), or a truncated hairpin (FAM-20/7oriT-half_hairpin) were incubated with WT_299 (Table 3.3, Figure 3.3a). Each substrate was cleaved at the *nic* site, but the percent product generated increased if the inverted repeat (IR) did not form (Table 3.3, Figure 3.9). This same pattern was observed when measuring the DNA nicking activity of Y→F mutants of WT_299 (Figure 3.5). Both wild type and mutant constructs nicked a greater overall percentage of FAM-20/7oriT-half_hairpin as compared to FAM-35/7oriT-hairpin, and the relative DNA nicking activity of the majority of Y→F mutant versus to wild type was similar between substrates. Of note, two mutants did not follow this trend. First, Y19 + Y27 nicked the substrate FAM-20/7oriT-half_hairpin

more effectively than the substrate FAM-35/7oriT-hairpin, though mutant activity in the presence of both substrates was significantly reduced relative to wild type. Second, Y18 + 26 nicked the substrate FAM-35/7oriT-hairpin at wild type levels, but its activity fell to 60% of wild type when nicking FAM-20/7oriT-half_hairpin (Figure 3.5). As demonstrated for other homologous conjugative systems, the DNA nicking reaction of the relaxase exists in equilibrium with ligation of the nicked products, which regenerates the original DNA substrate^{35,37}. Since the wild type pCU1 TraI relaxase nicks a greater percentage of shorter, linear DNA substrates, it appears that a longer substrate including a complete hairpin shifts the nicking/ligation equilibrium of the pCU1 TraI relaxase in favor of ligation. However, in all situations, the substrate is nicked at the appropriate *nic* site.

Second, to examine the effect of nucleic acid sequence on pCU1 TraI relaxase-mediated DNA nicking, WT_299 was incubated with either an oriT-containing substrate or a control substrate that did not contain the pCU1 *oriT* sequence (Figure 3.3A). WT_299 only cleaved the *oriT*-containing substrate (Figure 3.9, Table 3.3). In the same manner, during DNA cross-over assays, WT_299 could only successfully ligate *oriT*-containing substrates (Figure 3.3B, Figure 3.9, Table 3.3). In fact, both the substrate spanning the *nic* site (referred to as donor substrate) and the substrate containing the *oriT* sequence 5' the *nic* site (referred to as the recipient) had to contain the pCU1 *oriT* sequence and the complete *nic* site, for successful TraI-mediated ligation. If either donor or recipient did not encode the *oriT*, or if the recipient was truncated prior to the *nic* site, no cross-over products were observed. DNA cross-over experiments involving two labeled DNA substrates, both spanning the *nic* site, tested the ability of pCU1 TraI relaxase to first cleave both substrates, and then ligate the products such that two novel DNA constructs would be generated, an unlabeled construct and a doubly-labeled construct (Figure 3.3B). In particular, the 5'-labeled FAM-20/7oriT-hairpin and the 3'-labeled 20/10oriT-half_hairpin-TAM (or 35/10oriT-hairpin-TAM) confirmed that WT_299 was capable of cleaving both substrates, crossing-over and ligating the nicked products, and then generating a substrate with both a 5' FAM and 3' TAM label (Figure 3.3B, Table 3.3).

To determine the specific bases within the pCU1 *oriT* sequence required for relaxase-mediated DNA nicking, a series of unlabeled DNA substrates were designed in which mutations were made within one or multiple regions of the wild type pCU1 *oriT* sequence (Table 3.3, Figure 3.10). These regions were designated “Hairpin,” “TAG,” “Pentanucleotide,” or “Post Nic,” and corresponded to sections of the F and R388 plasmid *oriT* previously shown to influence DNA nicking, binding, or plasmid transfer^{32, 35, 42, 43}. In light of the sequence-dependent nicking activity reported for the F and R388 relaxases, the TAG and Pentanucleotide regions were predicted to strongly influence DNA nicking by the pCU1 relaxase, while the Hairpin and Post Nic regions were expected to have only a limited effect. WT_299 was then incubated with each of these unlabeled mutant DNA substrates that spanned the *nic* site (the donor substrate) and a fluorescently-labeled recipient substrate (FAM-20oriT_half-hairpin) that encoded the pCU1 *oriT* and terminated at the *nic* site. This experiment thus measured the ability of the relaxase to nick the donor substrate and generate a 27 nucleotide long cross-over product (Figure 3.3B). The percent cross-over product generated from each mutant substrate was compared to that detected during a DNA cross-over experiment between a wild type donor substrate (35/7oriT-hairpin) and the same recipient substrate. Mutations within the Hairpin region, Post Nic region, or TAG region alone did not decrease overall relaxase activity. Mutation of the two flanking nucleotides within the Pentanucleotide region also had no inhibitory effect on enzyme activity. However, the mutation of bases within the core Pentanucleotide region eliminated enzyme activity, as did the mutation of a flanking nucleotide within the Pentanucleotide region along with a mutation within the Hairpin region or the Post Nic region (Table 3.3, Figure 3.10).

A number of the mutant substrates were nicked erroneously by WT_299 prior to ligation onto 20oriT-half_hairpin, thus creating products either longer or shorter than the expected 27mer (Table 3.3). In particular, these mis-nicked products were observed if the mutant substrate contained a duplicated or shifted Pentanucleotide region. If the mutant substrate also contained a mutant TAG region, multiple mis-nicked products were observed. If the mutant substrate contained a TAG and

Pentanucleotide region that were shifted in concert, only one alternative product was generated. However, mutation or shifting of the TAG region alone did not consistently affect nicking activity. The likelihood of erroneous nicking was also not affected by the presence, absence, or location of the hairpin forming inverted repeat.

As indicated by the results of the DNA cross-over assays, the Pentanucleotide region of the pCU1 *oriT* appears crucial for sequence-specific nicking by the pCU1 TraI relaxase. This region is expected to wrap around the N-terminus of the pCU1 TraI relaxase, forming an intra-strand U-turn that has been previously described for both the F and R388 *oriT*^{29,43}. To illustrate the importance of the contacts formed between the Pentanucleotide region and the pCU1 TraI N-terminus, we compared the function of WT_299 to that of the mutant pCU1 TraI relaxase, Nterm_299, in which three residues (Ser-Asn-Ala) were appended to the protein N-terminus. The nicking activity of Nterm_299 was in fact inhibited relative to WT_299, and during cross-over experiments involving 35/7oriT-hairpin and FAM-35oriT-hairpin, Nterm_299 appeared to nick the DNA substrate at erroneous *nic* sites (Figure 3.11). As described in Chapter 2, by comparing the structures of WT_299, Nterm_299, and the R388 TrwC relaxase³⁶ (PDB code 2CDM) (Figure 2.7), the N-terminus of Nterm_299 was shown to clearly clash with the DNA bound by the R388 relaxase. Therefore, these structural data clearly explains the functional inhibition observed during DNA nicking assays.

3.6 Summary of Data

In contrast to the high affinity, sequence-specific DNA binding of the F TraI and R388 TrwC relaxase enzymes^{32,35,39}, the pCU1 TraI relaxase binds DNA with low affinity. A second pCU1-encoded protein, TraK⁴⁶, may work in concert with pCU1 TraI to achieve tight, sequence-specific binding at the pCU1 *oriT*. Since pCU1 TraI binds DNA weakly with no sequence specificity, it was initially unclear whether or not this protein would nick DNA in the anticipated sequence-specific manner. However, DNA nicking assays indicated that the pCU1 TraI relaxase nicks DNA in a highly sequence specific manner, and bases within the Pentanucleotide region of the pCU1 *oriT* appear to cooperate with the N-terminus of WT_299 to position the *nic* site within the enzyme active site. In

the presence of a DNA substrate, the pCU1 TraI relaxase relies on the tyrosines of Motif 3, in particular Y26 or Y18 + Y19, and the conserved D84 to attack and cleave the scissile phosphate. Therefore, in summary, this work identified both the substrate bases and the protein residues responsible for first positioning a DNA substrate within the pCU1 TraI relaxase active site and then attacking and nicking the bound DNA substrate.

3.7 Tables and Figures

Table 3.1 DNA Binding Assay Substrates

Table 3.2 Thermal Stability of the pCU1 TraI Relaxase

Table 3.3 DNA Nicking and DNA Cross-over Assay Substrates

Table 3.4 Summary of Ion Metal Data

Table 3.5 Contribution of Individual Tyrosine and Aspartic Acid Residues to DNA Nicking

Figure 3.1 Circular Dichroism Spectra of the pCU1 TraI Relaxase

Figure 3.2 DNA Binding by the pCU1 TraI Relaxase

Figure 3.3 Schematics of DNA Nicking and Cross-over Experiments

Figure 3.4 Effect of Metal the DNA Nicking Activity of the pCU1 TraI Relaxase

Figure 3.5 DNA Nicking Activity of pCU1 TraI Relaxase Tyrosine Mutants

Figure 3.6 DNA Covalent Attachment Activity of pCU1 TraI Relaxase Tyrosine Mutants

Figure 3.7 DNA Nicking Activity of pCU1 TraI Relaxase Aspartic Acid Mutants

Figure 3.8 Model of DNA Nicking by the pCU1 TraI Relaxase Active Site

Figure 3.9 Representative Gels Illustrating the DNA Nicking and Ligation Activities of the pCU1 TraI Relaxase

Figure 3.10 Four Regions of the pCU1 *oriT* DNA Sequence

Figure 3.11 Importance of the pCU1 TraI Relaxase N-terminus during DNA Nicking

Substrate	Nucleic Acid Sequence	K _D (nM)	S.E.	R ₂
35/7oriT-hairpin	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G G T G T A T A G C A G	850	50	0.978
35oriT-hairpin	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G G T G T	890	60	0.976
35oriT-no_hairpin	C T G T A G T A G T A T G A C T A T C G C G C T G C G T T A G G T G T	810	50	0.994
20oriT-half_hairpin	T A T C G C G C T G C G T T A G G T G T	960	50	0.997
35mer-hairpin	C T A G C T C C G A G C A T A A G A G C T C G G A C T A C G T G A T C	810	40	0.996
35mer-no_hairpin	T G C G T G C G T A G T G T C T A T A G C G G A G A T C C T G G A G T	740	80	0.982
29mer	C C A A A C G A G C C A G C G A G C G A G C G A A C G C G	1180	70	0.995
24mer	C G A G C C A G C G A G C G A G C G A A C G C G	850	70	0.991
19mer	C A G C G A G C G A G C G A A C G C G	750	50	0.995
10mer	A G C G A A C G C G	1900	110	0.996

Table 3.1 DNA Binding Assay Substrates

The sequence, secondary structure, and the binding affinity, as measured by fluorescence anisotropy (FA)-based DNA binding assays, of WT_299 for each substrate are provided. Names of substrates containing a wild type pCU1 *nic* site and *oriT* include the phrase “oriT”. Names of mutant substrates include the phrase “mer”. In the text and accompanying figures, 5' fluorescein-labeled substrates are indicated by the prefix “FAM”. Mutated bases are in red. Sequences containing an inverted repeat that are predicted to form a hairpin are boxed. Each dissociation constant (K_D) reflects the affinity of WT_299 for the indicated substrate, with one exception. The K_D reported for 35/7oriT-hairpin reflects the affinity of Y18,19,26,27F_299 for this substrate.

Construct	Calculated T_m (°C)	
	Apo Enzyme	Enzyme + 6-fold excess DNA
WT_299	42.7 +/- 0.4	61.0 +/- 0.1
Nterm_299	44.6 +/- 0.3	54.6 +/- 0.3

Table 3.2 Thermal Stability of the pCU1 TraI Relaxase

Substrate	Nucleic Acid Sequence	Mis-Nicking	Nicking Activity	S.E.	Ligation Activity	S.E.
35/7oriT-hairpin	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G G T A G G T G T A T A G C A G		22.91	3.91	28.76	3.69
35/10oriT-hairpin	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G G T A G G T G T A T A G C A G G T T				9.06	2.24
35/7oriT-no_hairpin	C T G T A G T A G T A T G A C T A T C G C G C T G C G T T A G G T A G G T G T A T A G C A G		45.98	1.69	29.08	1.84
20/7oriT-half_hairpin	T A T C G C G C T G C G T T A G G T A G G T G T A T A G C A G		39.19	4.52		
20/10oriT-hairpin	T A T C G C G C T G C G T T A G G T A G G T G T A T A G C A G G T T				24.78	0.70
42mer-hairpin	C T A G C T C C G A G C A T A A G A G C T C G G A C T A C G T G A T C T A C T A T A		1.68	0.64		
35/7oriT-hairpin_mutant	A T A G C T C C G A G C A T A A G A G C T C G G A C T T A G G T G T A T A G C A G				42.99	1.81
35/7oriT-hairpin_shift	T A T C C G A G C A A G A G C T C G G C T T G C G T T A G G T G T A T A G A G				38.50	0.70
35/7oriT-TAG_L2+hairpin_mutant	C C A C G T A C G A G C A T A A G A G C T C G T G C T A G C A G A G T G T A T A G A G				36.52	1.85
35/7oriT-TAG_R2+hairpin_mutant	C C A G C T A C G A G C A T A A G A G C T C G T G C G T C A T A G T A T A G A G				31.49	0.33
F_plasmid_oriT-G1A	A A A A T C A G C A A A C T T G T T T T G C G T G G G T G T A G T G C T T				73.50	2.92
35/7oriT-TAG_mutant	T G T G A T A G C G T G A T T T A T C G C G C T G C G T C A T G T G T A T A G C A G				53.47	0.69
35/7oriT-TAG_L2	T G T G A T A G C G T G A T T T A T C G C G C T G C T A G C A G T G T A T A C C A G				40.11	1.95
35/7oriT-TAG_R2	T G T G A T A G C G T G A T T T A T C G C G C T G C G T C A T A G T A T A G C A G				25.70	1.17
35/7oriT-pentanuc_mutant_A	T G T G A T A G C G T G A T T T A T C G C G C T G C C T A G A C A C G T A G C A G				0.72	0.43
35/7oriT-pentanuc_mutant_B	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G T T G T T A G A G				1.44	1.30
35/7oriT-G144T,T141G	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G T T G T T A G A G				2.69	0.92
35/7oriT-G144T	G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G T T G T A T A G C A G				46.03	2.08
35/7oriT-A1T	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G G T G T T A G C A G				30.23	3.06
35/7oriT-postnic_mutant	T G T G A T A G C G T G A T T T A T C G C G C T G C G T T A G G T G T A A G T A G G				32.94	1.76
35/7oriT-hairpin+postnic_mutant	A T A G C T C C G A G C A T A A G A G C T C G G A C T T A G G T G T T A G T A G G				8.01	1.65
35/7oriT-4_region_mutant	C T A G C A C C G A G C A T A A G A G C T C G G A G T A G T G C G T A A T T A T A				1.29	0.44
35/7oriT-pentanuc_shift	T G T G A T A G C G T G A T T T A T C G C G C T C G G T T A G C A G T G T A T A G	yes			35.54	0.58
ForIT	A A A A T C A G C A A A A C T T G T T T T G C G T G G G T G T G G T G C T T	yes			27.06	0.43
35/7oriT-hairpin+pentanuc_shift	C T A G C A T C C G A G C A A G A G C T C G G T A G G T G T G T A A G T A T A	yes			27.83	2.36
35/7oriT-pentanuc_duplicate_A	C T A G C A C C G A G C A T A A G A G C T C G G A G T A G T G T G T A A G T A T A	yes			46.59	3.99
35/7oriT-pentanuc_duplicate_B	C T A G C T C C G A G C A T A A G A G C T C G G A C T G T G T G T A T G T A T A	yes			25.07	1.16

Table 3.3 DNA Nicking and DNA Cross-over Assay Substrates

The DNA sequence, secondary structure, and nicking or ligation activity of WT_299, as measured by DNA nicking and cross-over assays, for each substrate are provided. The names of substrates containing a wild type *nic* site and minimal *oriT* include the phrase “oriT”. The names of mutant substrates include the phrase “mer”. In the text and accompanying figures, fluorescently-labeled substrates are indicated by the prefix “FAM” (5' fluorescein label) or the suffix “TAM” (3' TAMRA label). Mutant substrates are listed according to the region of the pCU1 *oriT* which is mutated (see Figure 3.10 for a definition of these regions). Mutated bases are in red. Shading emphasizes the key areas of each mutant substrate relative to the wild type substrate. Sequences forming an inverted repeat that are predicted to form a hairpin are boxed in black. The “TAG” region of each substrate is boxed in green. The “Pentanucleotide” region of each substrate is boxed in red. The predicted *nic* site(s) for each substrate is indicated by black vertical double lines, and non-canonical *nic* sites, as observed in DNA cross-over assays, are indicated by red vertical double lines. “Mis-nicking” activity is reported for each substrate that generated a product during DNA cross-over assays that was larger or smaller than anticipated, indicating that the initial nicking reaction most likely occurred at a non-canonical *nic* site. DNA nicking and ligation activity is represented as percent activity as defined in the text. The ligation activity reported for each substrate was calculated from experiments involving the listed unlabeled substrate and FAM-20oriT-half_hairpin, with two exceptions. The ligation activity reported for 20/10oriT-half_hairpin and 35/10oriT-hairpin represents the percent doubly-labeled product generated during experiments involving these two 3'-TAM-labeled substrates and FAM-20/7oriT-half_hairpin. Errors reported for these two experiments and the nicking experiments involving 35/7oriT-hairpin and 20/7oriT-half_hairpin represent the standard deviation of the means of three experiments performed in triplicate. For all other data, each data point is the average of at least three replicate measurements, and the error is the standard error of these measurements.

Metal	Mg	Mn	Ca	Fe(III)	Ni	Cu	Zn
Conc. in Bacterial Cells (μM) ¹	>10,000	10	100	100	Low, n/a	10	100
Fold increase in purified sample over background ² (Protein: WT_299)	1.16	1.50	1.10	1.15	7.75	Not tested	45.12
Percent product generated ^{a,2} (Substrate: 35/7oriT-hairpin)	46.5	33.9	42.1	N/A	34.4	19.1	6.7
Satisfies electron density? ³	No	Yes	No	Yes	Yes	Yes	Yes
Preferred coordination ⁴	6	5,6	6	5,6 ^b	6	3,4	4

Table 3.4 Summary of Metal Ion Data

^a DNA nicking activity reported is the amount of product generated at the optimal concentration for each metal; Cu²⁺ becomes inhibitory as its concentration increases.

^b This analysis data includes both ferric and ferrous ions

¹ Finney, et al., 2003⁶⁸

² For details of method used, see Chapter 3, Sections 3.5.2, 3.5.3

³ For details of method used, see Chapter 2, Section 2.5.2

⁴ Harding, et al., 2001⁶⁹

Relaxase Construct	% Activity FAM-35/7	% Activity FAM-20/7	Role of Individual Residues during DNA Nicking
Y – Y – Y – Y (Wild type relaxase)	100	100	Y18 attacks scissile PO ₄ after D84 indirectly deprotonates Y18 via Y19 OR Y26 attacks scissile PO ₄ after D84 directly deprotonates Y26
Y – F – F – F	30**	20**	Y18 is too distant from D84 to be deprotonated, is limited to inefficient attack on scissile PO ₄
Y – Y – F – F	125	113	Y18 now benefits from indirect deprotonation by D84 via Y19, is capable of efficient attack on scissile PO ₄
Y – Y – F – Y	45**	55**	D84 deprotonates Y19. Y19 forms bifurcated interaction with Y18 (capable of efficient attack on scissile PO ₄) and Y27 (limited ability to attack scissile PO ₄)
F – F – Y – F	80	90	D84 deprotonates Y26, which in turn is capable of efficient attack on scissile PO ₄
F – F – Y – Y	100	110	No change in DNA nicking as compared to F – F – Y – F; Y27 is not deprotonated by either D84 or Y26
F – Y – Y – Y	35**	70**	D84 deprotonates both Y26 (capable of efficient attack on scissile PO ₄) and Y19 (does not attack scissile PO ₄)
F – Y – F – F	0**	0**	Y19 cannot not attack scissile PO ₄
F – F – F – Y	0**	0**	Y27 is not deprotonated by D84
F – Y – F – Y	25**	70**	Y19 forms bridge between D84 and Y27, which has a limited ability to attack scissile PO ₄
Y – F – Y – F	95	60**	Activity of this mutant primarily reflects that of Y26
Y – Y – Y – F	125	90	Activity of this mutant reflects that of both Y26 and Y18+Y19
Y – F – Y – Y	115	95	Activity of this mutant primarily reflects that of Y26
D84A Mutants			D84 mediated deprotonation of tyrosines is eliminated AND Net charge on active site increases
D84A – Y – Y – Y – Y	65‡	60‡	Nicking by WT tyrosines (Y26 and Y18+19) is less efficient due to elimination of tyrosine deprotonation by D84
D84A – Y – F – F – F	60†	70‡	Nicking by Y18 is more efficient due to increase in net active site charge (note: Y – F – F – F is not deprotonated by D84)
D84A – Y – Y – F – F	35‡	45‡	Nicking by Y18+19 is less efficient due to elimination of tyrosine deprotonation by D84
D84A – F – F – Y – F	25‡	35‡	Nicking by Y26 is less efficient due to elimination of tyrosine deprotonation by D84
D84A – F – F – Y – Y	40‡	40‡	Nicking by Y26 is less efficient due to elimination of tyrosine deprotonation by D84

Table 3.5. Contribution of Individual Tyrosine and Aspartic Acid Residues to DNA Nicking

A summary of the DNA nicking activity of each tyrosine and aspartic acid mutant relaxase construct is provided for two substrates (FAM-35/7oriT-hairpin and FAM-20/7oriT-half hairpin), as well as the proposed role of each of the four tyrosines and the active site aspartic acid during DNA nicking. % Activity represents DNA nicking activity relative to WT₂₉₉. ** indicates nicking activity is inhibited relative to WT with 95% confidence. ‡ indicates nicking activity of the D84A, Y→F mutant is inhibited relative to its D84, Y→F counterpart with 95% confidence. † indicates nicking activity of the D84A, Y→F mutant is enhanced relative to its D84, Y→F counterpart with 90% confidence.

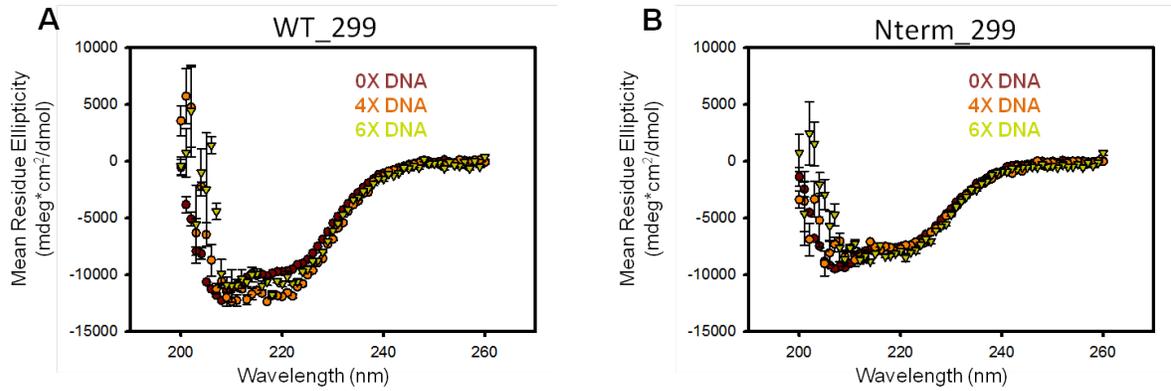


Figure 3.1 Circular Dichroism Spectra of the pCU1 TraI Relaxase

The calculated mean residue ellipticity (MRE, $\text{mdeg} \cdot \text{cm}^2 / \text{dmol}$) of relaxase and relaxase + DNA (35oriT-hairpin) is plotted as a function of wavelength. Unbound enzyme is represented by dark red data points; ~97% (4-fold DNA excess) and ~100% (6-fold DNA excess) DNA-bound enzyme is represented by orange and yellow data points, respectively. Each data point is the average of three background subtracted MRE values (where background subtracted = experimental sample signal – control sample signal, where control samples are equivalent to experimental samples, but without the protein component), and each error bar represents the standard error of the variance of three background subtracted MRE data points.

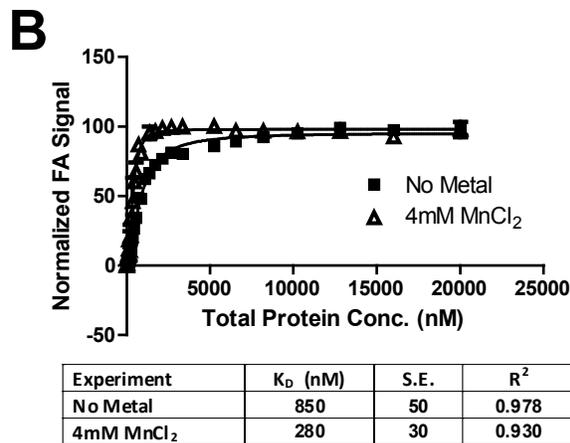
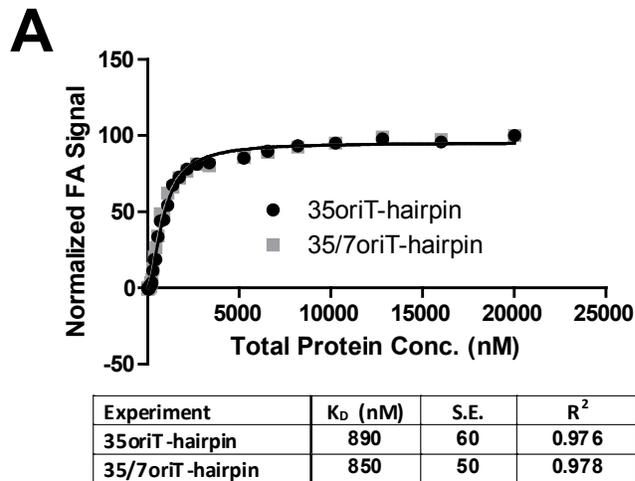


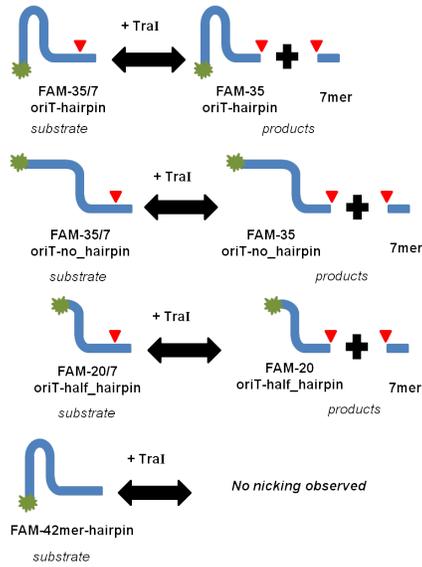
Figure 3.2 DNA Binding by the pCU1 TraI Relaxase

DNA binding curves were generated by monitoring the change in fluorescence anisotropy of a 5'-fluorescein (FAM)-labeled DNA substrate as a function of relaxase protein concentration. The DNA binding affinity (represented as the dissociation constant, K_D , in nM) was calculated by fitting the curves with Equation 3.2 using nonlinear regression.

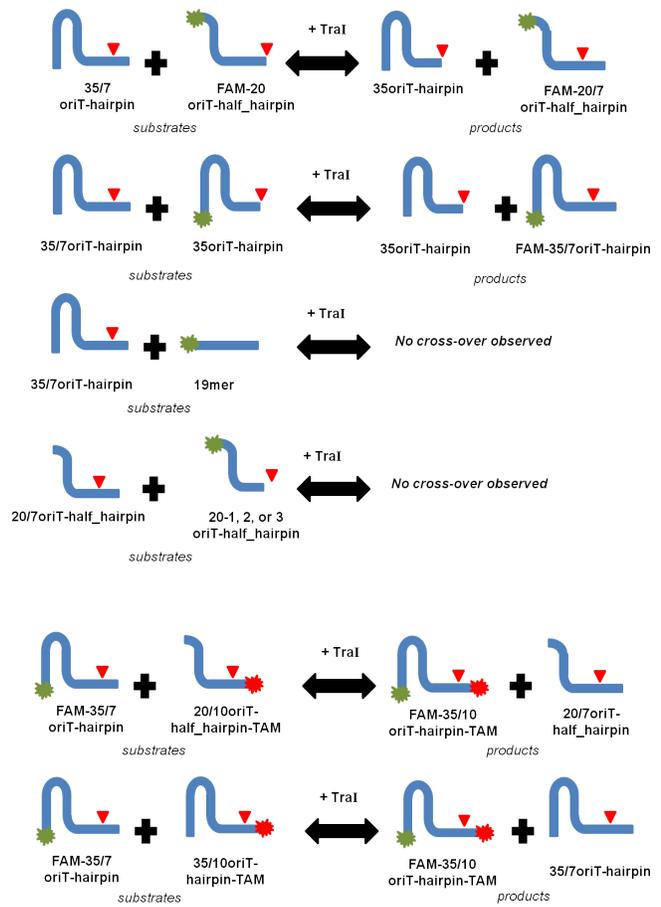
A) 35oriT-hairpin and 35/7oriT-hairpin are bound by WT₂₉₉ and Y18,19,26,27F₂₉₉, respectively. The substrate 35/7oriT-hairpin extends 7 nucleotides beyond the pCU1 *oriT nic* site, while 35oriT-hairpin is truncated at the *nic* site.

B) 35/7oriT-hairpin is bound by Y18,19,26,27F₂₉₉ in the presence or absence of 4 mM MgCl₂.

A. Schematics of DNA Nicking Assays



B. Schematics of DNA Cross-over Assays



C. Schematics of DNA Covalent Attachment Assays

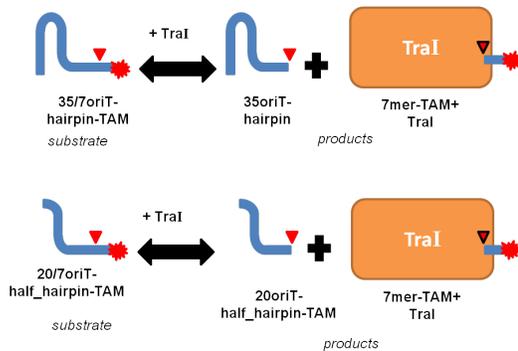


Figure 3.3 Schematics of DNA Nicking and Cross-over Experiments

DNA substrates are drawn as blue lines. TraI is drawn as an orange rectangle. The location of the fluorescein label on fluorescent substrates is indicated by a green star, the location of the TAMRA label is indicated by a red star, and the substrate *nic* site by a red triangle.

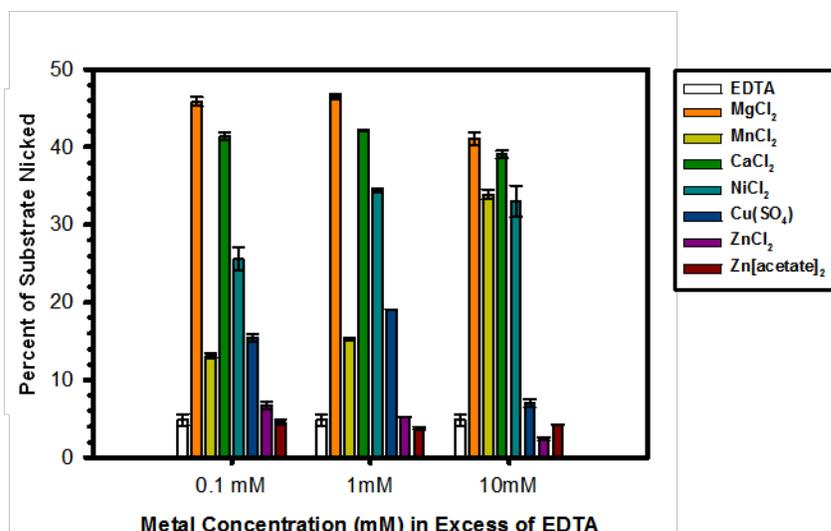


Figure 3.4 Effect of Metal on the DNA Nicking Activity of the pCU1 TraI Relaxase

To determine the DNA nicking activity of the pCU1 TraI relaxase (WT_299) after treatment with EDTA and in the presence of the indicated concentration of metal, 5 μ M enzyme was incubated with 1 μ M substrate (FAM-35/7oriT-hairpin) for 1 h at 37 °C, and products were then separated on denaturing polyacrylamide gels. Each experiment was performed in triplicate; the error is the standard error of these three measurements. Additional reaction conditions and methods can be found in Sections 3.5.2, 3.5.3.

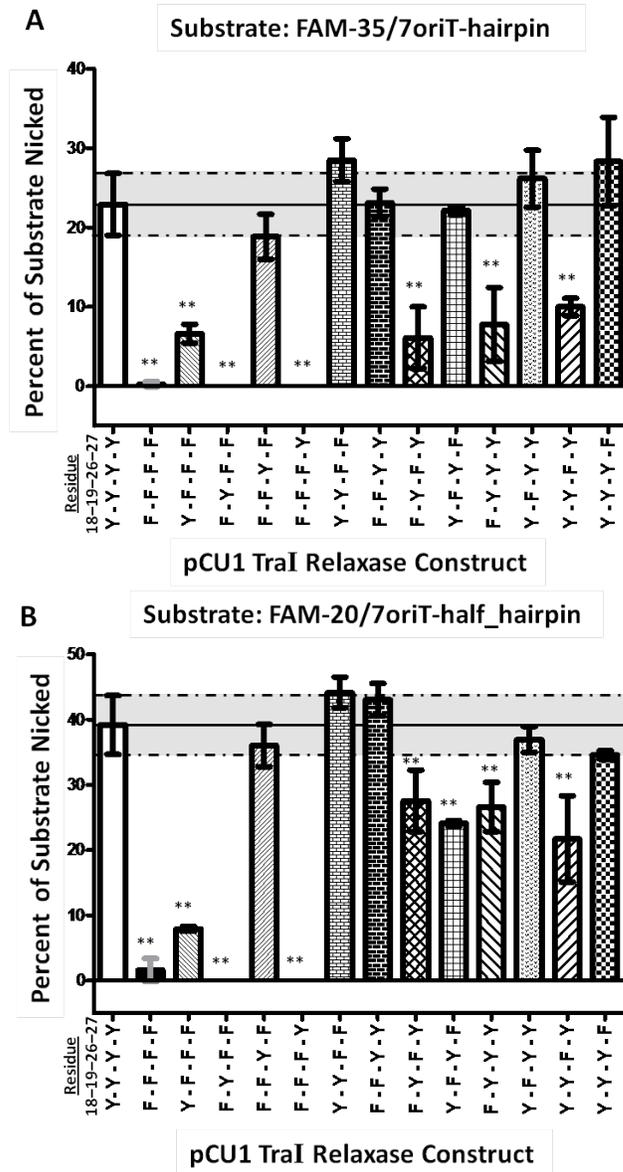


Figure 3.5 DNA Nicking Activity of pCU1 TraI Relaxase Tyrosine Mutants

The DNA nicking activity of each TraI construct is represented as the percent DNA substrate nicked and error bars represent the standard deviation of three experiments, each performed in triplicate. Each construct is labeled according to whether a tyrosine (Y) or phenylalanine (F) is present at residues 18, 19, 26, and 27. ** indicates nicking activity is inhibited relative to WT with 95% confidence. Solid and dashed lines represent the average nicking activity of WT and one standard deviation above and below this value, respectively. Experimental details can be found in Sections 3.5.2, 3.5.4.

A) The substrate tested, FAM-35/7oriT-hairpin, spans the pCU1 *nic* site and forms a hairpin upstream of the *nic* site.

B) The substrate tested, FAM-20/7oriT-half_hairpin, spans the pCU1 *nic* site and incorporates only one arm of the hairpin.

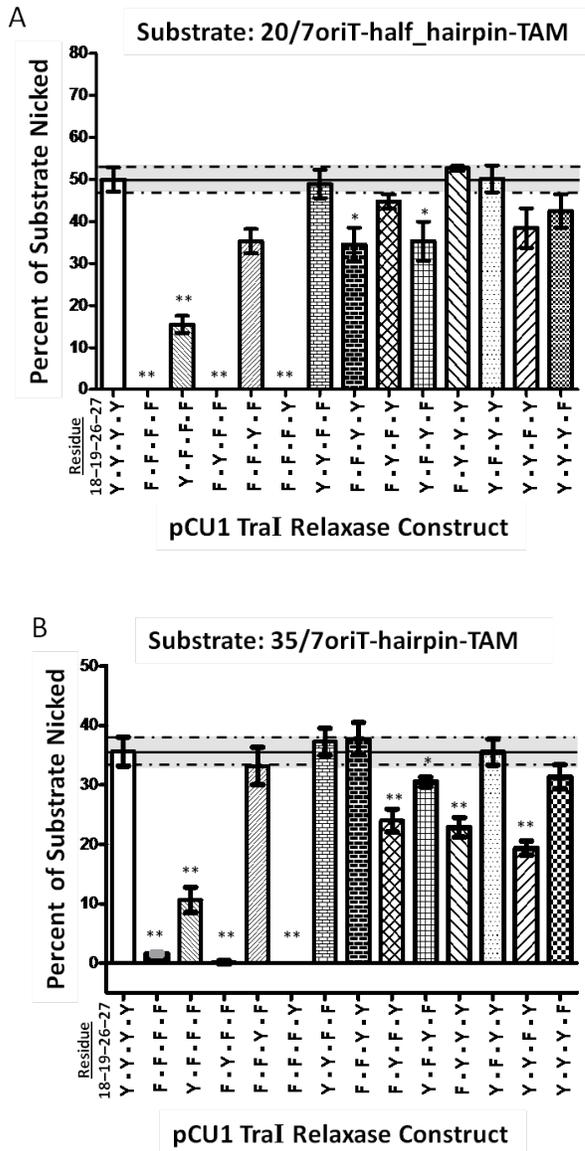


Figure 3.6 DNA Covalent Attachment Activity of pCU1 TraI Relaxase Tyrosine Mutants

The DNA covalent attachment activity of each TraI construct is represented as the percent substrate nicked and error bars represent the standard deviation of three experiments, each performed in triplicate. Each construct is labeled according to whether a tyrosine (Y) or phenylalanine (F) is present at residues 18, 19, 26, and 27. ** and * indicates nicking activity is inhibited relative to WT with 95% or 90% confidence, respectively. Solid and dashed lines represent the average nicking activity of WT and one standard deviation above and below this value, respectively. Experimental details can be found in Sections 3.5.2, 3.5.4.

A) The substrate tested, 20/7oriT-half_hairpin-TAM, spans the pCU1 *nic* site and incorporates only one arm of the hairpin.

B) The substrate tested, 35/7oriT-hairpin-TAM, spans the pCU1 *nic* site and forms a hairpin upstream of the *nic* site.

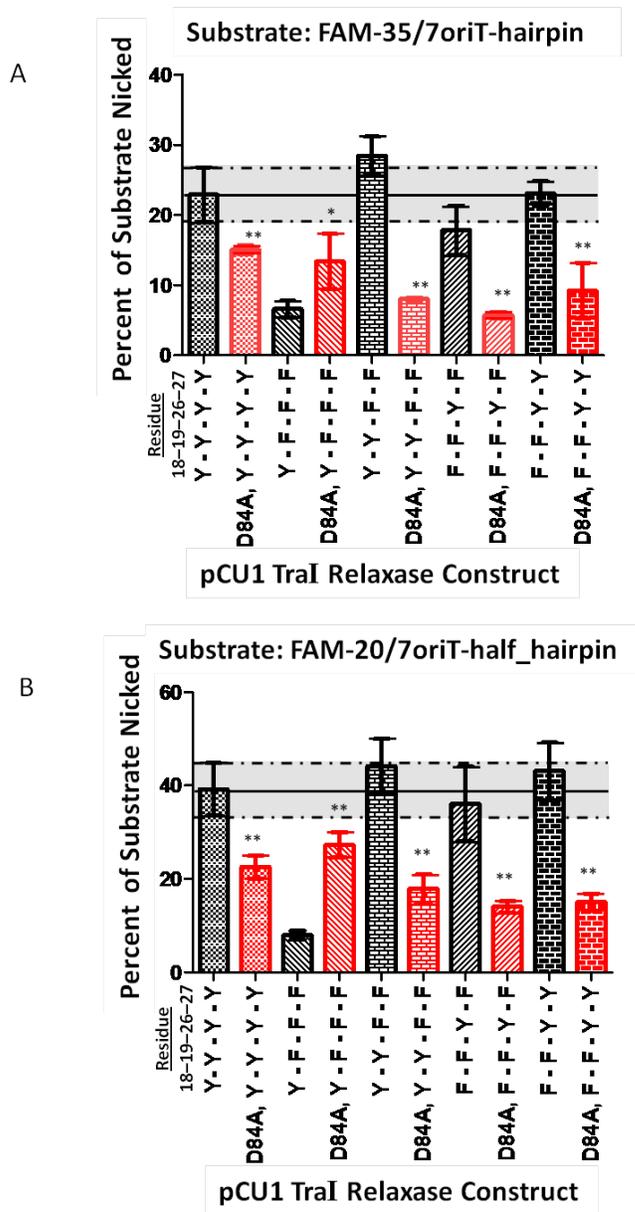


Figure 3.7 DNA Nicking Activity of pCU1 TraI Relaxase Aspartic Acid Mutants

The DNA nicking activity of each TraI construct is represented as the percent DNA substrate nicked and error bars represent the standard deviation of three experiments performed in triplicate. Each construct is labeled according to whether a tyrosine (Y) or phenylalanine (F) is present at residues 18, 19, 26, and 27. Data generated by D84A containing mutants is represented by red bars. ** and * indicates nicking activity is altered relative to the corresponding tyrosine mutant with 95% and 90% confidence, respectively. Experimental details can be found in Sections 3.5.2, 3.5.5.

A) The substrate tested, FAM-35/7oriT-hairpin, spans the pCU1 *nic* site and forms a hairpin upstream of the *nic* site.

B) The substrate tested, FAM-20/7oriT-half_hairpin, spans the pCU1 *nic* site and incorporates only one arm of the hairpin.

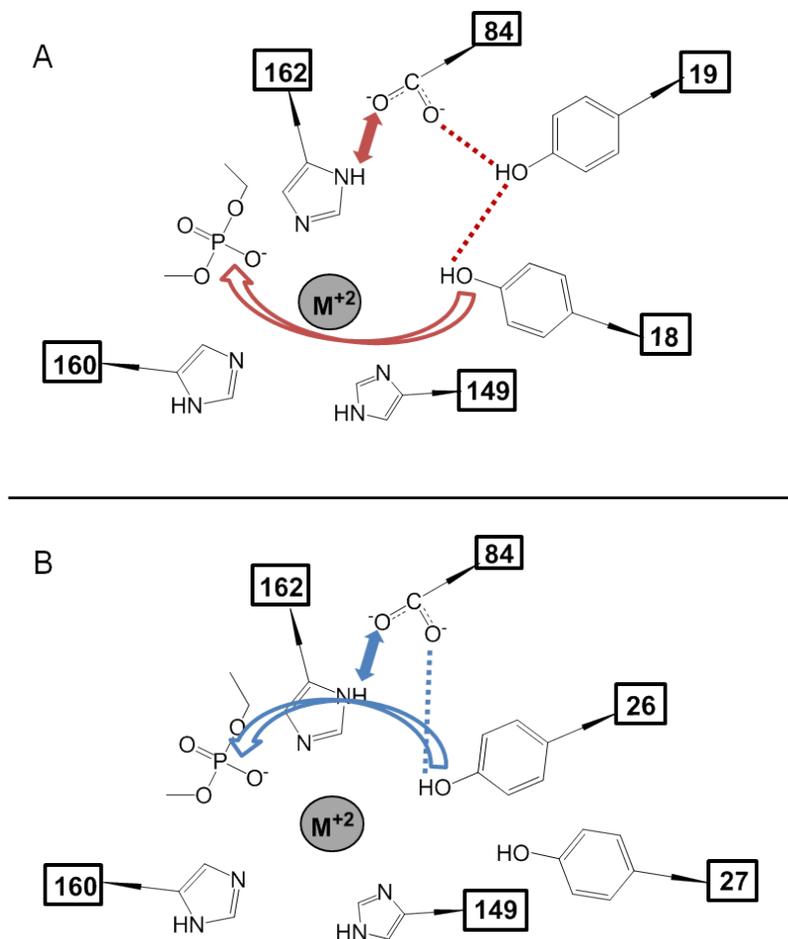


Figure 3.8 Model of DNA Nicking by the pCU1 TraI Relaxase Active Site

A) The proposed orientation of tyrosines 18 and 19 relative to the active site base, aspartic acid 84, and the scissile phosphate of the DNA substrate, based upon DNA nicking activity data. Red dotted lines and bidirectional arrows indicate the presence of residue-residue interactions. The red curved arrow indicates the S_N2 attack of Y18 on the scissile phosphate.

B) As in (A), the proposed orientation of tyrosines 26 and 27 relative to the active site base, aspartic acid 84, and the scissile phosphate of the DNA substrate, based upon DNA nicking activity data. Blue dotted lines and bidirectional arrows indicate the presence of residue-residue interactions. The blue curved arrow indicates the S_N2 attack of Y26 on the scissile phosphate.

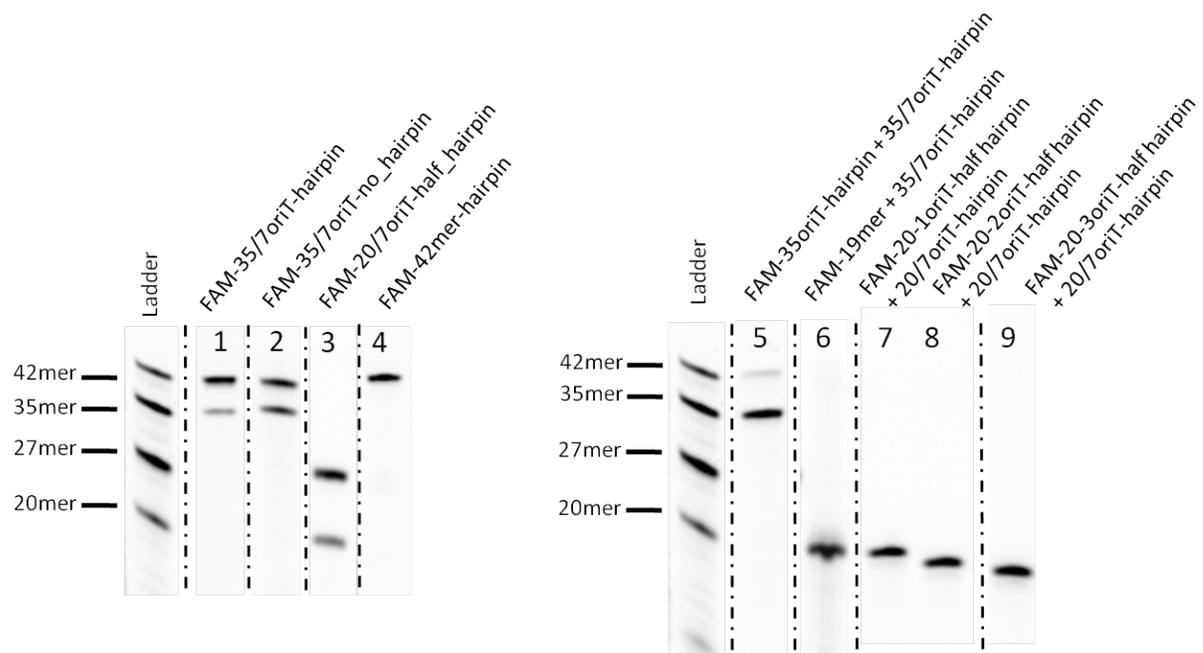


Figure 3.9 Representative Gels Illustrating the DNA Nicking and Ligation Activities of the pCU1 TraI Relaxase

DNA nicking and cross-over substrates and products were separated on denaturing polyacrylamide gels prior to quantification. Representative gels are provided here. Dashed lines indicate that the lanes provided are from different gels or different sections of one gel. Each lane is aligned with the provided ladder as it appeared on the original gel. Experimental details can be found in Sections 3.5.2, 3.5.7.

In lanes 1, 2 and 4, WT_299 was incubated with a 5' fluorescently-labeled *oriT*-containing 42 nucleotide (nt) long substrate that, if nicked, would form a fluorescently-labeled 35mer product.

In lane 3, WT_299 was incubated with a 5' fluorescently-labeled *oriT*-containing 27nt long substrate that, if nicked, would form a fluorescently-labeled 20mer product.

In lane 5, WT_299 was incubated with an unlabeled *oriT*-containing 42nt long substrate and a 5' fluorescently-labeled *oriT*-containing 35nt long substrate. Upon nicking the 42mer and successfully ligating the nicked product onto the 5' fluorescently-labeled substrate, a 5' fluorescently-labeled 42mer would be generated.

In lane 6, WT_299 was incubated with an unlabeled *oriT*-containing 42mer and a 5' fluorescently-labeled 19mer not encoding the *oriT* sequence. Upon nicking the 42mer and successfully ligating the nicked product onto the fluorescently-labeled substrate, a 5' fluorescently-labeled 26mer should be generated.

In lane 7-9, WT_299 was incubated with an unlabeled *oriT*-containing 27mer and a 5' fluorescently-labeled *oriT*-containing 19, 18, or 17mer truncated 1, 2, or 3 bases prior to the *nic* site. Upon nicking the 27mer and successfully ligating the nicked product onto the 5' fluorescently-labeled substrate, a 5' fluorescently-labeled 26, 25, or 24mer should be generated.

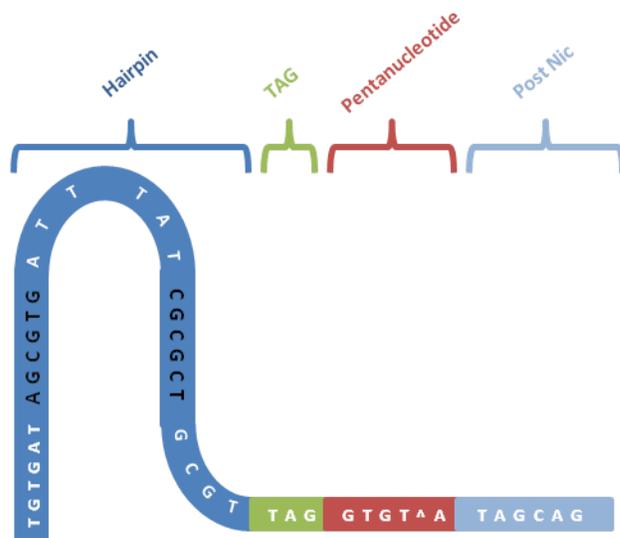


Figure 3.10 Four Regions of the pCU1 *oriT* DNA Sequence

The *oriT* sequence of plasmid pCU1 is divided into four regions (Hairpin in dark blue, TAG in green, Pentanucleotide in red, and Post Nic in light blue). The bases forming the inverted repeat are shaded black and the location of the predicted hairpin is shown. The location of the *nic* site (T[^]A) is indicated by a caret.

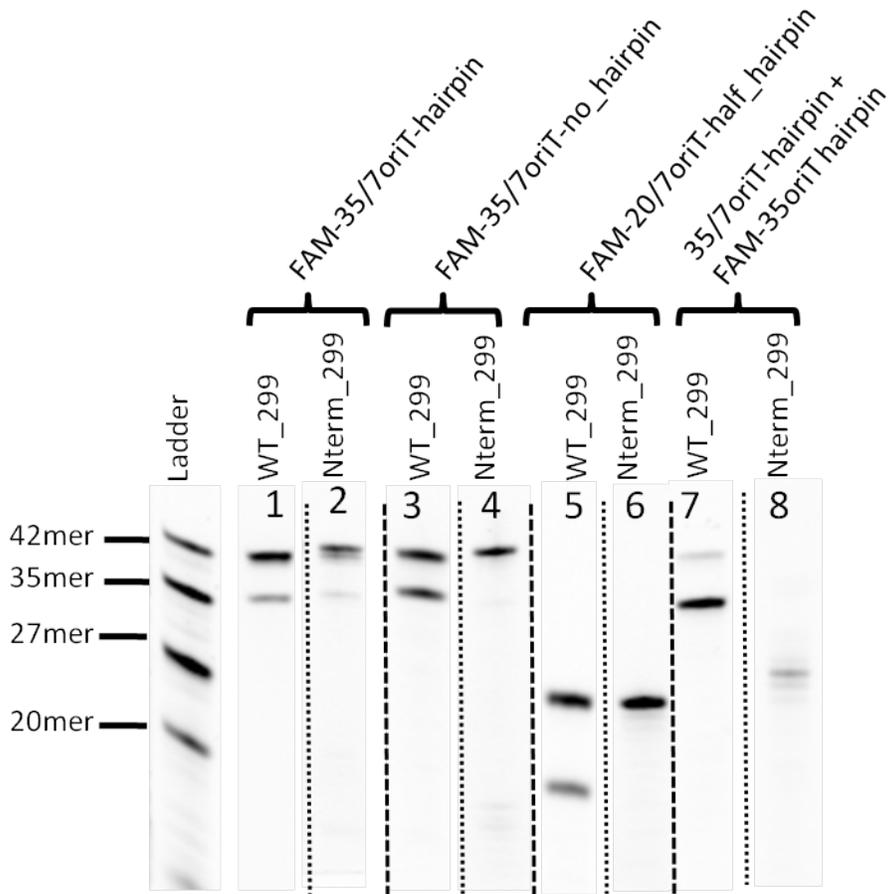


Figure 3.11 Importance of the pCU1 TraI Relaxase N-terminus during DNA Nicking

DNA nicking and cross-over substrates and products were separated on denaturing polyacrylamide gels prior to quantification. Representative gels are provided here. Dashed lines indicate that the lanes provided are from different gels or different sections of one gel. Each lane is aligned with the provided ladder as it appeared on the original gel. Experimental details can be found in Sections 3.5.2, 3.5.7.

In lanes 1-4, enzyme was incubated with a 5' fluorescently-labeled *oriT*-containing 42mer substrate that, if nicked, would form a 35mer product.

In lanes 5 and 6, enzyme was incubated with a *oriT*-containing 5' fluorescently-labeled 27mer substrate that, if nicked, would form a 20mer product.

In lanes 7 and 8, enzyme was incubated with an unlabeled *oriT*-containing 42mer and a 5' fluorescently-labeled *oriT*-containing 35mer. Upon nicking the 42mer and successfully ligating the nicked product onto the 5' fluorescently-labeled substrate, a 5' fluorescently-labeled 42mer would be generated.

Chapter 4: DNA Binding and DNA-Dependent ATPase Activities of the pCU1 TraI Helicase

4.1. The Classification of Helicases

In addition to the DNA nicking and ligation activities of the relaxase, CPT relies upon the DNA unwinding activity of a DNA helicase to separate the T-strand from the parent strand. DNA helicases harness the energy released upon hydrolysis of ATP to translocate along one or both strands of a dsDNA substrate, driving apart the two strands⁵⁵. They are classified into six superfamilies based upon the presence and sequence of amino acid motifs characteristic of helicases (Figure 4.1)⁷¹. A closely related group of enzymes, the translocases, contain many of the motifs originally assigned solely to helicases⁷¹.

Within each helicase superfamily, individual family members are further defined based upon their mechanism of action when separating DNA strands. If the enzyme primarily contacts one of the two DNA strands, it is classified as a type α helicase, while if the enzyme contacts both strands, it is considered a type β helicase. When traveling along the DNA strand, if the helicase travels in the 3' to 5' direction, it is classified as a type A helicase, but if it travels 5' to 3', it is classified as a type B helicase⁷¹.

Helicases vary in the oligomeric state they assume when actively separating DNA strands. The oligomeric state preferred by the helicase strongly influences the mechanism the enzyme will employ when translocating along a DNA strand. Monomeric helicases primarily utilize the “inchworm” or “hand-over-hand” model, during which the enzyme uses the hydrolysis of ATP to undergo conformational changes that result in the ratcheting of the enzyme along the DNA strand. As a result, monomeric helicases are primarily type α helicases (Figure 4.2A). Dimeric helicases primarily use the rolling model of translocation. Each monomer switches between ssDNA-bound and dsDNA-bound states as the helicase “walks” down the DNA strand. The monomer in the dsDNA-

bound state melts the bound DNA helix and then releases one strand, as the second monomer then binds to dsDNA downstream of the first monomer. This cycle is driven by the hydrolysis of ATP within each monomer. Dimeric helicases are often highly processive, since one monomer is bound to the DNA substrate at all times (Figure 4.2B)⁷¹⁻⁷³. Hexameric helicases actually thread one strand of the DNA helix through the ring formed by the six monomers. The helicase then uses the hydrolysis of ATP to drive conformational changes in the monomers, which in turn translocate along the bound DNA strand, driving the helix apart (Figure 4.2C)⁷¹⁻⁷³. Finally, the RecBCD helicase is a heterotrimer of helicases, though one monomer (RecC) is not functional. Therefore, RecB and RecD together form a dimeric helicase. However, RecBD does not use the rolling model of DNA translocation typical of dimeric helicases; instead, both monomers use the inchworm model, but on separate strands. Since RecD is a type B helicase and RecC is a type A helicase, together they translocate in the same net direction along dsDNA (Figure 4.2D)⁷³.

Of the 6 superfamilies, superfamily 1 (SF1) and superfamily 2 (SF2) are the largest and contain monomeric and dimeric helicases. Superfamilies 3,4,5, and 6 are primarily hexameric ring helicases. SF1 and SF2 helicases can be either type A or type B, though of the known SF1 helicases, all are type α ⁷¹. The conjugative helicases of the F, R388, and pCU1 plasmids are classified as SF1B helicases and are expected to use the inchworm model of translocation⁷⁴⁻⁷⁶.

4.2. Introduction to Helicase Structure and Function

As stated in Section 4.1, helicases function as monomers, dimers, or hexamers. Regardless of their overall oligomeric state, most helicase monomers consist of two α - β domains, often referred to as domain 1A and domain 2A. The basic fold assumed by each of the core domains is conserved across helicase superfamilies and resembles that of the RecA protein. Thus, together the two core domains are often referred to as the “RecA core” of the helicase enzyme. Most commonly, only the N-terminal α - β domain (domain 1A) is capable of ATP binding and hydrolysis, while both domains bind and translocate along a DNA strand. Some helicase monomers also contain additional domains that “decorate” the RecA core (Figure 4.3)⁷⁷.

The characteristic motifs used to identify and classify helicases are found within domains 1A and 1B of the RecA core. Together they allow the helicase to hydrolyze ATP, bind DNA, and translate the energy of ATP hydrolysis to movement along the DNA strand. While each superfamily contains a unique number of motifs, the motifs required for ATP hydrolysis are highly conserved. Motif I (the Walker A motif) is predicted to coordinate the phosphate and Mg^{2+} ion of the bound MgATP/MgADP species. Motif II (the Walker B motif) is predicted to coordinate the Mg^{2+} ion and provide a catalytic base (Table 4.1)^{72,77}.

SF1 helicases, such as the pCU1 TraI helicase, contain seven well characterized motifs that are shared with SF2 helicases, as well as an additional motif recently proposed (TxGx) (Table 4.1)^{72,77}. The motifs involved primarily in DNA binding fall along the interface of the two α - β domains (Motifs Ia, III, IV, V, TxGx, and QxxR). The ATP binding site and the Walker A and Walker B motifs are found in the N-terminal α - β domain (1A), on the face of the domain opposite of that of the DNA binding site. Structural and mutational data describing the arginine residue of Motif VI have implicated this motif, and in particular this residue, as the linker between Walker A- and Walker B-mediated ATP binding and hydrolysis and DNA binding and translocation^{72,77}.

4.3. Analysis of the Sequence and Predicted Structure of the pCU1 TraI Helicase

By visually inspecting the amino acid sequence of pCU1 TraI and by aligning the sequence of pCU1 TraI with those of homologous conjugative helicases, each of the seven SF1 helicase motifs were tentatively identified within the predicted helicase domain of the pCU1 TraI enzyme (Table 4.1). As described in Section 4.5, the identities of Motifs I (Walker A) and II (Walker B) were confirmed by mutagenesis and ATPase activity assays. The N-terminal (residue 311) and C-terminal (residue 932) extent of the minimal helicase domain was determined by ATPase activity assays, again as described in Section 4.5.

The structure prediction software PHYRE generated two putative structures of the pCU1 TraI helicase using the input residues pCU1 TraI 311-1078 and pCU1 TraI 311-932 (Figure 4.5A). The two models were based upon the structure of the SF1 helicase RecD (PDB code 1W36, chain D), with

E-values of 6.4×10^{-26} (WT_311-1078) and 8.9×10^{-28} (WT_311-932) and an estimated precision of 100% for both submissions (sequence identity of aligned residues = 15% and 16% respectively). The N-terminal 35 residues and residues 787-841 of both WT_311-1078 and WT_311-932 were not included in the model; residues 346-445 were aligned with the accessory N-terminal domain of RecD, while the remaining residues formed the RecA core domains 1A and 2A and the accessory 1B domain of RecD (Figure 4.5A). pCU1 TraI helicase residues 787-841, which were omitted from the model, likely form accessory domain 2B, which was disordered and not modeled into the RecD structure. The final 152 residues of WT_311-1078 and final 6 residues of WT_311-932 were not incorporated into the model; as a result the two predicted structures were practically identical over the 424 C α positions included in the model. The putative pCU1 TraI ATPase motifs (Walker A and B) identified by sequence alignment were indeed located at the ATP binding site of RecD, and the locations of the predicted pCU1 TraI helicase DNA binding motifs corresponded well to those of RecD, as did the arginine residue of motif VI (Figure 4.5B). A schematic of the pCU1 TraI helicase in Figure 4.4C specifies the location of each putative helicase motif relative to the predicted domain structure, and allows the predicted secondary structure of the pCU1 TraI helicase to be compared to that of other helicases in Figure 4.3. Of note, the predicted secondary structure of pCU1 TraI's RecA core differs slightly from that presented in Figure 4.3. For example, the pCU1 TraI helicase domain 2A only contains four beta sheets, as compared to the five of the canonical RecA core. However, the order and arrangement of 6 of the 7 SF1 helicase motifs is maintained.

The model of the pCU1 TraI helicase was compared to the structure of the RecD2 helicase (PDB codes 3GPL, 3GP8), which has been solved in complex with DNA and with DNA and the nonhydrolyzable ATP analogue adenylyl-imidodiphosphate (ADPNP). The SF1B helicase RecD2 is a structurally similar homolog of RecD that is predicted to function as a monomer. From the structure of the helicase-DNA complex, residues responsible for mediating RecD2 translocation along the bound DNA were identified. By comparing the RecD2-DNA complex with the structure of the RecD2-DNA-ADPNP, Salkrishnan et al. were able to propose a mechanism by which RecD2

progressed through a series of conformational changes to translocate along the bound DNA strand as a result of ATP binding, hydrolysis and release^{78, 79}. In Figure 4.4C, the predicted structure of the pCU1 TraI helicase in complex with the RecD2 DNA substrate is provided, illustrating the probable location of the pCU1 helicase DNA binding site. Several putative pCU1 TraI helicase DNA binding motifs and its predicted domain 1B are highlighted.

4.4. Characterization of DNA Binding by the pCU1 TraI Helicase

4.4.1. Introduction to DNA Binding by the pCU1 TraI Helicase

While DNA binding by the conjugative relaxase has been extensively examined^{19, 32, 35, 39}, DNA binding by the conjugative helicase has only just recently attracted attention^{27, 76}. The majority of this work has focused on the helicase of the F plasmid TraI. The F plasmid TraI consists of an N-terminal relaxase, a C-terminal helicase, and a central domain that is predicted to form a traditional RecA helicase core that lacks ATPase activity. The full length F TraI protein contains one dsDNA binding site within the C-terminal helicase domain and two ssDNA binding sites, one within the relaxase and one within the central domain. However, these sites are not distinct, but instead represent three segments of one extended binding groove^{27, 76}. As discussed in Chapter 1, many conjugative systems, such as plasmids R388 and pCU1, also encode a DNA helicase C-terminal of their conjugative relaxase, as part of a multidomain protein. In the cases of plasmids R388 and pCU1, only one helicase domain has been identified within the multidomain relaxase-helicase protein, and the domain is predicted to contain only one RecA core (Figure 4.3, Figure 4.4, Table 4.1). It is unknown how a protein containing a relaxase domain and one helicase domain would integrate these two when binding DNA. We characterized the DNA binding activity of the pCU1 TraI helicase and the full length pCU1 TraI in order to compare DNA binding by pCU1 TraI to that of the larger F TraI protein. In particular, we used a combination of size-exclusion chromatography with static light scattering, electrophoretic mobility shift assays, and fluorescence anisotropy-based binding assays to probe the interaction between pCU1 TraI and its DNA substrates.

4.4.2. Construct Design and Cloning of the pCU1 TraI Helicase

The full length pCU1 TraI is 1078 residues in length. While residues 483-932 were shown to encompass the seven canonical helicase motifs (Table 4.1), the extent of the functional helicase ATPase domain was determined by activity assays, as described in the following sections. Primers were designed to isolate residues 1-1078, 1-932, 207-1078, 226-1078, 311-1078, 483-1078, and 483-932 of pCU1 TraI using the Vector NTI Express 10.0.1 suite of programs (Invitrogen, 2005) and were then commercially synthesized (IDT). PCR products were cloned into the pTYB2 vector of the IMPACT (New England Biolabs) system between *NdeI* and *SmaI* sites to generate the constructs WT_1078, WT_932, WT_207-1078, WT_226-1078, WT_311-1078, WT_483-1078 and WT_483-932. Due to the sequence of pTYB2, the proteins were thus C-terminally fused to an intein and a chitin binding domain (CBD) affinity tag. All mutant constructs were generated by QuikChange site directed mutagenesis (Stratagene) of WT_1078. All cloning and mutagenesis was verified by sequencing at the UNC-CH Genome Analysis Facility. All sequence data were analyzed, organized and stored in the Database Explorer 2.0, a component of Vector NTI Advance 10.0.01 (Invitrogen, 2005).

4.4.3 Expression and Purification of the pCU1 TraI Helicase

For both wild type and mutant constructs, 1 L TB broth was inoculated with 10 mL of a saturated overnight culture of *Escherichia coli* BL21(DE3) containing the expression plasmid. Cells were grown under antibiotic selection (100 µg/mL ampicillin) at 37°C with vigorous shaking until the cell density reached an OD₆₀₀ of 0.6. IPTG was then added to a final concentration of 0.1 mM and the temperature was dropped to 18°C. The protein was overexpressed at 18°C for 16 h, after which the cells were harvested by centrifugation (15 min at 4,500 rpm at 4°C) and resuspended in Buffer C (500 mM NaCl, 20 mM Tris-HCl pH 7.5, 10% glycerol, 5 mM EDTA, 0.01% azide) at a ratio of 10 mL buffer per 1 L of original culture. The cells were then frozen and stored at -80°C.

Before purification, 50 mL of resuspended cells were thawed in the presence of lysozyme, DNase (50 µg/mL) and a cocktail of protease inhibitors (Roche, 1 tablet per 50 mL cells) and lysed on ice using a Sonic Dismembrator Model 500 (Fisher Scientific) (2 min of 0.5 s pulses at 60%

intensity). The soluble fraction of the lysis was isolated by centrifugation (70 min at 17,000 rpm at 4°C). All subsequent purification steps were performed at 4°C. The fusion protein was purified on chitin resin (NEB) using a batch bind method followed by an extended wash step (2 x 400 mL) over a gravity column. Buffer C was used to wash the chitin resin during both batch bind and wash steps. The intein and CBD tags were removed by incubating the chitin resin in Buffer C with 50 mM D,L-dithiothreitol in Buffer C for 16 h to induce self-cleavage by the intein and consequent release of the protein from the intein and CBD. Protein was then eluted in Buffer C from the chitin resin by gravity column chromatography.

A final size exclusion chromatography purification step was then performed. Protein was concentrated to ~ 7 mg/mL, and 5 mL of protein was loaded onto a HiLoad 16/60 Superdex 200 column (GE Healthcare) pre-equilibrated with Buffer S (500 mM NaCl, 20 mM Tris-HCl pH 7.5, 5% glycerol, 5 mM EDTA, 0.01% azide) on an ATKExpress FPLC (GE Healthcare). Protein was eluted from the column in Buffer S in 2 mL fractions at a 1.2 mL/min flow rate. Protein was concentrated to ~ 50 μ M (5 mg/mL), and a final dialysis into Buffer D (2 x 1 h dialyses of 5mL protein again 1 L Buffer D: 100 mM NaCl, 20 mM Tris-HCl pH 7.5, 5% glycerol, 0.01% azide) was performed to decrease the salt concentration and remove EDTA from the sample. The protein was immediately flash frozen in liquid nitrogen and stored in 60 μ L aliquots at -80°C. All protein samples were separated by SDS-PAGE, stained with Coomassie Brilliant Blue, de-stained, visualized and found to be > 95% pure. All water was obtained from the laboratory Barnstead E-pure water filtration system at > 17 megohm (ddH₂O).

4.4.4. Effect of DNA on the Oligomeric State of pCU1 TraI

In order to initially probe the oligomerization state of the full length pCU1 TraI (WT_1078) in the presence and absence of DNA, WT_1078 was separated by size exclusion chromatography over a Wyeth silica-based column (model # WTC-030s5) attached to an Amersham Biosciences AKTA FPLC in the presence and absence of a DNA substrate (35oriT-hairpin). In particular, WT_1078 was diluted to 20.8 μ M (2.5 mg/mL) in Buffer D. Three samples of WT_1078, containing

either 0 μM , 20.8 μM , or 41.6 μM 35oriT-hairpin, were analyzed. The molecular weight of each eluting peak was then estimated by static light scattering using a Wyatt DAWN EOS light scattering instrument interfaced to a Wyatt Optilab refractometer, and Wyatt dynamic light scattering module (Table 4.2). In the absence of DNA, one peak was observed, with an estimated molecular weight (MW) of 115 kD (MW of one WT_1078 monomer = 120 kD). Upon the addition of an increasing concentration of DNA, an additional, earlier-eluting peak was observed, with an estimated molecular weight twice that of the initial peak. At TraI to DNA ratios of 1:0.5, 1:1, and 1:2, 20-30% of the total mass of the material injected eluted in the earlier dimer peak, and 70-80% eluted in the original monomer peak (Table 4.2).

These results indicated that alone in solution WT_1078 exists as a monomer, and that in the presence of DNA, a fraction of the TraI population dimerizes. The sensitivity of this technique is insufficient to determine the number of DNA molecules bound per TraI molecule in each peak, and free DNA (if present) eluted in the void volume of the column; therefore, the number of DNA molecules present in each peak was undefined. However, since the ratio of monomer peak size to dimer peak size remained stable as DNA concentration increased from 1TraI:0.5 DNA to 1TraI:2 DNA, it is likely that TraI binds, at most, one molar equivalent of DNA, and it is possible that multiple TraI molecules may associate with one DNA molecule.

The oligomerization state of pCU1 TraI was further examined by following the change in electrophoretic mobility of WT_1078 in the presence or absence of the DNA substrate, FAM-35oriT-hairpin. In general, each 10 μL sample contained 0.6 mg/mL purified WT_1078 and an increasing concentration of 5' fluorescein-labeled DNA substrate in 50 mM NaCl, 17.8 mM Tris-HCl pH 7.5, 3.9% glycerol, 0.25 mg/mL BSA, and 3.9 mM EDTA. Each sample was incubated at 37°C for 20 min, and then separated by polyacrylamide gel electrophoresis (PAGE) under non-denaturing conditions (resolving gel: 1X bis-Tris gel buffer (0.375 M Tris pH 8.8), 6% acrylamide, 0.075% APS, 0.075% TEMED; stacking gel: 1X bis-Tris gel buffer (0.125 M Tris, pH 6.8), 3.75% acrylamide, 0.125% APS, 0.125% TEMED; gels were pre-run with the running buffer (25 mM Tris base, 200 mM

glycine, pH 8.3) prior to loading sample). Gels were first visualized using a VersaDoc Imaging System, 4400 MP (BioRad) and the accompanying Quantity One software (BioRad) to determine the location of DNA-containing bands. Gels were then stained with Coomassie Brilliant Blue, de-stained, and visualized to determine the location of protein-containing bands.

Upon the addition of an equimolar amount of DNA to WT_1078, the mobility of the protein sample was shifted relative to that of the protein alone. A band corresponding to the unbound protein was not detected on the gel, though a faint band corresponding to free DNA could be detected. When an excess of DNA (1 protein to 2 DNA) was added to WT_1078, no unbound protein was observed and a dark band corresponding to free DNA was observed. These results support the conclusion that TraI binds 35oriT-hairpin with a maximum stoichiometry of 1, and 35oriT-hairpin may be able to bind two TraI molecules per DNA molecule (Table 4.2).

Together, these two techniques indicated that while pCU1 TraI exists as a monomer in solution, in the presence of DNA, two pCU1 TraI monomers can dimerize. It is unclear, however, if each monomer of the pCU1 TraI dimer binds one DNA substrate, or if two monomers are bound to one DNA substrate. It is unlikely from these data that one pCU1 TraI monomer binds to multiple DNA substrates. It was also unknown what effect DNA substrates other than 35oriT-hairpin might have on the behavior of pCU1 TraI. To attempt to answer some of these remaining questions as well as to determine the binding affinity of pCU1 TraI for a panel of DNA substrates, DNA binding by the full length pCU1 TraI and its putative helicase domain were analyzed using fluorescence anisotropy-based DNA binding assays.

4.4.5. DNA Binding Stoichiometry and Affinity of pCU1 TraI and the pCU1 TraI Helicase

The affinity of the helicase domain of pCU1 TraI (WT_311-1078) and full length pCU1 TraI (WT_1078) for a panel of DNA substrates (Table 4.3) was measured using fluorescence anisotropy-based DNA binding assays. In particular, the change in fluorescence anisotropy (FA) of a panel of 5' fluorescein-labeled DNA substrates was monitored as protein concentration was increased. The method used was similar to that described in Section 3.4.2. To generate binding curves and calculate

a dissociation constant (K_D or apparent K_D) for each experiment, normalized data were plotted as average FA vs. total protein concentration. Each data point is the average of at least 3 replicates, with error bars representing the standard error of these replicates. Curves generated during binding affinity experiments were fit to Equation 3.2 (see Section 3.4.2) or 4.1, depending on the shape of the curve and apparent binding mode that was observed.

$$\text{Equation 4.1 } f = \frac{(max)x^h}{K_{ap}^h + x^h}$$

where f , average FA signal detected; x , total protein concentration; K_{ap} , apparent K_D ; max , average FA signal of sample at a saturating concentration of protein; and h , the Hill coefficient. Note that Equation 4.1 calculates an apparent K_D , since it does not take into account protein depletion with complex formation. In contrast, equation 3.1 calculates an exact K_D . Fits were generated by nonlinear regression in Graphpad PRISM v5.03 (Graphpad, 2010).

The helicase domain WT_311-1078 bound all DNA substrates in a length-dependent manner, and with a higher affinity as compared to the relaxase domain WT_299 (Table 4.3). WT_1078 exhibited a higher DNA binding affinity as compared to either individual domain, and the shape of the binding curve observed, when [TraI] was plotted vs. FA, had a strong sigmoidal appearance, as compared to the hyperbolic appearance of the WT_299 and WT_311-1078 curves (Table 4.3, Figure 4.5). Due to the sigmoidal shape of the WT_1078 binding curves, Equation 4.1 was used to fit the data and generate an apparent K_D for the DNA-TraI interaction. Therefore, the binding affinity reported for WT_1078 is approximate and may underestimate the affinity of this pCU1 TraI construct for DNA. Both WT_311-1078 and WT_1078 failed to bind DNA substrates shorter than 15 nucleotides, and both bound FAM-15mer weakly, indicating that the minimal DNA binding site size for these two proteins is likely 15 nucleotides (nt).

The sigmoidal appearance of the WT_1078 DNA binding curve could be a result of multiple pCU1 TraI molecules binding to one DNA molecule. As the length of the DNA substrate investigated increased, the sigmoid character of the curve increased, favoring the possibility that

multiple TraI molecules were binding to one DNA molecule. However, due to the nature of the FA-based technique, it is difficult to unambiguously decipher the meaning of sigmoidal binding curves. Therefore, a series of DNA binding density experiments were performed, by which the stoichiometry of DNA binding by pCU1 TraI could be determined.

Binding density experiments were performed as described in detail by Lohman et al.⁸⁰ and Jezewska et al. 2006⁸¹. For these experiments, the DNA substrate is referred to as the macromolecule, and WT_1078 is referred to as the ligand. The goal of these assays is to determine the number of ligands bound to a macromolecule. This relationship is expressed as the degree of binding, where a degree of binding of 1 would indicate one ligand is bound to one macromolecule. Correspondingly, a maximum degree of binding represents the maximum number of ligands that can bind one macromolecule under optimal ligand and macromolecule concentrations.

To determine the maximum degree of binding of TraI (the ligand) on DNA (the macromolecule), the fluorescence anisotropy (FA) of 50 nM and 300 nM 5' fluorescein (FAM)-labeled DNA macromolecules was monitored as a function of increasing WT_1078 (ligand) concentration. As a result, a pair of DNA binding curves was generated for each DNA macromolecule investigated (Figure 4.6). For 21 FA values falling between 10% and 80% of the maximum FA of each pair of binding curves, the corresponding ligand (TraI) concentrations (L_{T1} , L_{T2}) and macromolecule (DNA) concentrations (M_{T1} , M_{T2}) were recorded. The higher of the two macromolecule concentrations is designated experiment 1 ($M_{T1} = 300$ nM), and the lower ($M_{T2} = 50$ nM) is experiment 2. These pairs of values (L_{T1} , M_{T1} and L_{T2} , M_{T2}) were used to calculate the average degree of binding of TraI ligands on the DNA macromolecule at each FA value, using equation 4.2,

$$\text{Equation 4.2} \quad \sum v_i = \frac{(L_T - L_F)}{M_T} = \frac{(L_{T1} - L_{T2})}{(M_{T1} - M_{T2})}$$

where Σv_i , average degree of binding at each FA value; L_T , total ligand; L_F , free ligand; M_T , total macromolecule; as defined above, for each pair of values, M_{T1} corresponded to 300 nM DNA macromolecule and M_{T2} corresponded to 50 nM DNA macromolecule.

The average degree of binding (Σv_i) was then plotted versus its corresponding FA value (Figure 4.7). Linear regression of the resulting points generated a line of best fit. Experimental points deviating significantly from a straight line (those corresponding to high TraI concentration) were not used when generating the line of best fit, as these were likely to introduce error into subsequent calculations. This line of best fit was then extrapolated to the maximum FA value observed for the DNA macromolecule (Figure 4.6) to determine the maximum Σv_i (Figure 4.7). Thus, a maximum Σv_i was estimated for each DNA substrate of interest. Maximum Σv_i was then plotted versus DNA substrate length, sequence, and substrate structure to determine the relationship between the maximum average degree of binding of TraI and its DNA substrate (Figure 4.8).

As can be seen from Figure 4.8, the maximum average degree of binding increased linearly with substrate length for all linear substrates. Substrates 14 nucleotides and shorter were not sufficiently bound by WT_1078 to generate a binding curve that could be accurately analyzed (Figure 4.6). This finding correlated with binding affinity data, which illustrated that WT_1078 bound FAM-15mer, but not FAM-10mer. For substrates predicted to form a hairpin, the maximum degree of binding was lower relative to that of linear substrates with the same number of nucleotides. From these data, it appears that the number of pCU1 TraI molecules bound to a DNA substrate increases linearly with DNA substrate length, and the presence of a hairpin in the substrate decreases the effective length of the substrate as perceived by pCU1 TraI. The presence of the pCU1 *oriT* sequence does not appear to affect the number of pCU1 TraI molecules bound to the DNA substrate.

4.4.6. Summary of DNA Binding by pCU1 TraI and the pCU1 TraI Helicase

The following model of DNA binding by pCU1 TraI summarizes the data presented above. Both relaxase (WT_299) and helicase (WT_311-1078) domains of pCU1 TraI bind DNA (Figure 4.5). Full length pCU1 TraI and the pCU1 TraI helicase require an estimated minimal DNA binding

site size of 15 nucleotides (nt) for binding (Figure 4.5, Figure 4.6). Since the DNA binding affinity of WT_1078 and WT_311-1078 is greater than that of WT_299, and WT_299 bound DNA substrates as short as 10 nt, the this 15 nt site size likely reflects that of a higher affinity helicase binding site. Since WT_299 bound FAM-10mer, its minimal site size is likely 10 nt or less, which is in fact the length of the DNA substrate observed bound by the F TraI relaxase (PDB code 2A0I). However, as evidenced by structures of the TrwC relaxase in complex with DNA, the relaxase can bind up to 24 nt (PDB code 2CDM). Therefore, these data indicate that together the pCU1 TraI relaxase and helicase domains could bind anywhere between 25 and 40 nt, when both binding sites are fully occupied.

However, binding density experiments indicated that 20 nt-long DNA macromolecules could bind on average 1.5 full length TraI ligands. Once the DNA reached 35 nt in length, two pCU1 TraI ligands bound one DNA molecule (Figure 4.8). Analysis by SEC-SLS and EMSA indicated that the substrate 35oriT-hairpin bound between one and two TraI molecules, though a 1:1 ratio of DNA:TraI appeared to be favored. Therefore, these data would suggest the full length TraI contains a shorter DNA binding site size, likely closer to 20 nt.

These differing site size estimations can be reconciled if it is assumed that either the individual relaxase and helicase binding sites interact in the context of the full length enzyme, or they are not fully occupied when two TraI monomers bind one DNA molecule. Under either assumption, an equilibrium would exist between free DNA molecules and TraI-DNA complexes, and for DNA substrates 20 nt or longer, a second equilibrium would exist between DNA occupied by one TraI molecule and that occupied by multiple TraI molecules. This second equilibrium is dependent on DNA length, and appears to favor the singly-bound species for DNA substrates shorter than 35 nt (Table 4.2).

It is still unclear the role protein-protein interactions play as multiple TraI molecules assemble on one DNA molecule. It has been proposed that the final 150 residues of the homologous TrwC relaxase-helicase enzyme serve as a dimerization domain^{42, 74}. However, after truncation of the corresponding residues of pCU1 TraI to generate the construct WT_932, the sigmoidal character of

DNA binding curves is maintained (Figure 4.5), indicating DNA-induced TraI dimerization may still occur. So, if protein-protein interactions form between TraI monomers, it is currently unclear where this interface is located.

4.5. Characterization of pCU1 TraI Helicase ATPase Activity

4.5.1. Introduction to pCU1 TraI Helicase ATPase Activity

The pCU1 TraI helicase translocates along DNA, driving apart the two strands of the pCU1 plasmid with the energy generated by ATP hydrolysis. Based upon the number of its helicase motifs and the amino acid sequence of each, the TraI helicase is classified as a SF1B helicase (Figure 4.1, 4.4, Table 4.1). As a result, it is expected to translocate along one strand of the pCU1 plasmid in the 5' to 3' direction, most likely functioning as a monomer and using the inchworm model of translocation^{71, 72, 77} (Figure 4.2). The conjugative helicases of plasmid R388 TrwC and the F plasmid TraI are also classified as SF1B helicases^{74, 82}. The F TraI helicase translocates along ssDNA in the 5' to 3' direction at up to 1100 base pairs (bp) per second with a step size of 6-8 bp, and can utilize any of the four NTPs. It requires a significant 5' ssDNA overhang when unwinding dsDNA substrates, and its activity increases as the length of the overhang is extended. The primary helicase domain of F TraI is found between residues ~950-1476, though a second Rec-A core is located between residues 309-950^{27, 76, 82}. The R388 TrwC helicase also translocates along ssDNA in the 5' to 3' direction, and can utilize any of the 4 NTPs, though it has a strong preference for ATP. The minimal helicase construct begins near residue 200 of TrwC and extends to the C-terminus⁷⁴.

To characterize the ATPase activity of the pCU1 TraI helicase, we incorporated pCU1 TraI into an NADH-coupled ATPase activity assay (Figure 4.9). In particular, we used this assay to identify the optimal DNA and NTP substrate for the pCU1 TraI helicase, to determine extent of the minimal helicase within pCU1 TraI, and to verify the identity of the pCU1 TraI ATPase Walker A and Walker B motifs. In subsequent work outlined in Section 4.6., we used these data to optimize a second ATPase activity assay that was easily adaptable to a high throughput screen for compounds capable of inhibiting the ATPase activity of the pCU1 TraI helicase.

Recently, multiple helicases have been targeted with inhibitors in the pursuit of therapies for diseases ranging from viral hepatitis to gram positive bacterial infections (Figure 4.1)⁸³⁻⁹³. For example, a variety of nucleotide mimics were shown to inhibit the NS3 helicase of the hepatitis C virus⁹⁴⁻⁹⁷, and a set of coumarin-based molecules was shown to target the replicative helicase DnaB of *Bacillus anthracis*⁸⁴. While the majority compounds inhibiting helicases to date are either weakly selective or lack potency, they have all identified synthetic and biological routes by which specific helicases could be targeted^{83, 85, 87-89, 91-97}, and, in fact, HSV helicase-primase inhibitors have already shown efficacy in animal models^{86, 87, 90}. Therefore, these efforts demonstrate that the helicase family of enzymes is indeed a viable therapeutic target. In addition, due to the relative paucity of research investigating helicase-specific inhibitors, these enzymes could provide a new avenue in the unending search for novel therapeutic leads in the areas of cancer and infectious diseases. In particular, we suggest that conjugative helicases such as the pCU1 TraI helicase represent potential targets for future antimicrobials. In Section 4.6, initial work to identify such compounds is described.

4.5.2. Design of an NADH-coupled ATPase Activity Assay

A coupled enzyme assay involving pyruvate kinase and lactate dehydrogenase, as described by Kiiianitsa 2003 (Figure 4.9), was optimized for characterizing the ATPase activity of the pCU1 TraI helicase⁹⁸. Coupled enzymes (pyruvate kinase (PK), lactate dehydrogenase (LDH)), phosphoenolpyruvate (PEP), nicotinamide adenine dinucleotide (NADH), and all four NTPs (ATP, TTP, CTP, GTP) were obtained from Sigma. The full length pCU1 TraI construct and pCU1 TraI helicase constructs were cloned, expressed and purified as described in Sections 4.4.2 – 4.4.3. The DNA substrates incorporated into the assay were commercially synthesized (IDT); the sequences are provided in Table 4.3. DNA substrates were resuspended in Buffer R (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.05 mM EDTA, 0.01% azide), heated to 95°C for 10 min, and then allowed to slow cool to room temperature. ATPase assay buffer (Buffer P) consisted of 25 mM NaCl, 50 mM Tris-acetate pH 7.5, 10 mM Mg(acetate), 25 mM K₂PO₄, 0.1 mg/mL BSA.

During the NADH-coupled ATPase assay, ATP is hydrolyzed to ADP by pCU1 TraI in the presence of ssDNA. ATP is regenerated from ADP by PK as it converts PEP to pyruvate. Pyruvate is then converted to lactate by LDH as it simultaneously oxidizes NADH to $\text{NAD}^+ + \text{H}$. The oxidation of NADH to $\text{NAD}^+ + \text{H}$ is detected as a decrease in absorption at 350 nm (A_{350}) since NADH absorbs at this wavelength, but NAD^+ does not. There is a 1:1 relationship between ATP hydrolysis and NADH oxidation, therefore the drop in A_{350} is directly proportional to the ATPase activity of pCU1 TraI. The drop in absorbance over time was monitored at 350 nm by a PHERAstar fluorescence plate reader (BMG Labtech) and was used to calculate the rate of ATP hydrolysis by pCU1 TraI using Equation 4.3⁹⁹:

$$\text{Equation 4.3 ATPase rate [ATP x min}^{-1}] = -\frac{dA_{350}}{dt} \left[\frac{\text{OD}}{\text{min}} \right] \times K_{path}^{-1} \times \text{moles}^{-1} \text{ ATPase}$$

where K_{path} is the molar absorption coefficient for NADH for a given optical pathlength, which was calculated to be $3965.75 \text{ OD} \cdot \text{mol}^{-1}$ for the assay volume (75 μL) in the 384-well clear-bottom assay plates (Corning) used in the assay. dA_{350}/dt was corrected for background NADH hydrolysis and NTP hydrolysis by subtracting a no protein control from all experimental data points.

4.5.3. Optimization and Application of an NADH-coupled ATPase Activity Assay

Assay conditions were systematically varied over a series of experiments to determine the concentrations of pCU1 TraI, ssDNA, NTP, NADH, PEP, PK, and LDH that would maximize pCU1 TraI ATPase activity (as determined by Equation 4.3) and allow for optimal detection of NADH oxidation. The ideal length of ssDNA and the identity of NTPs hydrolyzed by the pCU1 TraI helicase were also determined. Using these optimized conditions, the extent of the minimal pCU1 TraI helicase was established, and key residues within the Walker A and Walker B motifs were identified (Table 4.4, Figure 4.10, 4.11).

The impact of varying these experimental conditions on the ATPase activity of each pCU1 TraI helicase construct is outlined in Table 4.4. As expected, raw ATPase activity (mol ATP/min) was directly proportional to the concentration of pCU1 TraI (data not shown). The specific ATPase

activity (mol ATP/min/mol TraI) increased with the length of the ssDNA substrate used (Figure 4.10). Due to practical constraints, a 60 nt-long ssDNA molecule (60mer) was chosen as the final substrate. The ATPase activity of WT_1078 was not significantly impacted by DNA substrate concentration when the concentration of DNA was equal to or greater than that of pCU1 TraI. However, for other helicase constructs that did not bind DNA as tightly as WT_1078 (Table 4.3), ATPase activity increased with DNA concentration (Figure 4.11, Table 4.4).

When analyzed with this NADH-coupled enzyme assay, pCU1 TraI was only capable of hydrolyzing ATP, showing little or no activity above background in the presence of CTP, GTP, or TTP (Figure 4.10); and ATPase activity was actually inhibited at high ATP concentrations (Table 4.4). However, these observations are complicated by the presence of PK. PK is a metal-dependent enzyme that catalyzes the unidirectional conversion of PEP to pyruvate, but can be inhibited by high concentrations of ATP relative to ADP, and is much more active in the presence of ATP as compared to other NTPs¹⁰⁰. Therefore, while pCU1 TraI is definitely most active in the presence of ATP, this specific coupled enzyme assay is not capable of determining if it has limited activity in the presence of other NTPs.

Upon optimization, assay conditions consisted of the following in a 75 μ L reaction volume: 100 nM pCU1 TraI, 1.25 mM ATP, 100 nM 60mer ssDNA, 0.7 mM PEP, 0.3 mM NADH, 50 mM PK, 50 mM LDH. pCU1 TraI was allowed to bind to its ssDNA substrate for 10 min at 37°C prior to the addition of ATP. Following the addition of ATP, the ΔA_{350} was monitored at 37°C for 120 min. Under these conditions, the minimal extent of the pCU1 TraI helicase domain was determined to fall between residues 311 and 932 (Table 4.4, Figure 4.11). The impact of mutations within the predicted Walker A and Walker B motifs in the context of the full length enzyme (WT_1078) is illustrated in Figure 4.11. Mutation of a conserved lysine (pCU1 TraI K507) within the Walker A motif, which is expected to coordinate the gamma phosphate of the bound ATP, eliminated ATPase activity. Mutation of a conserved glutamic acid (pCU1 TraI E569) within the Walker B, which is expected to serve as a catalytic base, also eliminated ATPase activity. However, mutation of a conserved aspartic

acid (pCU1 TraI D568) within the Walker B motif, which is predicted to coordinate the bound Mg^{2+} ion, inhibited ATPase activity if changed to alanine, but had no effect on ATPase activity if changed to asparagine. These results verified the location of both Walker A and B motifs within the pCU1 TraI helicase domain. Since the D568N mutant had little or no effect on pCU1 TraI ATPase activity, an asparagine appears capable of effectively coordinating the bound Mg^{2+} within the ATPase active site.

4.6. Design, Optimization, and Validation of a High Throughput Assay of pCU1 TraI Helicase ATPase Activity

4.6.1. ADP Quest Assay Design

After characterizing the ATPase activity of pCU1 TraI using the NADH-coupled ATPase activity assay, a second ATPase activity assay, ADP Quest¹⁰¹, was optimized for use with pCU1 TraI. This assay also uses a series of coupled enzymes to convert ATP hydrolysis by pCU1 TraI to a final detectable product; however, it provides the advantage of incorporating into its final step the conversion of Amplex Red to resorufin, a fluorescent compound which can then be detected (Figure 4.12). In general, fluorescence detection provides greater sensitivity and a wider dynamic range as compared to detection by absorption. Therefore, the ADP Quest assay required smaller reaction volumes and significantly lower reagent concentrations as compared to the NADH-coupled ATPase activity assay.

The ADP Quest assay design is described by Charter et al., 2006 and is outlined in Figure 4.12¹⁰¹. The initial step is identical to that of the NADH-coupled assay described in Section 4.5.1. Following ATP hydrolysis to ADP by pCU1 TraI, pyruvate kinase (PK) regenerates ATP from ADP while simultaneously converting PEP to pyruvate, PO_4 , and O_2 . Pyruvate is then broken down to acetyl phosphate, CO_2 , and H_2O_2 by pyruvate oxidase. Horseradish peroxidase uses the oxidizing power of H_2O_2 to convert Amplex Red, which is not fluorescent, to the fluorescent compound resorufin. Following excitation at 530 nm, emission by resorufin at 590 nm was detected using a PHERAstar plate reader (BMG Labtech), an Envision 2103 Multilabel Reader (Perkin Elmer), or an

Aquest plate reader (LJL BioSystems). Both the Envision and Aquest readers are located in the Center for Integrative Chemical Biology and Drug Discovery (CICBDD) at the University of North Carolina at Chapel Hill. Many of the experimental steps described below were performed at the CICBDD. A standard curve of nine known concentrations of ADP was used to determine the relationship between the fluorescence of each sample and the concentration of ADP present.

Reagents A and B were obtained from DiscoverRx Corp. (Fremont, CA). Reagent A contains 50 mM phosphate buffer, pH 7.0, 250 μ M Amplex Red, and 1 U/mL catalase. Reagent B contains 50 mM phosphate buffer, pH 7.0, 2.5 mM $MgCl_2$, 25 μ M flavine adenine dinucleotide, 250 μ M thiamine pyrophosphate, 250 μ M PEP, 25 U/mL horseradish peroxidase, 50 U/mL pyruvate kinase, 15 U/mL pyruvate oxidase, and 1 U/mL catalase. ATP was obtained from Sigma. Full length pCU1 TraI (WT_1078) was cloned, expressed and purified as described in Sections 4.4.2 – 4.4.3. The 60mer ssDNA substrate incorporated into the assay was commercially synthesized (IDT); the sequence is provided in Table 4.3. DNA was resuspended in Buffer R (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.05 mM EDTA, 0.01% azide), heated to 95°C for 10 min, and then allowed to slow cool to room temperature. ATPase assay buffer (Buffer ADP) consisted of 20 mM NaCl, 15 mM HEPES pH 7.5, 0.3 mM $MgCl_2$ (once optimized), and 0.1 mg/mL BSA. Experiments were performed initially in 384-well black assay plates (Corning) and then, upon further miniaturization of the assay, black 384-well small volume plates were used (Greiner).

4.6.2. Results of ADP Quest Assay Optimization and Validation

The ADP Quest assay was optimized for use with pCU1 TraI WT_1078 with goal of identifying reaction conditions that would support efficient pCU1 TraI ATPase activity and allow for straight forward incorporation of the assay into a high throughput screen (HTS) for ATPase inhibitors. Upon miniaturization of the assay to a reaction volume of 21.5 μ L, the optimized assay conditions consisted of the following: 4 nM pCU1 TraI (in buffer ADP), 25 nM DNA substrate (in buffer ADP), 10 μ M ATP (in buffer ADP; ATP K_M = 10-12 μ M), 4 μ L Reagent A, 2.5 μ L Reagent B, and a 10 mM $MgCl_2$ quench (in buffer ADP). A pCU1 TraI concentration of 4 nM generated signal

significantly above that of the background signal generated by Reagents A and B, and was also sufficiently low such that a linear increase in signal due to TraI ATPase activity was observed for >60 min. Decreasing the volumes of Reagents A and B from that suggested by the manufacturer improved the linearity of the ADP standard curve (4 μL of Reagent A was used as compared to the suggested 7.5 μL ; 2.5 μL of Reagent B as compared to the suggested 15 μL). EDTA was initially used to quench the TraI ATPase activity. However, EDTA was found to also inhibit pyruvate kinase and pyruvate oxidase. Therefore, 10 mM MgCl_2 was chosen as the assay quench, since 4 nM pCU1 TraI was inhibited by 10 mM MgCl_2 , but the three coupled enzymes in Reagent B were still fully functional under these conditions.

Since the compounds which pCU1 TraI would be screened against were stored in DMSO, a DMSO tolerance test was performed. pCU1 TraI ATPase activity was unaffected by a DMSO concentration $\leq 5\%$. Finally, the order of addition of components was optimized to generate the following protocol:

1. 2 μL 6.7% DMSO, mimicking the addition of compound stored in DMSO
2. 8 μL 6.5 nM TraI stock
3. 3 μL 108.3 nM DNA + 43.3 nM ATP (START solution)
4. Adhesive seal applied to each plate
5. Plate incubated at 37°C for 60 min
6. 2 μL 104 mM MgCl_2 quench
7. Plate cooled at 20°C for 30 min
8. 4.5 μL Reagent A
9. 2 μL Reagent B
10. Plate incubated at 20°C for 60 min
11. Fluorescent signal read by plate reader

Following steps 3, 5, and 8, each plate was spun at 1000 rpm for 25 s.

Following optimization, the assay was validated to establish the HIGH (maximum enzyme activity) values and LOW (EDTA-quenched enzyme) values of the assay, as well as to determine how well separated these values were. Z' factor analysis was used to determine the degree of separation of LOW and HIGH values, as well as the amount of error associated with each. The Z' factor is defined by Equation 4.4:

$$\text{Equation 4.4 } Z' \text{ factor} = 1 - \frac{3(\sigma_H + \sigma_L)}{\mu_H - \mu_L}$$

where, σ_H is the standard deviation of the HIGH controls, σ_L is the standard deviation of the LOW controls, μ_H is the mean of the HIGH controls, and μ_L is the mean of the LOW controls¹⁰². Therefore, during validation, 2x384-well plates of LOW controls and 2x384-well plates of HIGH controls were prepared on successive days until the calculated Z' factors exceeded 0.5 to 0.6 on 3 days (Figure 4.13A). LOW wells contained pCU1 TraI inhibited with 10 mM MgCl₂ prior to addition of the START solution, and HIGH wells contained pCU1 TraI quenched with 10 mM MgCl₂ after 60 min of ATPase activity. For LOW controls, the above protocol was adjusted such that 2 μ L 104 mM MgCl₂ quench was added to each well prior to the addition of the START solution. During subsequent high throughput screens, if plates failed to meet the established HIGH, LOW, and Z' factor standards, the screen would be repeated. During validation, all plates were prepared using a Multidrop Combi with a small tube plastic tip dispensing cassette (Thermo Scientific).

After the initial validation standard was achieved on three separate days, a round of HTS validation was performed. This final validation process mimicked as closely as possible the subsequent high throughput screen (HTS) protocol, minus the addition of compounds into the 384-well assay plates. Therefore, it incorporated the Nanoscreen NSX-1536 384 drop setter (Multimek), controlled by the Nanoscreen ALHS software, into the first step of the established protocol. This drop setter would be used in subsequent HTS assays to add compound into each plate, though during HTS validation it simply added DMSO to each plate. During HTS validation, the plate layout that would be used during high throughput screening of compounds was also adopted. In this plate

design, columns 1 and 2 were LOW controls, columns 23 and 24 were HIGH controls, and columns 3 to 22 would contain the compounds to be screened. During HTS validation, columns 3 to 22 contained a final concentration of 1.03% DMSO but no compound. Three plates were prepared according to this design, HTS validation was performed, and upon calculating the Z' factors for each, and all three plates exceeded the Z' factor standard of 0.5 or greater (Figure 4.13B).

4.7. A High Throughput Screen for Inhibitors of the pCU1 TraI Helicase ATPase Activity

Using the validated assay design as described in Section 4.6, pCU1 TraI helicase ATPase activity was screened against two libraries of compounds, the Library of Pharmaceutically Active Compounds (LOPAC, 1280 compounds) and the KINASE screen (~5000 compounds). The KINASE screen consists of compounds predicted to be active against kinases. The LOPAC library was analyzed in duplicate to verify the reproducibility of hits identified by the validated assay (Figure 4.14). The LOPAC library was also counter-screened against the ADP Quest assay components in the absence of pCU1 TraI in order to identify compounds inhibiting the activity of one or more of the coupled enzymes, as opposed to inhibiting the activity of the TraI ATPase. All compounds in both the LOPAC and KINASE screens initially found to inhibit the pCU1 TraI ATPase activity were then eliminated if they also were found to inhibit the coupled enzymes in the ADP Quest assay or a kinase that had been previously characterized and screened at the CICBDD.

After eliminating false positives using this counter-screening protocol, the top 100 compounds, as measured by percent inhibition of pCU1 TraI were identified. These 100 were then analyzed for structural and chemical similarity. Thirty one compounds were chosen that both followed a chemical series and significantly inhibited the activity of pCU1 TraI relative to other kinases and the ADP Quest components. These 31 were then further analyzed to determine the concentration of compound required to achieve 50% inhibition of the TraI ATPase. In particular, 10 concentrations of each compound (30 μM – 0.15 nM) were tested for their ability to inhibit the TraI ATPase, as well as the ADP Quest components alone. Using this protocol, 6 compounds were identified as false positives, as they did not inhibit the pCU1 TraI ATPase at any of the concentrations

used. Two compounds significantly inhibited the ADP Quest components, as compared to pCU1 TraI ATPase. Of the remaining 23 compounds, 9 compounds had an IC₅₀ (concentration achieving 50% inhibition) against pCU1 TraI ATPase activity that was 5-fold higher than that required to inhibit 50% of the activity of the ADP Quest reagents. The structures of these compounds are shown in Figure 4.15. Many of the compounds resemble nucleotide analogues and contain two multi-ring systems linked by amide or amine groups. Several also contain sulfoxide and ether groups. Their overall structure is reminiscent of inhibitors identified against other helicases (see Section 4.5.1).

4.8. Future Steps for the Identification and Characterization of pCU1 TraI ATPase Inhibitors

The screening process described in Section 4.7 identified 23 compounds capable of selectively inhibiting the ATPase activity of pCU1 TraI. These compounds will be analyzed next using a secondary dsDNA strand separation assay. During this assay, pCU1 TraI will separate two annealed ssDNA substrates, one of which is fluorescently or radioactively labeled. Since the mobility of the labeled ssDNA strand is greater than that of the annealed substrate in a gel matrix, separation of the labeled strand from the unlabeled strand will be detected by analyzing the products of the reaction by gel electrophoresis. Each of the 23 compounds identified in Section 4.7 will be tested for its ability to inhibit the strand separation activity of pCU1 TraI. The structures of compounds showing efficacy in this secondary assay will be compared to those in a larger library of 100,000 compounds available at the CICBDD. Compounds in this larger library that have structures similar to those showing efficacy in the DNA strand separation assay will then be screened against the ATPase activity of pCU1 TraI. This cycle of primary assay (ATPase), secondary assay (DNA strand separation), and identification of additional compounds will be continued in order to identify compounds with the ability to inhibit the ATPase and DNA strand separation activities of pCU1 TraI at nanomolar concentrations. As necessary, the synthetic group at the CICBDD can generate additional compounds based upon the structures of effective inhibitors identified within the compound libraries.

4.9. Tables and Figures

Table 4.1 Conserved SF1/SF2 Helicase Motifs

Table 4.2 pCU1 TraI Oligomerization and DNA Binding

Table 4.3 DNA Substrates for DNA Binding and ATPase Activity Assays

Table 4.4 Impact of NADH-coupled ATPase Activity Assay Conditions on pCU1 TraI ATPase Activity

Figure 4.1 Classification of Helicases

Figure 4.2 Four Models of Helicase DNA Unwinding Mechanisms (from Matson 2003)

Figure 4.3 Structures of Representative Helicases and the Conserved RecA Core (from Caruthers 2002)

Figure 4.4 Predicted Structure of the pCU1 TraI Helicase

Figure 4.5 DNA Binding Curves

Figure 4.6 Representative DNA Binding Density Curves

Figure 4.7 Average Degree of Binding as a Function of Fluorescence Anisotropy

Figure 4.8 Maximum Average Degree of Binding as a Function of Substrate Length, Structure

Figure 4.9 Design of the NADH-coupled ATPase Activity Assay

Figure 4.10 pCU1 TraI ATPase Activity as a Function of ssDNA Length, NTP Concentration

Figure 4.11 Impact of pCU1 TraI Construct Length and Mutagenesis of the Walker A and B Motifs on pCU1 TraI ATPase Activity

Figure 4.12 Design of the ADP Quest ATPase Activity Assay

Figure 4.13 Validation of the ADP Quest ATPase Activity Assay Design

Figure 4.14 High Throughput Screening of the LOPAC against pCU1 TraI ATPase Activity

Figure 4.15 Putative Inhibitors of the pCU1 TraI ATPase Identified through High Throughput Screening

Motif	I (Walker A)	Ia	II (Walker B)	III	IV	V	VI
Canonical SF1/2 Sequence	++xGxAGoG <u>K</u> S/T	xx+xxxoo	+++ <u>D</u> E <u>x</u> o	++++GDxoQ	xx+xooxR	xx <u>I</u> +xxxoG+o+oox	VA+ <u>T</u> R <u>x</u> oo
pCU1 Sequence	AAHGYAGTG <u>K</u> S	PY <u>G</u> TQKKAL	VFI <u>D</u> EAG	AVFLGDTSQKTAV <u>E</u> A	ATLIISGTNASR	AT <u>V</u> <u>H</u> <u>K</u> SQGLTCDRV	YVGIS <u>R</u> ARHE
	498-508	530-538	565-571	592-606	683-694	854-868	885-894

Table 4.1 Conserved SF1/SF2 Helicase Motifs

The seven conserved helicase motifs of SF1 and SF2 helicases are listed in the top row, with the canonical sequence of each provided in the second row. The putative pCU1 sequence of each motif is provided in the third row, and below each are the corresponding residue numbers. Key residues within each motif are in bold and underlined. The first four motifs (blue background) are located within domain 1A. The last three motifs (red background) are located within domain 2A.

SEC-SLS Data				EMSA Data		
Exp	Peak	MW	% mass	Protein- only band	DNA-only band	Protein+ DNA band
WT_1078	1	115	96.8	<input checked="" type="checkbox"/>		
WT_1078 + DNA (1:0.25)	1	123	98.4			
	2		1.6			
WT_1078 + DNA (1:0.5)	1	121.5	74.6			<input checked="" type="checkbox"/>
	2	251.3	25.4			
WT_1078 + DNA (1:1)	1	107	71		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	2	189	29			
WT_1078 + DNA (1:2)	1	112	75		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	2	245	25			

Table 4.2 pCU1 TraI Oligomerization and DNA Binding

The molar ratio of WT_1078 to DNA is provided in parentheses for each experiment. For each SEC-SLS experiment, the number of peaks observed, as well as the estimated molecular weight (MW) and percent of total mass eluting in each peak is provided. WT_1078 + DNA (1:0.25 molar ratio) eluted as two peaks, but the second peak was too small to accurately determine an estimated MW. For each EMSA sample, the presence of a protein-only band (observed only by Coomassie stain), a DNA-only band (observed only by fluorescence detection), and a protein+DNA band (observed by both Coomassie stain and fluorescence detection) is indicated. WT_1078 + DNA (1:1 molar ratio) generated a very faint DNA-only band.

Substrate	WT_299										WT_311-1078										WT_1078										WT_932									
	K _o (nM) 95% CI										K _o (nM) 95% CI										App. K _o (nM) 95% CI										App. K _o (nM) 95% CI									
FAM-42orf1-no_hairpin	A A T T C A G T G T A T T A G T A T G A C T A T C T C G C T G C G T T A G G T G T																																							
FAM-357orf1-hairpin	T G T G A T A G C G T G A T T A T C G C G C T G C G T T A G G T G T A T A G C A G																																							
FAM-357orf1-no_hairpin	T G T A T T A G T A T G A C T A T C T C G C T G C G T T A G G T G T A T A A T A G																																							
FAM-35orf1-hairpin	T G T G A T A G C G T G A T T A T C G C G C T G C G T T A G G T G T																																							
FAM-20orf1-half_hairpin	T A T C G C G C T G C G T T A G G T G T																																							
FAM-42mer-hairpin	C T A G C T C C G A G C A T A A G A G C I C G G A C T A C G T G A T C T A C T A T A																																							
FAM-42mer-no_hairpin	A C A G A T G C G T G C G A T A G A C T C A T A T A G C G G T G A T A C T G G A G T																																							
FAM-35mer-hairpin	C T A G C T C G A G C A T A A G A G C T C G G A C T A C G T G A T C																																							
FAM-35mer-no_hairpin	T G C G T G C G T A G T G T C T A T A G C G G A G A T C C T G G A G T																																							
FAM-29mer	C C A A A C G A G C C A G C G A G C G A G C G A A C G C G																																							
FAM-24mer	C G A G C A G C G A G C G A G C G A A C G C G																																							
FAM-20mer	C G A G T A C C G G T T A C C C T A G A																																							
FAM-19mer	C A G C G A G C G A G C G A A C G C G																																							
FAM-18mer	C G A G T A C C G G T T A C C C T A																																							
FAM-16mer	C G A G T A C C G G T A C C C																																							
FAM-15mer	C G A G T A C C G G T T A C C																																							
FAM-10mer	A G C G A A C G C G																																							
FAM-35:35:46:DNA	G A G T C G G C A T C C G G A T C T A G G G T A A C C G G T A C T C G										A G C G A A C G C G										G A G T A C C G G T T A C C C T A G										G A G T A C C G G T T A C C C T A G									
FAM-37:23:45:DNA	C T C A G C C G T A G C C C T A G A T C C C A T T G C C A T G A G C										C T C A G C C G T A G C C C T A G A T C C C A T T G C C A T G A G C										C T C A G C C G T A G C C C T A G A T C C C A T T G C C A T G A G C										C T C A G C C G T A G C C C T A G A T C C C A T T G C C A T G A G C									
FAM-45:23:46:DNA	T C G A T C T A G C A T C C G G A T C T A G G G T A A C C G G T A C T C G										C C T A G A T C C C A T T G C C A T G A G C										C C T A G A T C C C A T T G C C A T G A G C										C C T A G A T C C C A T T G C C A T G A G C									
FAM-60:23:45:DNA	G C G A A C T G T C G A G T C G G C A T C C G G A T C T A G G G T A A C C G G T A C T C G										C C T A G A T C C C A T T G C C A T G A G C										C C T A G A T C C C A T T G C C A T G A G C										C C T A G A T C C C A T T G C C A T G A G C									
FAM-60:23:45:DNA	G C G A T A C G T A C T G T C G A T C C A T G T C G A T C T A G C A T C C G G A T C T A G G G T A A C C G G T A C T C G										C C T A G A T C C C A T T G C C A T G A G C										C C T A G A T C C C A T T G C C A T G A G C										C C T A G A T C C C A T T G C C A T G A G C									

Table 4.3 DNA Substrates for DNA Binding and ATPase Activity Assays

For each substrate, the sequence and secondary structure are provided. Names of substrates containing a wild type *nic* site and minimal *oriT* include the phrase “oriT.” Names of mutant substrates include the phrase “mer.” “FAM” indicates that the substrates were fluorescently-labeled on the 5' end with the fluorophore fluorescein. Double-stranded substrates were labeled only on the shorter of the two strands. Mutated bases of *oriT*-containing substrates are in red font. Sequences containing an inverted repeat that are predicted to form a hairpin are boxed. The location of the pCU1 *nic* site is indicated by a bold red vertical line, where applicable. Dissociation constants (K_D or apparent K_D) as measured by fluorescence anisotropy (FA)-based binding assays are provided for the indicated pCU1 TraI construct and substrate. WT_1078 and WT_932 binding affinity data are given as the apparent K_D (App. K_D). Due to the sigmoidal character of these binding curves, they were fit with Equation 4.1, which generates an apparent K_D since it does not take into account protein depletion with complex formation. The 37mer, 45mer, and 60mer ssDNA substrates used during ATPase assays were identical in sequence to the top (longer) strand of the dsDNA substrates listed in the last three rows of the table. The 23mer ssDNA substrate used in ATPase assays was identical in sequence to the bottom (shorter) strand of these dsDNA substrates.

TraI Construct	1.25 mM ATP				2.5 mM ATP			
	100 nM 60mer		500 nM 60mer		100 nM 60mer		500 nM 60mer	
WT_1078	19.58	(4.45)	20.49	(4.26)	15.89	(5.39)	17.88	(6.73)
WT_932	11.12	(2.60)	12.75	(2.16)	7.66	(4.60)	9.97	(5.44)
WT_207-1078	5.01	(0.70)	11.88	(3.17)	9.27	(1.71)	15.76	(4.37)
WT_226-1078	4.33	(1.17)	11.06	(2.35)	10.12	(2.42)	18.49	(2.81)
WT_311-1078	4.64	(2.42)	12.16	(4.07)	13.49	(3.95)	24.87	(1.01)
WT_483-1078	0.15	(2.03)	1.51	(2.05)	-0.49	(0.45)	-0.19	(0.58)
WT_483-932	0.99	(0.21)	2.51	(1.29)	0.36	(1.14)	0.45	(0.58)
WT_299	1.68	(1.56)	2.63	(1.72)	1.75	(1.23)	1.70	(0.89)

Table 4.4 Impact of NADH-coupled ATPase Activity Assay Conditions on pCU1 TraI ATPase Activity

For each pCU1 TraI construct listed, the impact of ATP concentration and ssDNA concentration is shown. Values listed are the average specific activity of each construct (mol ATP/min/mol TraI), as determined by three separate experiments, each performed in triplicate. Values in parenthesis represent the standard deviation of the average specific activity.

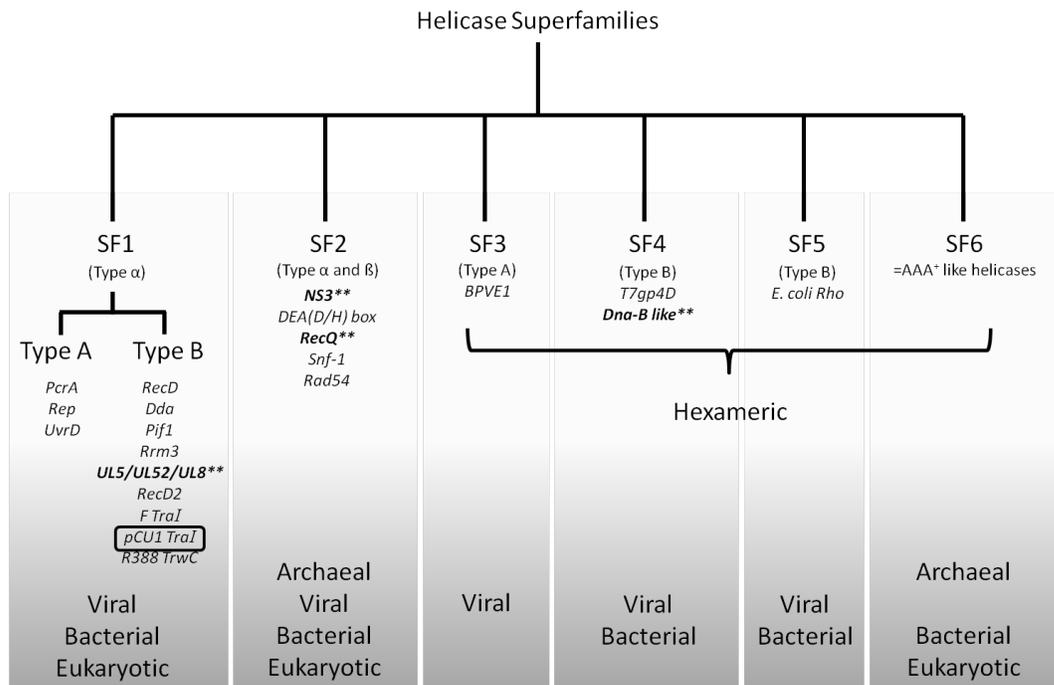


Figure 4.1 Classification of Helicases

The six superfamilies of helicases are listed, along with representative members of each superfamily and the class(es) of organism(s) containing members within each superfamily. Starred and bolded helicases have been targeted by helicase-specific inhibitors. In parentheses is provided the specific type (A or B; α or β) of helicase found within each class, as applicable. SF6 helicases are also referred to as AAA⁺ like helicases.

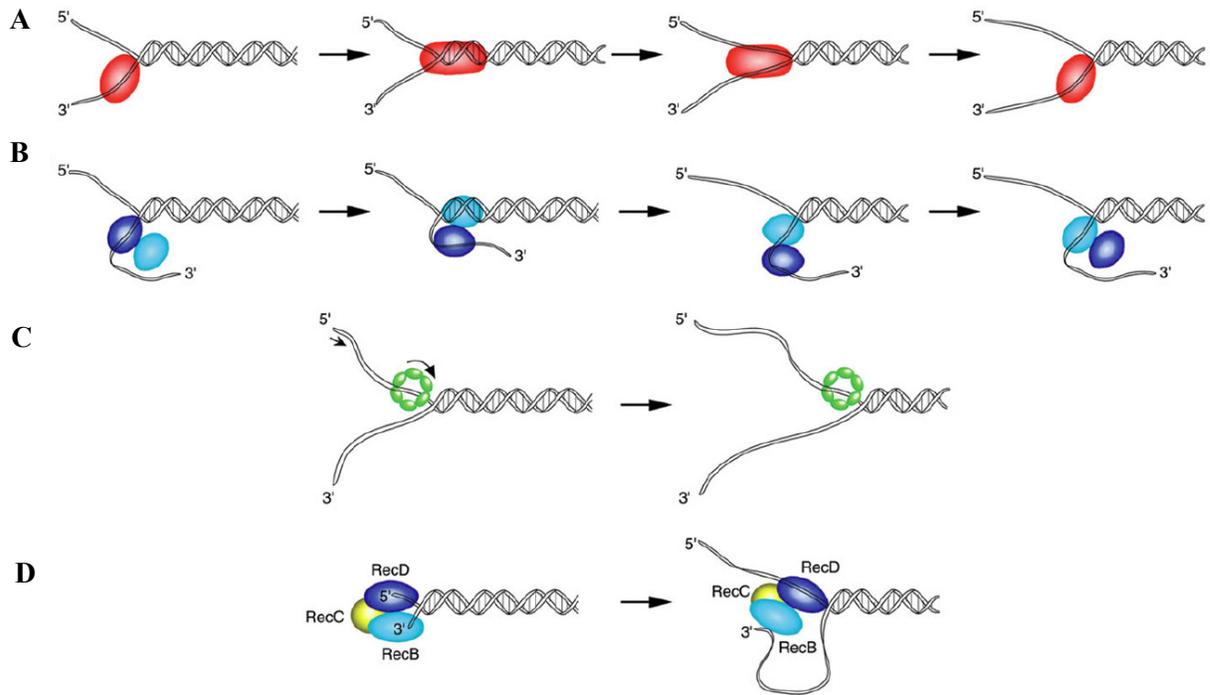


Figure 4.2 Four Models of Helicase DNA Unwinding Mechanisms (from Matson, 2003)

Models depicting four unwinding reaction mechanisms for DNA helicases. A) The inchworm mechanism for a monomeric helicase. The enzyme cycles through energetic and conformational states as it hydrolyzes NTP to couple unidirectional DNA translocation with duplex destabilization. B) The functional dimer model utilizes a dimeric helicase composed of two identical protomers. The protomers alternate in binding to ssDNA and duplex DNA as the protein translocates along and unwinds duplex DNA. C) A hexameric helicase, with ssDNA moving through the central channel of the hexamer, utilizes NTP hydrolysis to fuel unidirectional translocation along one strand while displacing the other strand to the outside of the enzyme. D) The RecBCD complex utilizes two active helicases, one moving along one strand in the 5'→3' direction and the other moving along the other strand in a 3'→5' direction. The two helicase motors have opposite unwinding polarities but move in the same overall direction as they translocate along the antiparallel strands of duplex DNA. The RecBCD helicase unwinds DNA at a faster rate than the RecB helicase resulting in the formation of the 'loop-tails' intermediate shown. (From Matson, 2003⁷³)

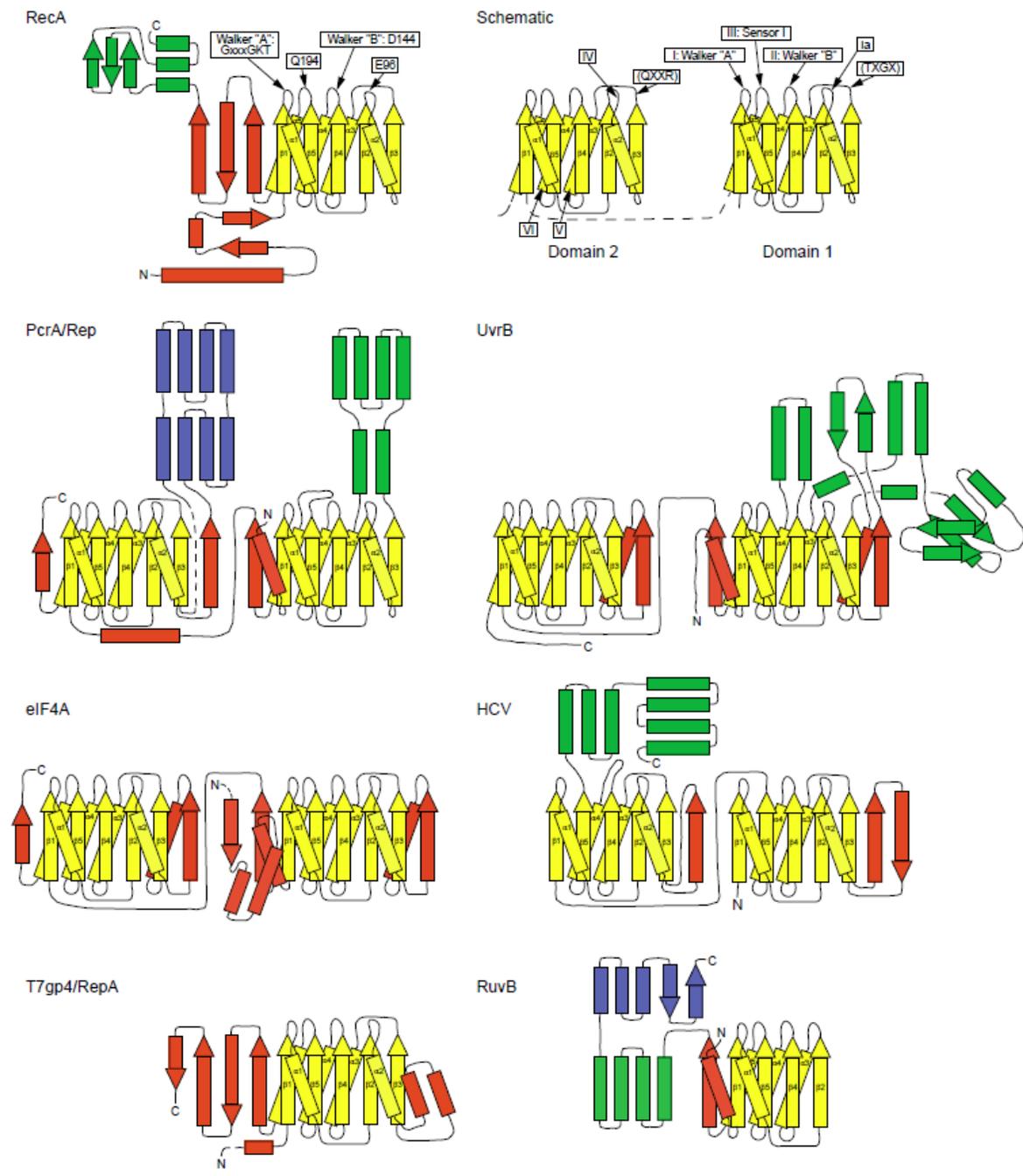
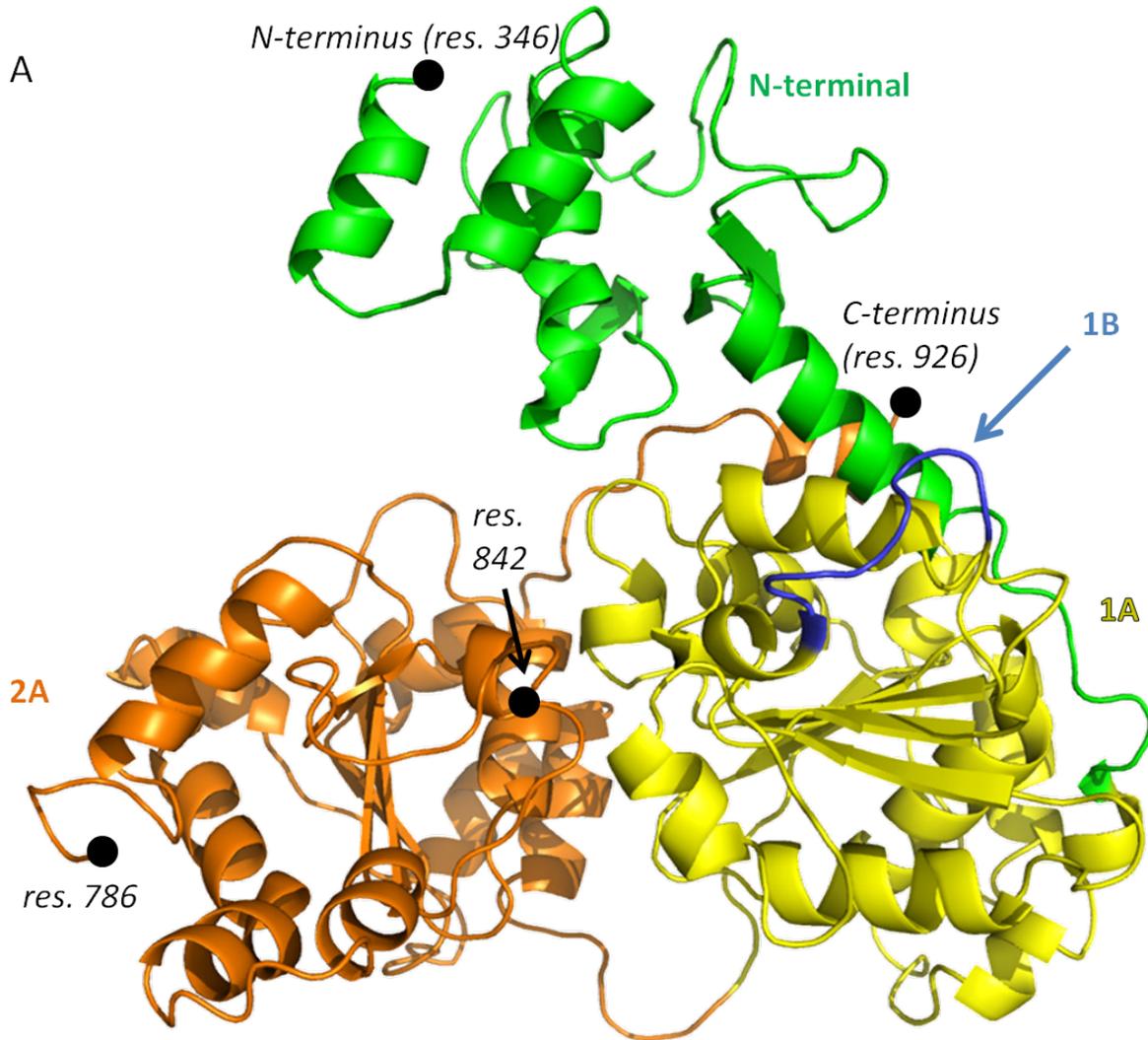
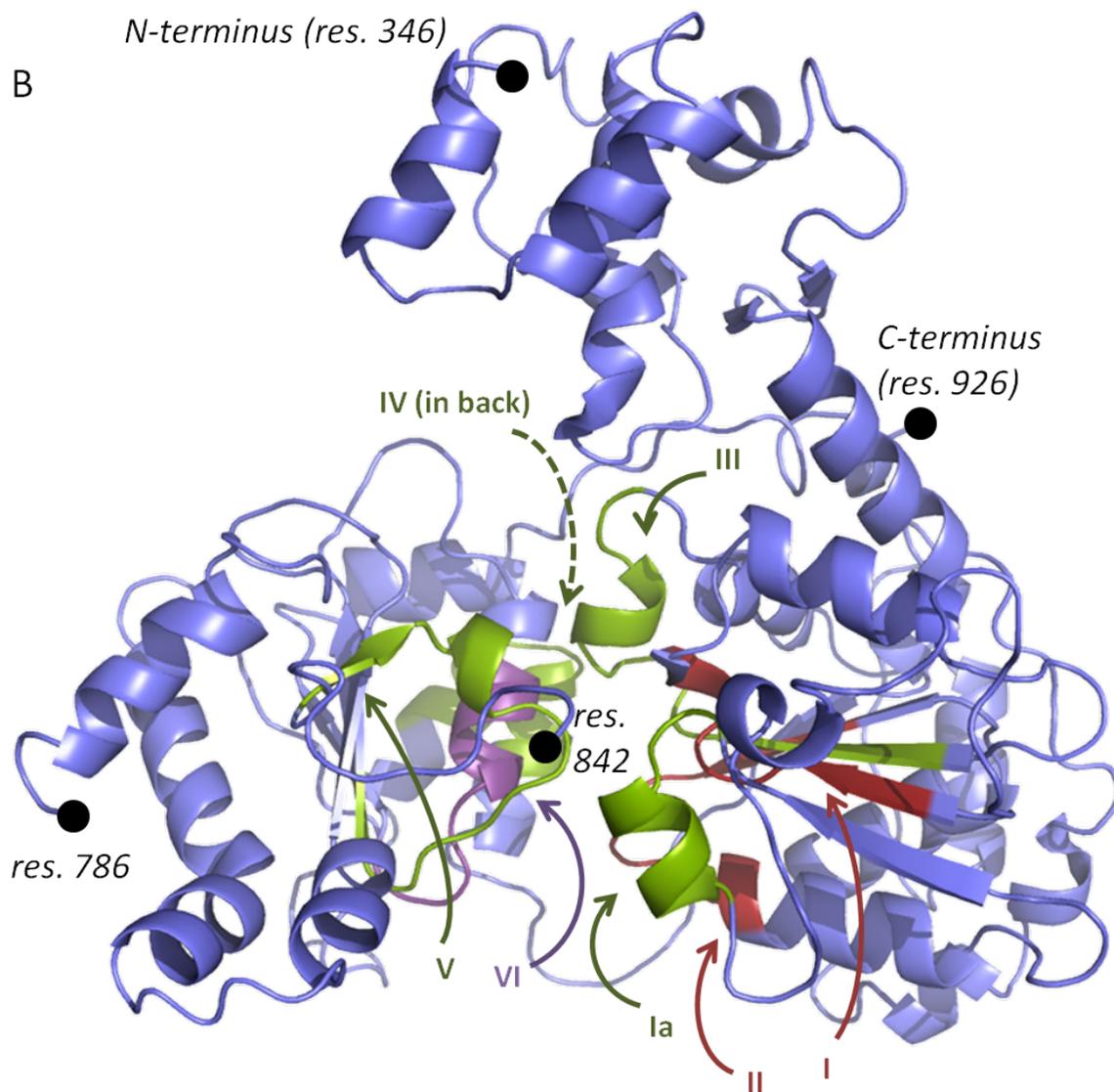


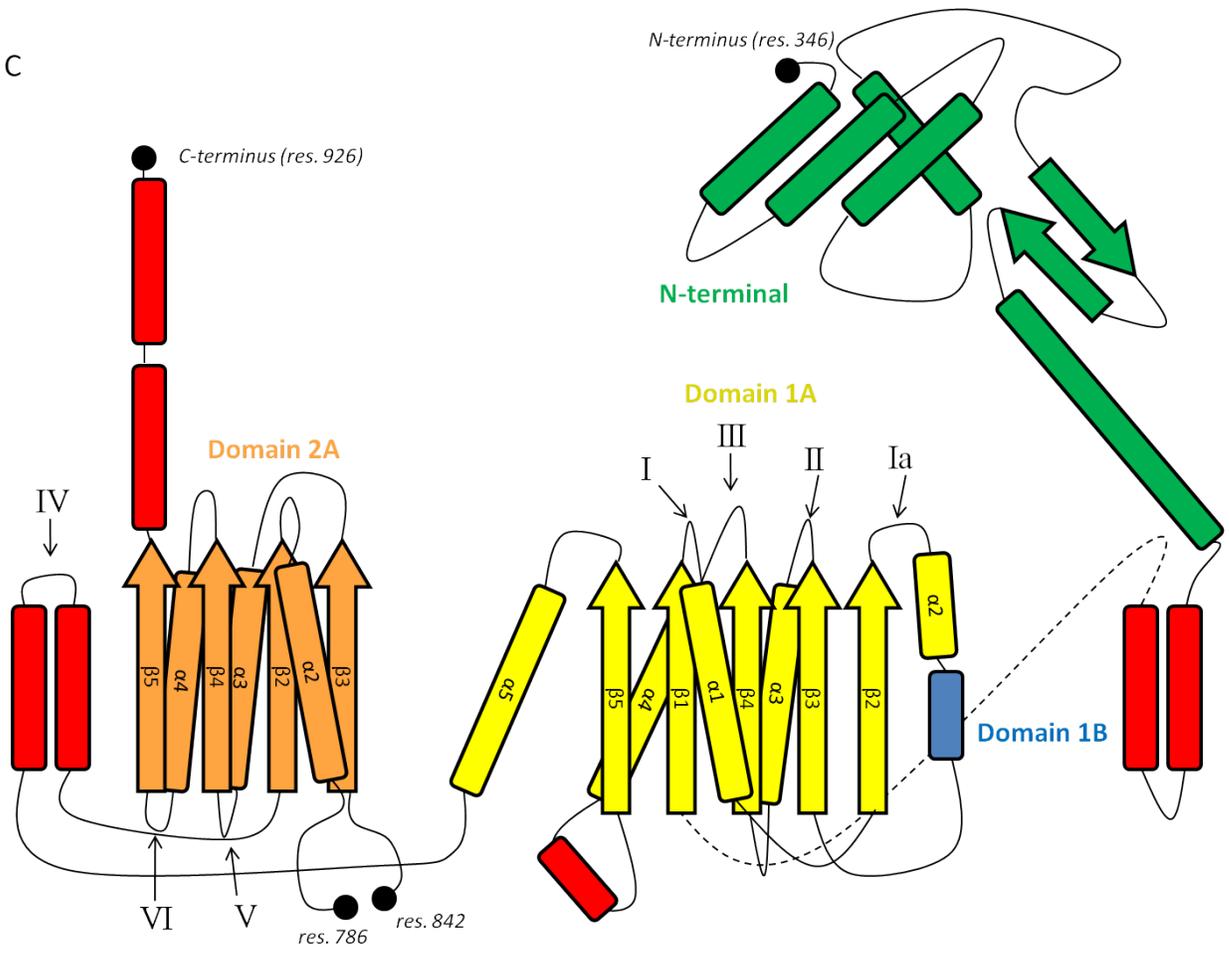
Figure 4.3 Structures of Representative Helicases and the Conserved RecA Core (from Caruthers, 2002⁷⁷)

Topology diagrams of representative helicases. Yellow, conserved RecA-related ‘core’; red, variable structural elements in domains with a RecA-like core; green and blue, additional structural domains. The schematic on the upper right summarizes the positions in the topology of the RecA-like core of the seven conserved motifs defined by Gorbalenya and Koonin for the SF-1 and SF-2 helicases, as well as of additional motifs that are helicase-family specific, shown in parentheses. (from Caruthers, 2002⁷⁷)





C



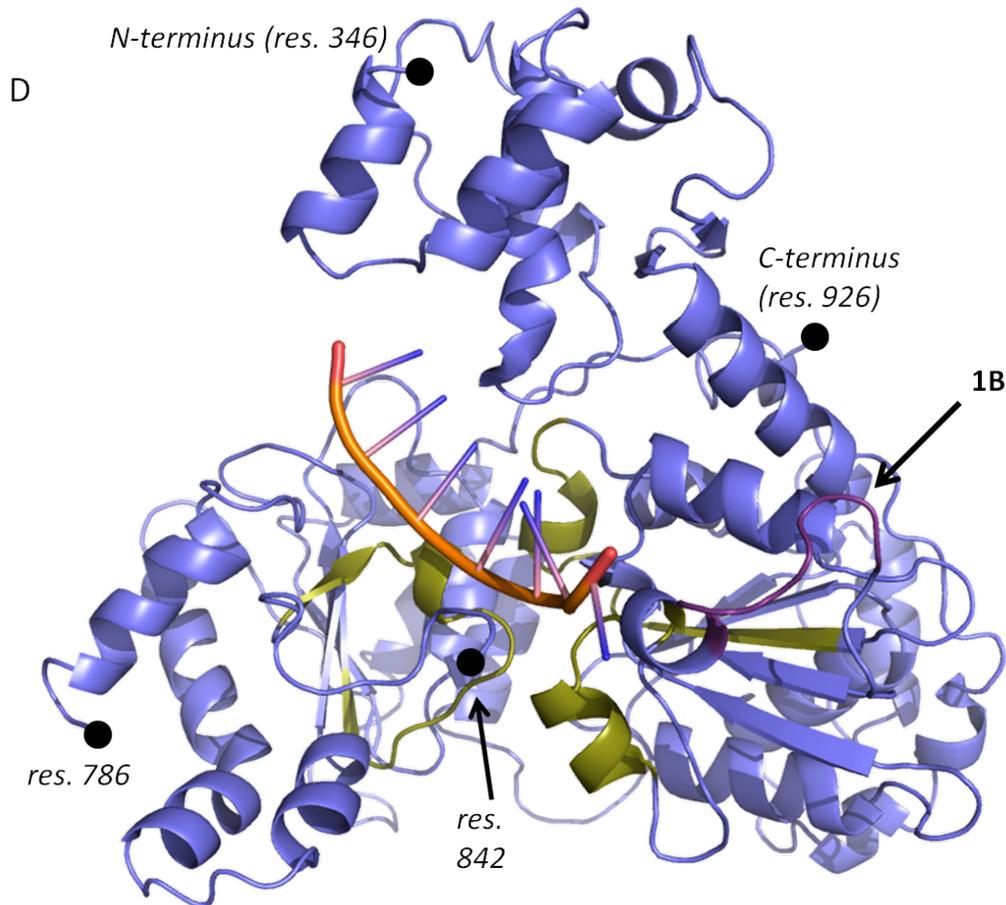


Figure 4.4 Predicted Structure of the pCU1 TraI Helicase

A) The predicted structure of the pCU1 helicase is shown. The N-terminal domain is colored green, and the RecA core domains 1A and 2A are in yellow and orange, respectively. The small domain 1B is in blue and is indicated by an arrow. Black circles mark the N- and C-termini and a gap in the model where intervening residues were not included. By comparing this model to that of RecD2, it appears these missing residues would form domain 2B.

B) The predicted structure of the pCU1 TraI helicase and the location of the seven canonical SF1 helicase motifs are shown: Motifs I and II (Walker A,B) are in red. DNA-binding motifs Ia, III, IV, and V are in green. Motif VI is in purple.

C) Schematic of the pCU1 TraI helicase secondary structure. β -sheets are represented by arrows, α -helices by rectangles, and loop regions by lines. Coloring is similar to that in Figure 4.3 and 4.4A. N-terminal domain is green, Domain 1A is yellow, Domain 2A is orange, Domain 1B is blue, and additional structures which do not correspond directly to that of the RecA core are red. β -sheets and α -helices are numbered in the order in which they appear in the model of pCU1 TraI helicase, as opposed to the numbering scheme in Figure 4.3. The location of each helicase motif is labeled.

D) The pCU1 TraI helicase and DNA from the DNA-bound structure of RecD2 are shown. In olive are helicase motifs involved in DNA-binding, as well as residues corresponding to those proposed by Saikrishnan, et al. (2009)⁷⁹ to participate in helicase translocation along the bound DNA. In purple is domain 1B (indicated by arrow), predicted to serve as a wedge capable of driving the DNA helix apart.

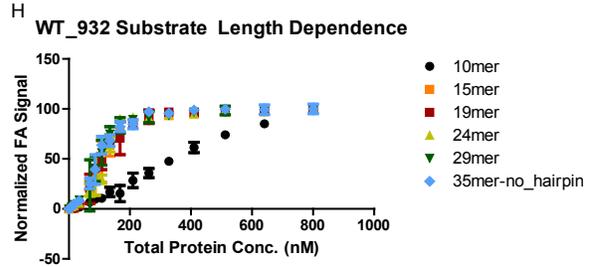
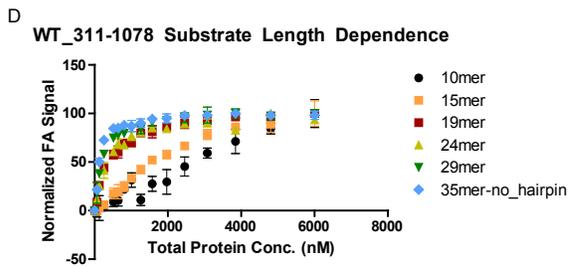
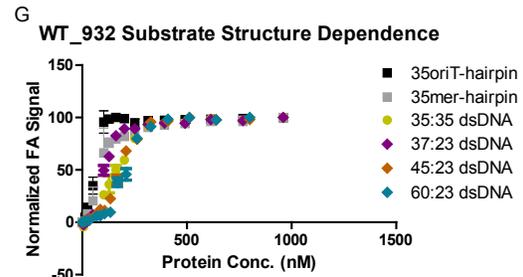
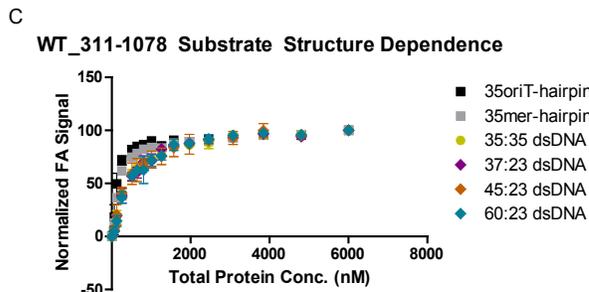
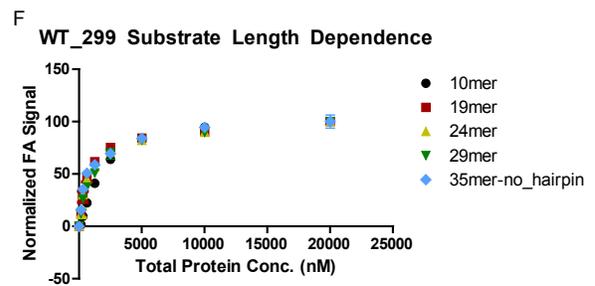
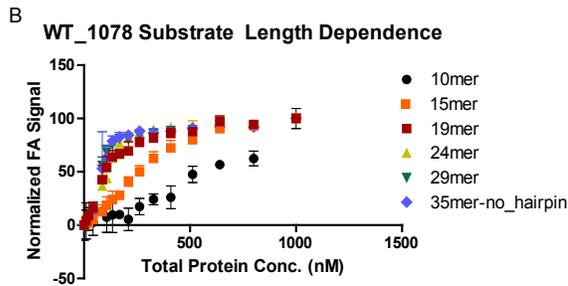
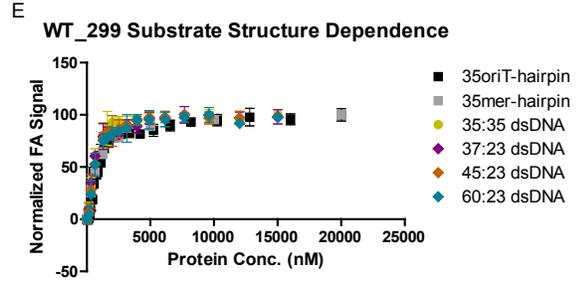
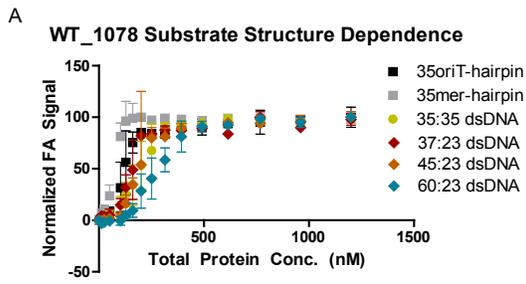


Figure 4.5 DNA Binding Curves

The DNA binding affinity of four pCU1 TraI constructs for a panel of DNA substrates (Table 4.3) was determined by fluorescence anisotropy (FA)-based DNA binding assays. DNA binding curves were generated by following the change in FA of a labeled DNA probe as protein concentration was increased. Each point is the average of one experiment performed in triplicate, with error bars representing the 95% CI of each point. Curves were normalized and then plotted. For each construct, two set of curves are shown. The first set illustrates the impact of DNA probe structure on DNA binding affinity, while the second set illustrates the impact of DNA probe length. **A,B)** WT_1078; **C,D)** WT_311-1078; **E,F)** WT_299; **G,H)** WT_932.

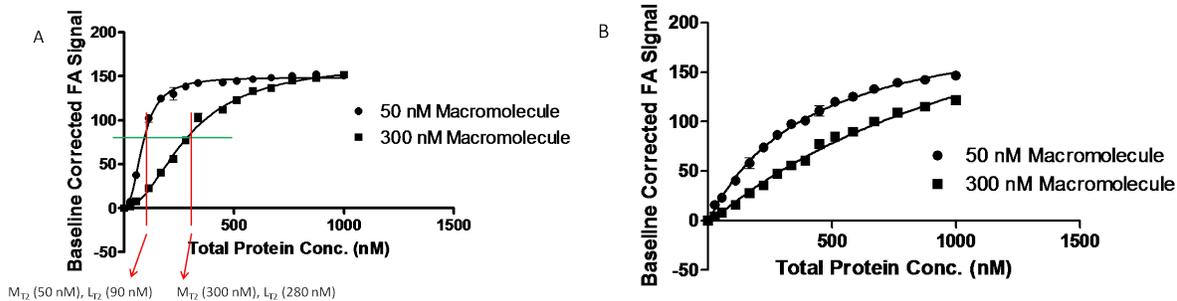


Figure 4.6 Representative DNA Binding Density Curves

A and B) The change in fluorescence anisotropy (FA) of a 50 nM and a 300 nM fluorescently-labeled DNA probe (macromolecule) is measured as a function of protein (ligand) concentration.

A) Macromolecule analyzed: FAM-20mer. The maximum FA signal generated by 50 nM and 300 nM macromolecule at high protein concentration is the same (~150). Therefore, the average degree of binding of WT_1078 on this macromolecule can be calculated for a series of FA values.

B) Macromolecule analyzed: FAM-14mer. In contrast, the maximum FA signal generated by 50 nM and 300 nM macromolecule at high protein concentration is not the same, making determination of the average degree of binding for a series of FA values error-prone.

A) To determine the average degree of binding at the FA value chosen (green line), the corresponding DNA macromolecule and protein ligand concentrations are determined for each curve (red lines). This is repeated 21 times for a series of FA values between 10% and 80% of the maximum FA value. These pairs of values are then used in Equation 4.2 to calculate the average degree of binding at each FA value.

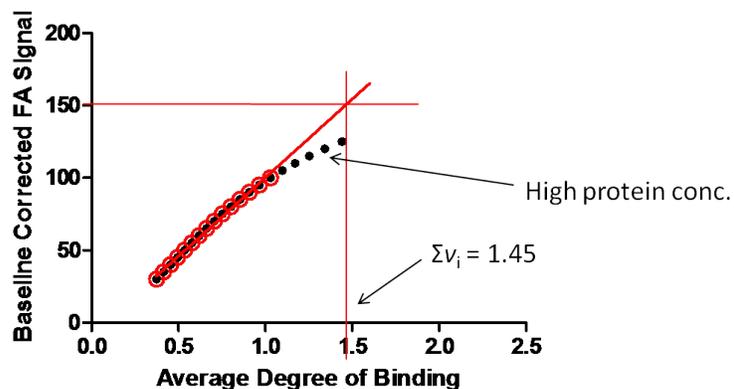


Figure 4.7 Average Degree of Binding as a Function of Fluorescence Anisotropy

For each DNA macromolecule analyzed, the average degree of binding ($\Sigma\nu_i$) is calculated for a series of 21 FA values (Figure 4.6). Each FA value is then plotted as a function of the corresponding calculated $\Sigma\nu_i$. A line of best fit (red diagonal) describes the relationship between $\Sigma\nu_i$ and FA. Data points deviating from the line of best fit (black points not circled in red) correspond to high protein concentration and are omitted to avoid introducing error into the calculations. From a plot of FA vs. protein concentration (Figure 4.6), the maximum FA value for each DNA macromolecule is known (Example: max FA = 150 for 20mer). The $\Sigma\nu_i$ at this maximum FA value is determined from the line of best fit, as indicated on the plot above, and represents the maximum $\Sigma\nu_i$ for that DNA macromolecule.

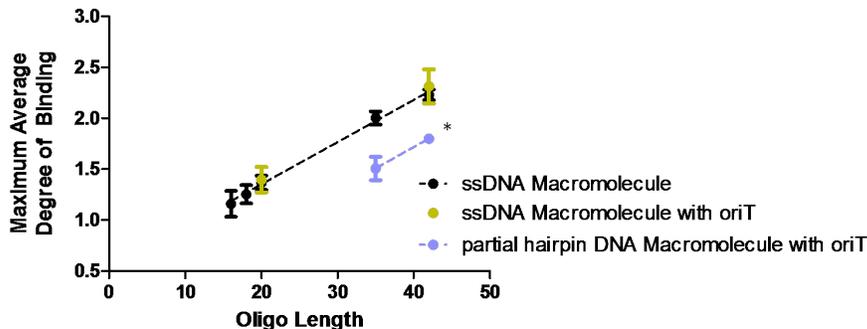


Figure 4.8 Maximum Average Degree of Binding as a Function of Substrate Length, Structure

The maximum average degree of binding (Σv_i) for each DNA substrate was determined as described in Figures 4.6 and 4.7. With one exception (*), each data point represents the average calculated maximum Σv_i for three experiments, each performed in triplicate, and the error bars represent the standard deviation of these calculations. *This value is the average of one experiment performed in triplicate. In black are ssDNA macromolecules of random sequence, in yellow are ssDNA macromolecules that contain the *oriT* sequence of pCU1 TraI, and in blue are hairpin-forming DNA macromolecules that contain the *oriT* sequence of pCU1 TraI. As can be seen, the maximum Σv_i (y-axis) increases as a function of DNA macromolecule length (x-axis). Hairpin-containing macromolecules appear, on average, 10 nt shorter than their ssDNA counterparts. The magnitude of this effect is expected to vary with the size of the hairpin formed.

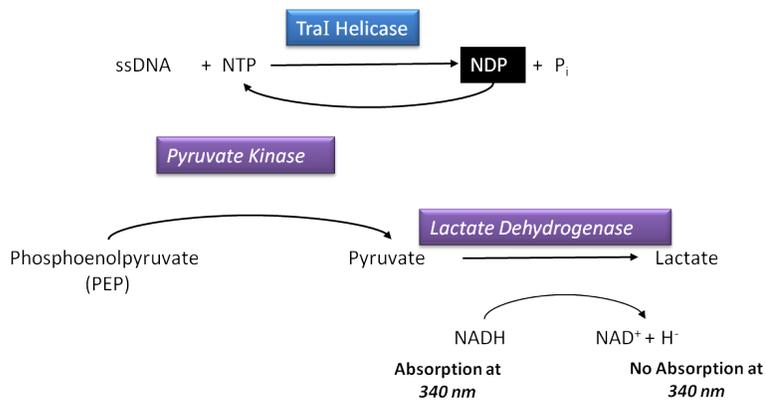
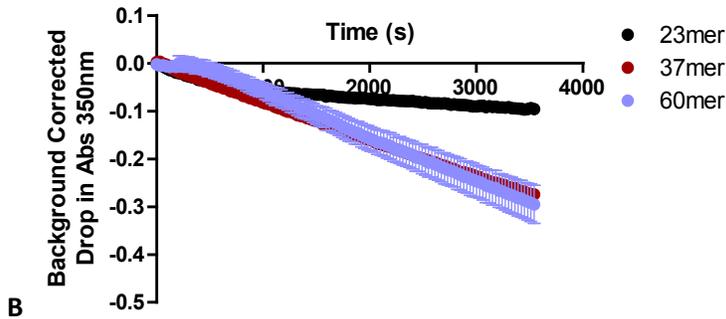


Figure 4.9 Design of the NADH-coupled ATPase Activity Assay

The pCU1 TraI helicase (blue square) hydrolyzes NTPs to NDPs in the presence of ssDNA. This activity is detected using two coupled enzymes, pyruvate kinase and lactate dehydrogenase. Pyruvate kinase converts NDPs back to NTP while also breaking down PEP to pyruvate. Lactate dehydrogenase converts pyruvate to lactate, while oxidizing NADH to NAD⁺. While NADH absorbs at 340 nm, NAD⁺ does not; thus the drop in absorbance at 340 nm corresponds to NTP hydrolysis by the pCU1 TraI helicase.

A

Impact of ssDNA Length on ATPase Activity



B

Impact of NTP on ATPase Activity

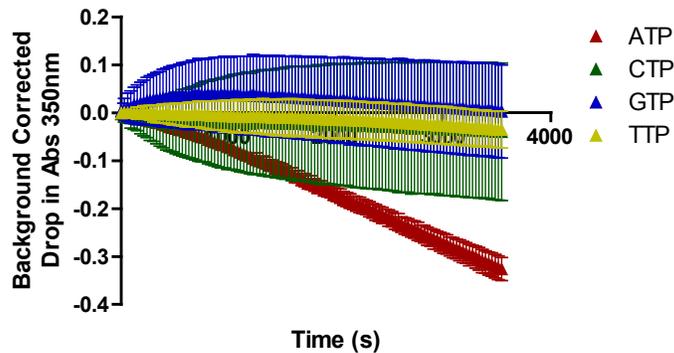


Figure 4.10 pCU1 TraI ATPase Activity as a Function of ssDNA Length, NTP Concentration

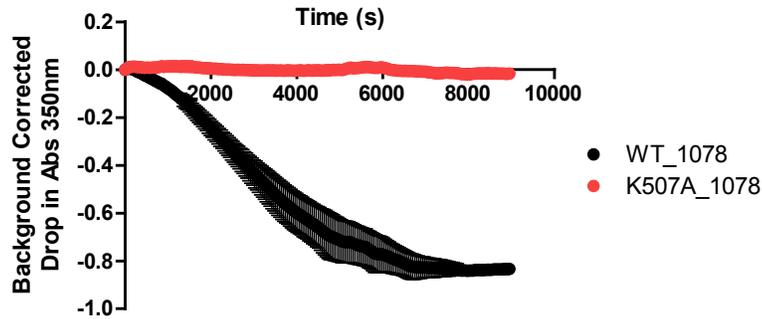
Using the NADH-coupled ATPase activity assay outlined in Figure 4.9, the impact of ssDNA length and NTP on the ATPase activity of the pCU1 TraI helicase was determined. ATPase activity is directly related to the drop in absorbance at 350 nm (y-axis) over time (x-axis). Each point is the average of three replicates, and has been corrected for background NADH oxidation and NTP hydrolysis by subtraction of a No TraI Control. Error bars were calculated using the equation: $\text{Error} = \text{Square root} [(\text{Raw error})^2 + (\text{No TraI Control error})^2]$, where raw error is the error associated with the experimental data prior to subtraction of the No TraI Control.

A) ATPase activity increases with ssDNA length; substrates longer than 60 nt long were not analyzed. Assay conditions: 100 nM TraI, 1.25 mM ATP, 100 nM 60mer ssDNA.

B) ATPase activity is stimulated by ATP, but not by the other three NTPs tested. The background NADH oxidation and NTP hydrolysis observed for CTP, TTP, and GTP was significantly higher than that for ATP, thus generating the large error bars observed above. Assay conditions: 100 nM TraI, 1.25 mM NTP, 100 nM 60mer ssDNA.

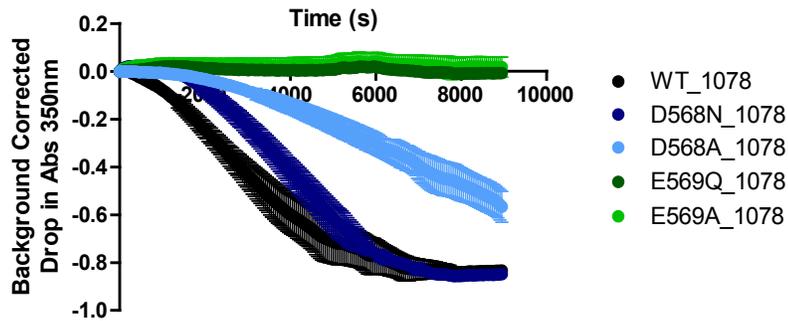
A

Impact of Mutations within the Walker A Motif on ATPase Activity



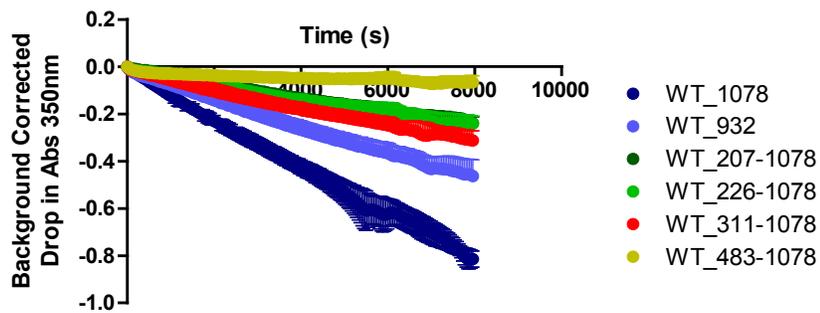
B

Impact of Mutations within the Walker B Motif on ATPase Activity



C

Impact of TraI Construct Length on ATPase Activity - Low [ssDNA]



D

Impact of TraI Construct Length on ATPase Activity - High [ssDNA]

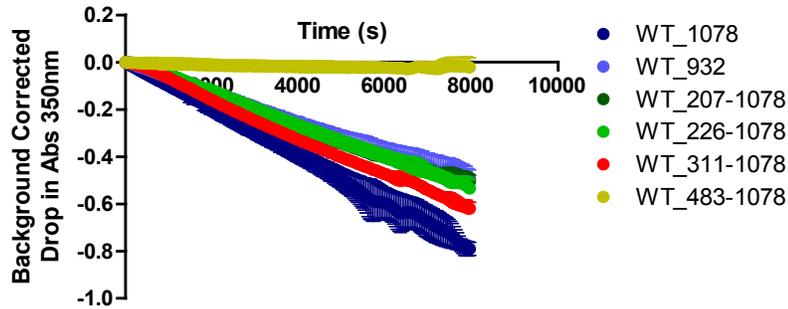


Figure 4.11 Impact of pCU1 TraI Construct Length and Mutagenesis of the Walker A and B Motifs on pCU1 TraI ATPase Activity

Using the NADH-coupled ATPase activity assay outlined in Figure 4.9, the impact of mutations made within the Walker A and Walker B ATPase motifs on the ATPase activity of pCU1 TraI, and the minimal length of the pCU1 TraI helicase, were determined. ATPase activity is directly related to the drop in absorbance at 350 nm (y-axis) over time (x-axis). Each point is the average of three replicates, and has been corrected for background NADH oxidation and NTP hydrolysis by subtraction of a No TraI Control. Error bars were calculated using the equation: $\text{Error} = \text{Square root} [(\text{Raw error})^2 + (\text{No TraI Control error})^2]$, where raw error is the error associated with the experimental data prior to subtraction of the No TraI Control.

A) ATPase activity is eliminated by mutation of K507 within the Walker A motif. Assay conditions: 100 nM TraI, 1.25 mM ATP, 500 nM 60mer ssDNA.

B) ATPase activity is eliminated by mutation of E569 within the Walker B motif, and is inhibited by the D568A mutation within the Walker B motif. The D568N mutation within the Walker B motif has only a limited impact on ATPase activity. Assay conditions: 100 nM TraI, 1.25 mM ATP, 60mer ssDNA.

C) The full length TraI (WT_1078) generates the greatest ATPase activity. ATPase activity is maintained in the truncated constructs WT_207-1078, WT_226-1078, WT_311-1078, and WT_932. It is eliminated by truncation of WT_1078 to WT_483-1078. Note these assay conditions include a low concentration of ssDNA as compared to Figure 4.11d. Assay conditions: 100 nM TraI, 1.25 mM NTP, 100 nM 60mer ssDNA.

D) In the presence of a higher ssDNA concentration, the full length TraI (WT_1078) still generates the greatest ATPase activity, but the ATPase activity of the truncated constructs WT_207-1078, WT_226-1078, and WT_311-1078 is increased. No significant change in the ATPase activity of WT_932 is observed. ATPase activity is still eliminated by truncation of WT_1078 to WT_483-1078. Assay conditions: 100 nM TraI, 1.25 mM NTP, 500 nM 60mer ssDNA.

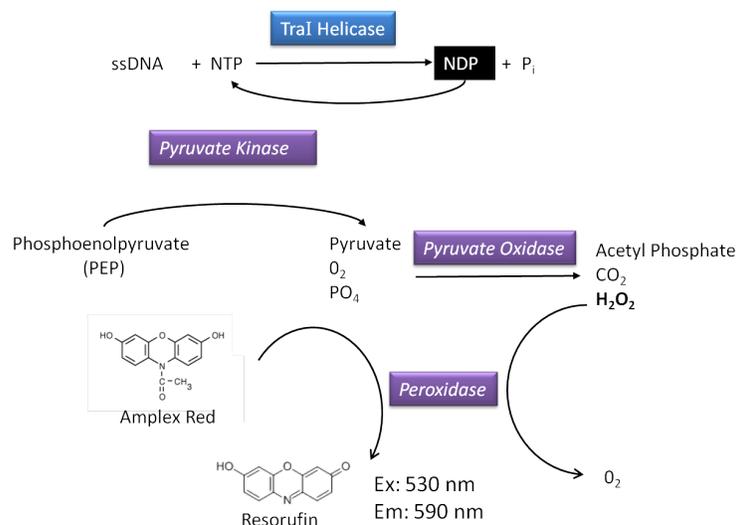


Figure 4.12 Design of the ADP Quest ATPase Activity Assay

The pCU1 TraI helicase (blue square) hydrolyzes NTPs to NDPs in the presence of ssDNA. This activity is detected using three linked enzymes, pyruvate kinase, pyruvate oxidase, and horseradish peroxidase. Pyruvate kinase converts NDPs back to NTP while also breaking down PEP to pyruvate. Pyruvate oxidase breaks pyruvate down to acetyl phosphate, CO₂, and H₂O₂. Horseradish peroxidase uses the oxidizing power of H₂O₂ to oxidize amplex red to resorufin. Resorufin emits at 590 nm upon excitation at 530 nm. Therefore, an increase in fluorescence by resorufin corresponds to NTP hydrolysis by the pCU1 TraI helicase.

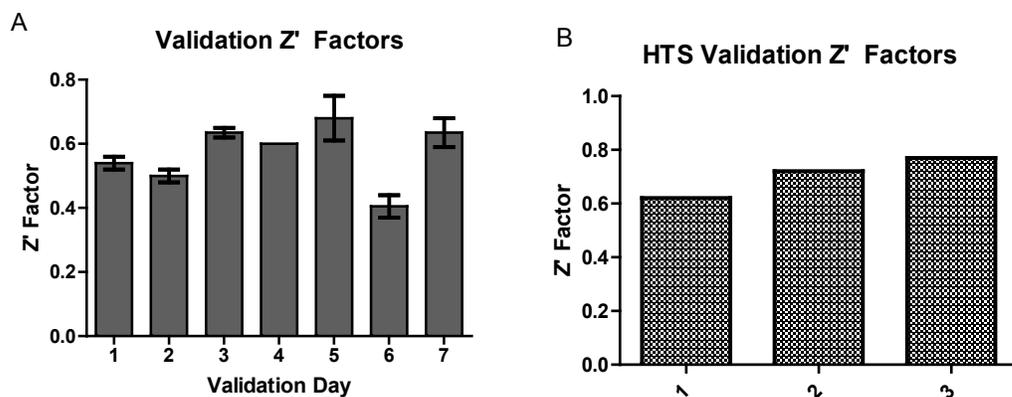


Figure 4.13 Validation of the ADP Quest ATPase Activity Assay Design

A) The calculated Z' factors for each of seven days of assay validation are shown. Four of the seven days meet or exceed the standard of a Z' factor of 0.5 or greater.

B) The calculated Z' factors for the three HTS validation plates are shown; each of the plates meets or exceeds the standard of a Z' factor of 0.5 or greater.

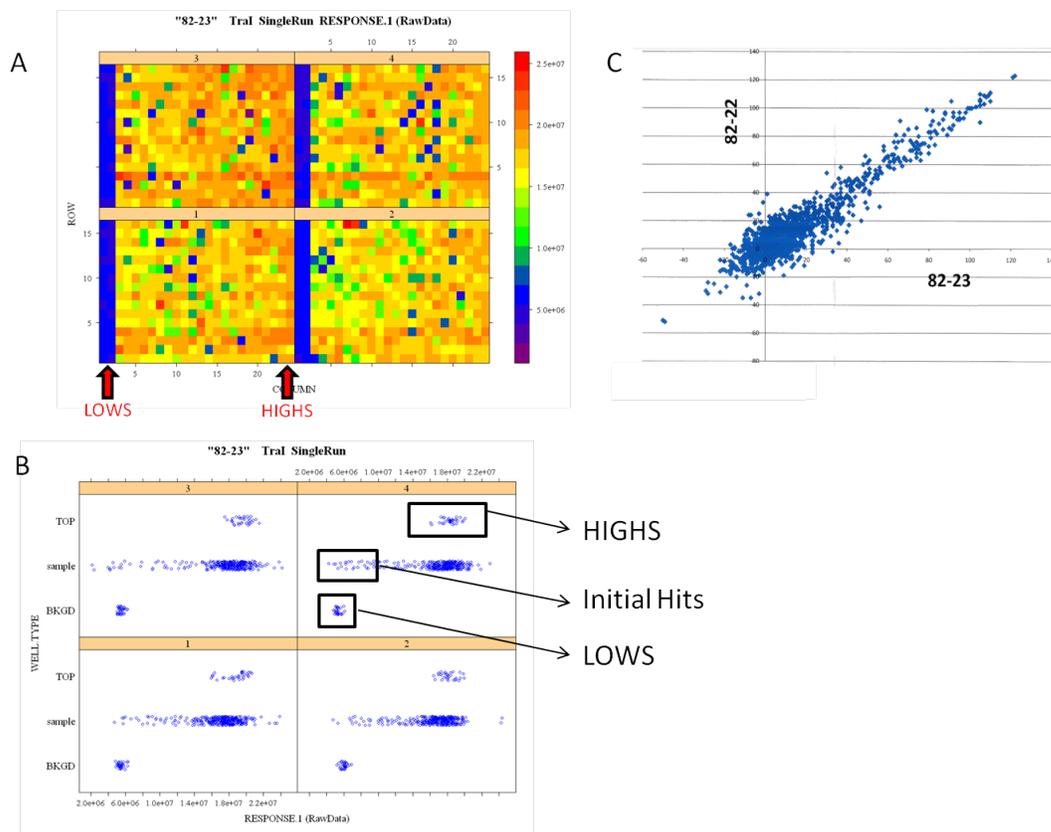


Figure 4.14 High Throughput Screening of the LOPAC against pCU1 TraI ATPase Activity

A) Plate views of the four plates composing one replicate of the LOPAC screen are shown. The LOW and HIGH columns of plate 1 are labeled. On the right, is a color gradient indicating the level of fluorescence detected for each well.

B) Strip plots of the four plates composing one replicate of the LOPAC screen are shown. Each dot corresponds to a well on each plate. For plate 4, The LOW, HIGH, and initial hits of the compound-containing wells are boxed and labeled.

C) Comparison of the duplicate runs of the four plates composing the LOPAC screen. On the y-axis is the signal detected by replicate 1; on the x-axis is the signal detected by replicate 2. Points falling along the diagonal generated the same or a similar signal on both plates.

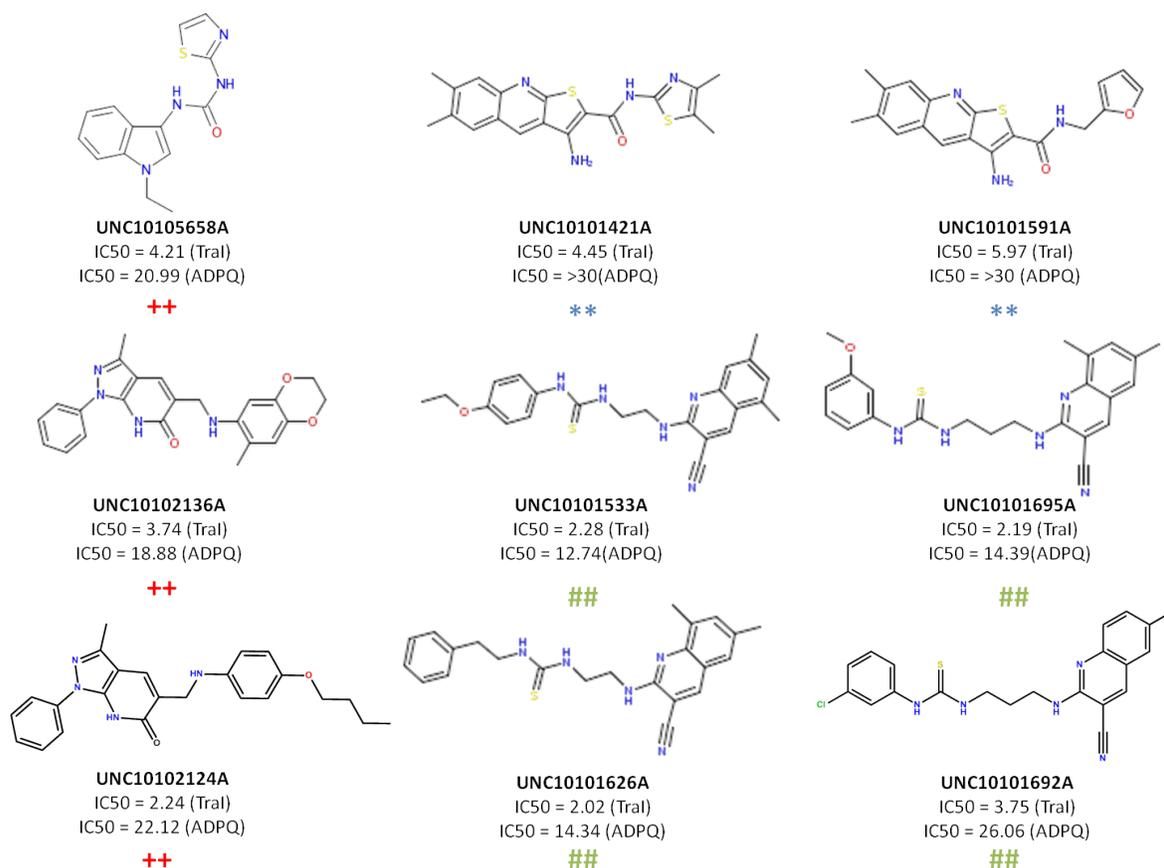


Figure 4.15 Putative Inhibitors of the pCU1 Tral ATPase Identified through High Throughput Screening

The structures of nine putative inhibitors of the pCU1 Tral ATPase, identified by high throughput screening, are provided. For each, the IC50 values, in μM , against the ATPase activity of pCU1 Tral and the components of the ADP Quest assay are provided. The compounds could be divided into three general chemical series, which are designated by symbols below each label (++, **, ##).

Chapter 5: Characterization of Plasmid pCU1 Transfer

5.1. Introduction to Plasmid Transfer

The relaxase and helicase domains of pCU1 TraI cooperate to transfer the T-strand of plasmid pCU1 into a recipient cell. In order to study the activities of these enzymes at the cellular level, a method to characterize pCU1 plasmid transfer between bacterial cells must be developed.

The transfer of a wide variety of plasmids between Gram negative and Gram positive organisms has been studied^{3, 23, 103-105}, often with the goal of identifying a point at which transfer can be inhibited^{6, 106-109}. The conditions required for optimal plasmid transfer vary with each system analyzed. For example, transfer of a specific plasmid between bacteria may occur primarily in liquid media, primarily on solid media, or both, and this transfer can be temperature dependent^{51, 75, 104, 110-112}. The rate of transfer is plasmid-specific, and this rate is often influenced by the relative and total concentrations of donor and recipient cell populations, as well as the identity of the donors and recipients^{9, 28, 104, 108, 111-116}. Plasmid-encoded antibiotic resistance is a commonly used marker to follow a transmissible plasmid from a donor cell to a recipient cell, though other markers including bioluminescence, or growth on media selective for auxotrophs have been used^{51, 107, 108, 117}.

5.2. Method Development for Characterization of Plasmid pCU1 Transfer

The following steps were taken in order to develop and optimize such a method capable of following the transfer of plasmid pCU1 between bacterial cells. First, a donor cell line and recipient cell line were established. The donor cell line is defined as the cell line containing a conjugative plasmid. The recipient cell line is defined as the cell line lacking a conjugative plasmid. Upon mating the donor and recipient cell lines, a third line is generated, the transconjugates, which are defined as recipient cells obtaining the conjugative plasmid during the course of the experiment. The recipient cell line was laboratory strain *Escherichia coli* BL21-AI (tetracycline resistant, Tet^R), and the donor cell line was

laboratory strain *Escherichia coli* HMS174(DE3) (rifamycin resistant, Rif^R) that was then transformed with the conjugative plasmid pCU1. pCU1 is resistant to streptomycin (Strep^R), spectomycin (Spec^R) and ampicillin (Amp^R). The plasmid pCU1 was obtained as a gift from S. Paterson (Carleton University).

Second, the minimal inhibitory concentration (MIC) of each antibiotic was confirmed for each cell line (Table 5.1). The following protocol was used to determine these MIC values.

1. 5 mL cultures of each cell line were grown under antibiotic selection in Luria Broth (LB) with shaking at 37°C for 16 h
2. 1.5 mL of each 5 mL culture was diluted into 50 mL LB subculture and grown without antibiotic selection with shaking at 37°C until the optical density at 600 nm (OD₆₀₀) reached 0.5 (approximately 2 h)
3. 10 µL of each subculture was diluted into 5 mL LB
4. A serial dilution of the antibiotic(s) of interest was performed in LB in a 96-well clear plastic U-bottom plate (Costar), such that each well contained 100 µL of diluted antibiotic following the serial dilution
5. 2 µL of the diluted cell population was then added to each well
6. The 96-well plate was covered with an adhesive AirPore Tape sheet (Qiagen) and incubated at 37°C for 16 h
7. Following incubation, each well was visually inspected for turbidity to determine the concentration of antibiotic(s) necessary to inhibit growth of the cell line.

Since donor and recipient cell lines were resistant to a unique set of antibiotics (Table 5.1), donor, recipient, and transconjugate (recipients obtaining the pCU1 plasmid) cells could be isolated using antibiotic selection. In particular, transconjugates and recipients were selected with tetracycline (T), transconjugates and donors with streptomycin (S), and transconjugates with T+S. Since transconjugates grew in the presence of both antibiotics, the final number or concentration of donors and recipients were obtained by subtracting the transconjugate population from that of donors + transconjugates and from that of recipients + transconjugates.

A protocol for following transfer of plasmid pCU1 from the HMS174(DE3) donor strain into the BL21-AI recipient strain was modified from that published for analysis of F plasmid transfer⁵¹. This protocol incorporates the use of Oxygen Biosensor plates to determine the size of each cell population (donor, recipient, transconjugate) at the end of the plasmid transfer assay. Oxygen Biosensor plates contain at the bottom of each well an oxygen-sensitive gel that is fluorescent (excitation 485 nm, emission 620 nm) when in an anaerobic environment. Therefore, upon the addition of a population of bacterial cells, the gel will fluoresce once the population is sufficiently large to consume a significant portion of the oxygen in solution, and the gel's fluorescent signal will continue to increase as the population increases (Figure 5.1). Therefore, if the plate is read continually following the addition of cells, the time of rise in fluorescent signal is directly related to the size of the population of cells initially placed into the well (Figure 5.2). In every three columns of the 96- or 384-well Oxygen Biosensor plate, tetracycline, streptomycin, and tetracycline + streptomycin was added, thus allowing for the selection of the three cell populations (recipients, donors, and transconjugates) on the multi-well plate.

The modified F plasmid protocol consisted of the following steps (Figure 5.3):

1. 5 mL cultures of donor (D) and recipient (R) were grown under antibiotic selection in LB with shaking at 37°C for 16 h
2. 2 mL of donor and 5 mL of recipient 5 mL cultures were diluted into 50 mL (donor) and 150 mL (recipient) LB subcultures and grown without antibiotic selection with shaking at 37°C until OD₆₀₀ reached 0.5 (approximately 2 h)
3. Subcultures were mixed using a 10 mL donor to 25 mL recipient (v/v) ratio and pelleted at 5000xg
4. Pelleted cells were resuspended in 8.75 mL LB to increase the cell density by a factor of 4
5. 250 µL of the resuspended mixture of cells was applied to each well of a 96-well filter-bottom plate (0.2 µm hydrophobic PVDF membrane, Corning) and the plate was centrifuged for 2 min at 3000xg to adhere the cells onto the filters

6. 250 μ L LB or 250 μ L LB + compound was then pipetted into specific wells of the 96-well filter-bottom plate to generate LB controls and experimental samples, respectively
7. Filter-bottom plates were then incubated at 37°C for a designated time period (2, 24, or 48 h)
8. After the designated number of hours, filter-bottom plates were spun for 2 min at 3000 \times g to remove LB and LB + compound from the cells
9. 250 μ L LB was then added to each well and cells were resuspended in this volume
10. 200 μ L of the resuspended cell volume was then diluted into 800 μ L of LB to dilute the cells 4-fold and generate the final control and experimental samples
11. Individual cell populations (donors, recipients, transconjugates) within control and experimental samples were then analyzed as described below

To determine the absolute number of donors, recipients, and transconjugates at the end of the experiment, LB controls were diluted in LB and plated onto selective media. Plates were incubated for 16 h at 37 °C. The number of colony forming units (CFUs) on each plate was used to calculate the final size of each cell population.

Experimental wells containing cells exposed to increasing concentrations of each compound were transferred to Oxygen Biosensor plates to determine the effect each compound had on the size of donor and recipient populations, and the number of transconjugates produced during the duration of the assay. These cells were not diluted prior to transfer. To determine the relationship between signal generated on an Oxygen Biosensor plate and the absolute size of the cell population, LB controls were serially diluted and then transferred to three columns of an Oxygen Biosensor plate containing S, T, and S+T, respectively. Therefore, the serially diluted LB control served as a standard curve for the remainder of the plate, allowing the fluorescent signal to be correlated with relative population size, and the numbers of CFUs appearing on the selective media were used to calibrate the fluorescent signal of the standard curve to a specific number of CFUs.

5.3. Analysis of Plasmid pCU1 Transfer

As detailed above, after each experiment, the size of the donor, recipient, and transconjugate cell populations was determined by plating the LB control onto selective media (Figure 5.3). The rate of plasmid transfer varied widely from experimental day to experimental day (Table 5.2). The two most likely causes of variability in plasmid transfer were the fluctuating size of the final donor and recipient cell populations and the slow rate of transfer of plasmid pCU1 from donor to recipient. Final donor and recipient population size varied between 2 and 10 fold from experimental day to experimental day, with this variability increasing as the incubation time of the assay lengthened. A small portion of this variability is a result of inconsistency in the number of donor and recipients initially mixed. The final OD₆₀₀ of input cells often ranged from 0.5 to 0.65, but this fractional change in initial population size can only partially explain the observed 2-10 fold changes observed. Since the rate of transfer, which is expressed as final number of transconjugates per final number of donors, was so slow, it was particularly sensitive to slight changes in the donor cell population.

Due to the extensive variability inherent in the current assay, it would be difficult, if not impossible, to quantify any effect (inhibitory or stimulatory) on transfer observed secondary to exposure to a compound. Therefore, in order to accurately characterize plasmid pCU1 transfer, and eventually the effect of any compounds on this transfer, the rate of plasmid transfer must be increased, and the impact of donor and recipient population size on the rate of transfer must be determined.

Since plasmid pCU1 was originally isolated from *Salmonella typhimurium*^{52,53}, it was possible that use of this species as donor or recipient, or both, would enhance the rate pCU1 plasmid transfer. However, initial experiments involving a strain of Group B *Salmonella* isolated from contaminated cheese (gift from L. Ramsey, Virginia Dept. of Agriculture and Consumer Services) revealed the irregular morphology of colonies produced by this species, which hindered accurately counting the number of CFUs on the control LB-agar plate. For future work, other strains of *Salmonella* may prove more tractable.

Related, but unpublished, work was performed by others in the Redinbo laboratory to characterize transfer of the F plasmid. This work highlighted an additional variable to consider when

performing plasmid transfer assays. It was determined that the ratio of donors to recipients and the absolute number of donor and recipient cells at the initiation of plasmid transfer experiments influenced the sensitivity of each population to certain compounds. At low concentrations, the apparent “kill rate” was greater for a certain population than at higher concentrations of that same cell type. Due to the variability in the size of donor and recipient populations during pCU1 plasmid transfer assays, this observation has significant import on the reliability of data describing the killing or survival rate of plasmid pCU1 donors and recipients in the presence of various compounds.

5.4 Tables and Figures

Table 5.1 Minimal Inhibitory Concentration of Antibiotics for Selected Cell Lines

Table 5.2 Number of Transconjugates Generated during Plasmid pCU1 Transfer Assay

Figure 5.1 Oxygen Biosensor Technology

Figure 5.2 Detection of Cell Population Size using the Oxygen Biosensor Plate Method

Figure 5.3 Plasmid pCU1 Transfer Protocol

	Ampicillin	Streptomycin	Spectinomycin	Rifampicin	Tetracyclin
HMS174	4	4	4	>128	1
HMS174 + pCU1	>128	>128	>128	>128	0.125-0.5
BL21 AI	4-8	8-16	4	8-16	32-64

Table 5.1 Minimal Inhibitory Concentration of Antibiotics for Selected Cell Lines

The minimal inhibitory concentrations (MICs) in $\mu\text{g/mL}$ of five antibiotics are provided for cell lines used during the analysis of plasmid pCU1 transfer. *Escherichia coli* HMS174(DE3) + pCU1 served as the donor cell line, while *Escherichia coli* BL21-AI served as the recipient cell line.

Incubation Time	# Transconjugates/#Donors
2hr	0.023 (+/- 0.03)
24hr	0.027 (+/- 0.06)
48hr	0.01 (+/- 0.015)

Table 5.2 Number of Transconjugates Generated during Plasmid pCU1 Transfer Assay

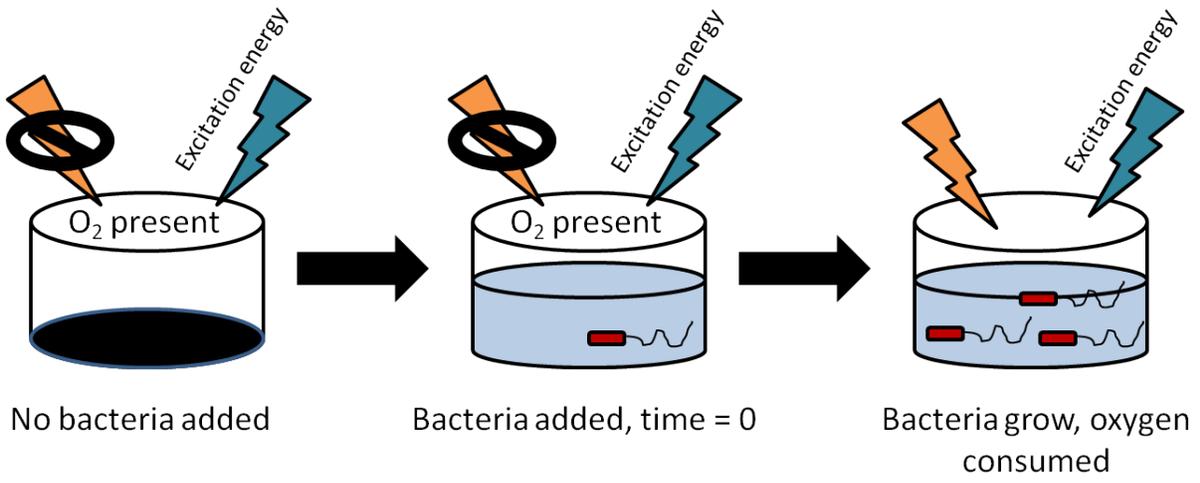


Figure 5.1 Oxygen Biosensor Technology

The gel material at the bottom of each well of the 96- and 384-well Oxygen Biosensor plates is fluorescent when in an anaerobic environment (Ex. 485 nm, Em. 620 nm), but this fluorescence is quenched by oxygen. Therefore, as bacterial populations grow and consume the oxygen in the well, the fluorescent signal increases.

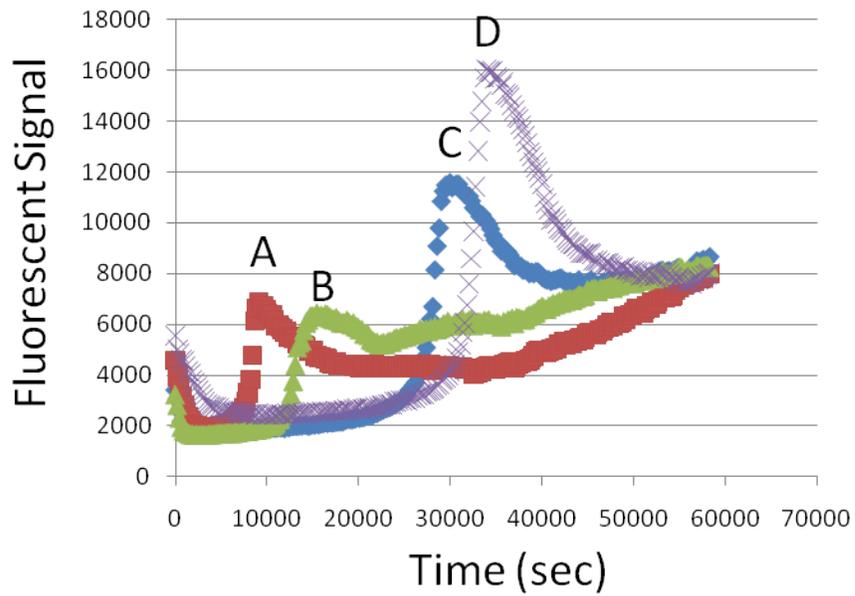


Figure 5.2 Detection of Cell Population Size using the Oxygen Biosensor Plate Method

The time of increase in fluorescent signal for each well is directly related to the initial size of the bacterial population placed into that well. According to the time of rise in signal seen above, the initial size of the cell populations were of the order $A > B > C > D$.

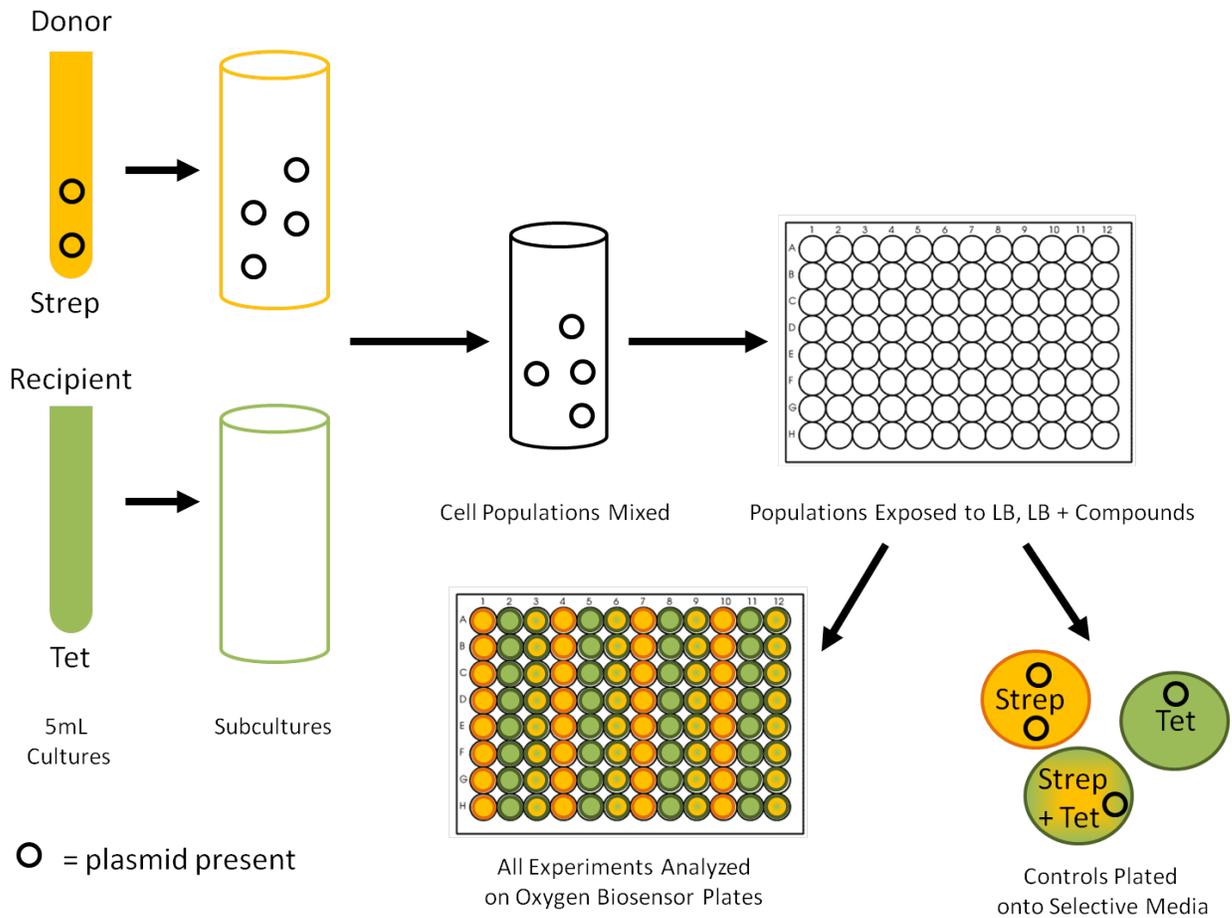


Figure 5.3 Plasmid pCU1 Transfer Protocol

To analyze the transfer of the conjugative plasmid pCU1 between bacterial cells, as well as the effect putative small molecule inhibitors may have on this transfer, donor (pCU1 plasmid plus) and recipient (no plasmid) cell lines are cultured and then mixed in the presence or absence of these compounds. After mixing, donor cells transfer the plasmid pCU1, which contains streptomycin resistance, to recipient cells, which are tetracycline resistant, to form a third population of cells called transconjugates, which are resistant to both tetracycline and streptomycin. The number of surviving donors, recipients, and transconjugates are determined by antibiotic selection on LB-agar plates and Oxygen Biosensor 96- and 384-well plates.

References

1. de la Cruz, F.; Davies, J., Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends in Microbiology*. **2000**, 8, (3), 128-133.
2. Lanka, E.; Wilkins, B. M., DNA processing reactions in bacterial conjugation. *Annual Review of Biochemistry* **1995**, 64, 141-169.
3. Smillie, C.; Garcillan-Barcia, M. P.; Francia, M. V.; Rocha, E. P. C.; de la Cruz, F., Mobility of plasmids. *Microbiology and Molecular Biology Reviews* **2010**, 74, (3), 434-452.
4. Waldor, M. K., Mobilizable genomic islands: going mobile with oriT mimicry. *Molecular Microbiology* **2010**, 78, (3), 537-540.
5. Larkin, C.; Datta, S.; Harley, M. J.; Anderson, B. J.; Ebie, A.; Hargreaves, V.; Schildbach, J. F., Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure* **2005**, 13, 1533-1544.
6. Williams, J. J.; Hergenrother, P. J., Exposing plasmids as the achilles' heel of drug-resistant bacteria. *Current Opinion in Chemical Biology* **2008**, 12, 389-399.
7. Mazel, D.; Davies, J., Antibiotic resistance in microbes. *Cellular and Molecular Life Sciences* **1999**, 56, (9-10), 742-754.
8. Francia, M. V.; Varsaki, A.; Garcillan-Barcia, M. P.; Latorre, A.; Drainas, C.; de la Cruz, F., A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiology Reviews* **2004**, 28, 79-100.
9. Waters, V. L., Conjugative transfer in the dissemination of beta-lactam and aminoglycoside resistance. *Frontiers in Bioscience* **1999**, 4, d433-456.
10. Garcillan-Barcia, M. P.; Francia, M. V.; de la Cruz, F., The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiology Reviews*. **2009**, 33, (3), 657-687.
11. Llosa, M.; Gomis-Ruth, F.; Coll, M.; de la Cruz, F., Bacterial conjugation: a two-step mechanism for DNA transport. *Molecular Microbiology* **2002**, 45, (1), 1-8.
12. Ilyina, T. V.; Koonin, E. V., Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria. *Nucleic Acids Research*. **1992**, 20, (13), 3279-3285.

13. Larkin, C.; Haft, R. J.; Harley, M. J.; Traxler, B.; Schildbach, J. F., Roles of active site residues and the HUH motif of the F plasmid TraI relaxase. *The Journal of Biological Chemistry*. **2007**, 282, (46), 33707-13.
14. Dyda, F.; Hickman, A. B., A Mob of Reqs. *Structure*. **2003**, 11, 1310-1311.
15. Koonin, E. V.; Ilyina, T. V., Computer-assisted dissection of rolling circle DNA replication. *BioSystems*. **1993**, 30, 241-268.
16. Monzingo, A. F.; Ozburn, A.; Xia, S.; Meyer, R. J.; Robertus, J. D., The structure of the minimal relaxase domain of MobA at 2.1 Å resolution. *Journal of Molecular Biology*. **2007**, 366, (1), 165-78.
17. Xia, S.; Robertus, J., Effect of divalent ions on the minimal relaxase domain of MobA. *Archives of Biochemistry and Biophysics*. **2009**, 488, 42-47.
18. Becker, E. C.; Meyer, R. J., MobA, the DNA strand transferase of plasmid R1162: the minimal domain required for DNA processing at the origin of transfer. *J Biol Chem* **2002**, 277, (17), 14575-80.
19. Bhattacharjee, M. K.; Meyer, R. J., Specific binding of MobA, a plasmid-encoded protein involved in the initiation and termination of conjugal DNA transfer, to single-stranded oriT DNA. *Nucleic Acids Research* **1993**, 21, (19), 4563-8.
20. Henderson, D.; Meyer, R., The MobA-linked primase is the only replication protein of R1162 required for conjugal mobilization. *J Bacteriol* **1999**, 181, (9), 2973-8.
21. Balzer, D., Pansegrau, W, Lanka, E, Essential Motifs of Relaxase (TraI) and TraG Proteins Involved in Conjugal Transfer of Plasmid RP4. *Journal of Bacteriology* **1994**, 176, (14), 4285-4295.
22. Pansegrau, W., Schroder, W, Lanka, E, Relaxase (TraI) of IncPα plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proceedings of the National Academy of Sciences* **1993**, 90, 2925-2929.
23. de la Cruz, F.; Frost, L. S.; Meyer, R. J.; Zechner, E. L., Conjugal DNA metabolism in Gram-negative bacteria. *FEMS Microbiology Reviews*. **2010**, 34, (1), 18-40.
24. Matson, S. W.; Ragonese, H., The F-plasmid TraI protein contains three functional domains required for conjugal DNA strand transfer. *Journal of Bacteriology* **2005**, 187, (2), 697-706.

25. Guogas, L. M.; Kennedy, S. A.; Lee, J.-H.; Redinbo, M. R., A novel fold in the TraI relaxase-helicase C-terminal domain is essential for conjugative DNA transfer. *Journal of Molecular Biology* **2009**, 386, 554-568.
26. Datta, S.; Larkin, C.; Schildbach, J. F., Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure*. **2003**, 11, 1369-1379.
27. Dostal, L.; Schildbach, J. F., Single-stranded DNA binding by F TraI relaxase and helicase domains is coordinating regulated. *Journal of Bacteriology*. **2010**, 192, (14), 3620-3628.
28. Dostal, L.; Shao, S.; Schildbach, J. F., Tracking F plasmid TraI relaxase processing reactions provides insight into F plasmid transfer. *Nucleic Acids Research* **2010**.
29. Hekman, K.; Guja, K.; Larkin, C.; Schildbach, J. F., An intrastrand three-DNA-base interaction is a key specificity determinant of F transfer initiation and of F TraI relaxase DNA recognition and cleavage. *Nucleic Acids Research* **2008**, 36, (14), 4565-72.
30. Ippen-Ihler, K. A., The conjugation system of F, the fertility factor of *Escherichia coli*. *Annual Review of Genetics* **1986**, 20, 593-624.
31. Matson, S. W.; Nelson, W. C.; Morton, B. S., Characterization of the reaction product of the oriT nicking reaction catalyzed by *Escherichia coli* DNA helicase I. *Journal of Bacteriology* **1993**, 175, (9), 2599-606.
32. Stern, J. C.; Schildbach, J. F., DNA recognition by F factor TraI36: highly sequence-specific binding of single-stranded DNA. *Biochemistry* **2001**, 40, (38), 11586-95.
33. Street, L. M.; Harley, M. J.; Stern, J. C.; Larkin, C.; Williams, S. L.; Miller, D. L.; Dohm, J. A.; Rodgers, M. E.; Schildbach, J. F., Subdomain organization and catalytic residues of the F factor TraI relaxase domain. *Biochimica et Biophysica Acta*. **2003**, 1646, (1-2), 86-99.
34. Willetts, N.; Skurray, R., The conjugation system of F-like plasmids. *Annual Review of Genetics* **1980**, 14, 41-76.
35. Williams, S. L.; Schildbach, J. F., Examination of an inverted repeat within the F factor origin of transfer: context dependence of F TraI relaxase DNA specificity. *Nucleic Acids Research*. **2006**, 34, (2), 426-35.
36. Boer, R.; Russi, S.; Guasch, A.; Lucas, M.; Blanco, A.; Perez-Luque, R.; Coll, M.; de la Cruz, F., 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. *Journal of Molecular Biology*. **2006**, 358, 857-869.

37. Gonzalez-Perez, B.; Lucas, M.; Cooke, L. A.; Vyle, J. S.; de la Cruz, F.; Moncalian, G., Analysis of DNA processing reactions in bacterial conjugation by using suicide oligonucleotides. *EMBO Journal*. **2007**, 26, (16), 3847-57.

38. Grandoso, G.; Avila, P.; Cayon, A.; Hernando, M. A.; Llosa, M.; de la Cruz, F., Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *Journal of Molecular Biology*. **2000**, 295, (5), 1163-72.

39. Guasch, A.; Lucas, M.; Moncalian, G.; Cabezas, M.; Perez-Luque, R.; Gomis-Ruth, F. X.; de la Cruz, F.; Coll, M., Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nature Structural Biology*. **2003**, 10, (12), 1002-10.

40. Llosa, M.; Bolland, S.; de la Cruz, F., Structural and functional analysis of the origin of conjugal transfer of the broad-host-range IncW plasmid R388 and comparison with the related IncN plasmid R46. *Molecular and General Genetics* **1991**, 226, (3), 473-483.

41. Llosa, M.; Grandoso, G.; de la Cruz, F., Nicking activity of TrwC directed against the origin of transfer of the IncW plasmid R388. *Journal of Molecular Biology* **1995**, 246, (1), 54-62.

42. Llosa, M.; Grandoso, G.; Hernando, M. A.; de la Cruz, F., Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *Journal of Molecular Biology*. **1996**, 264, (1), 56-67.

43. Lucas, M.; Gonzalez-Perez, B.; Cabezas, M.; Moncalian, G.; Rivas, G.; de la Cruz, F., Relaxase DNA binding and cleavage are two distinguishable steps in conjugative DNA processing that involve different sequence elements of the *nic* site. *The Journal of Biological Chemistry* **2010**.

44. Llosa, M.; Bolland, S.; de la Cruz, F., Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. *Journal of Molecular Biology* **1994**, 235, 448-464.

45. Paterson, E.; Iyer, V., Localization of the *nic* site of IncN conjugative plasmid pCU1 through formation of a hybrid *oriT*. *Journal of Bacteriology* **1997**, 179, (18), 5768-76.

46. Paterson, E.; More, M.; Pillay, G.; Cellini, C.; Woodgate, R.; Walker, G.; Iyer, V.; Winans, S., Genetic analysis of the mobilization and leading regions of the IncN plasmids pKM101 and pCU1. *Journal of Bacteriology* **1999**, 181, (8), 2572-83.

47. Willetts, N.; Wilkins, B., Processing of plasmid DNA during bacterial conjugation. *Microbiology Reviews*. **1984**, 48, (1), 24-41.
48. Llosa, M.; Zunzunegui, S.; de la Cruz, F., Conjugative coupling proteins interact with cognate and heterologous VerB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proceedings of the National Academy of Science USA* **2003**, 100, (18), 10465-10470.
49. Byrd, D. R.; Matson, S. W., Nicking by transesterification: the reaction catalysed by a relaxase. *Molecular Microbiology*. **1997**, 25, (6), 1011-22.
50. Pansegrau, W.; Lanka, E., Mechanisms of Initiation and Termination Reactions in Conjugative DNA Processing. *The Journal of Biological Chemistry*. **1996**, 271, (22), 13068-13076.
51. Lujan, S., Guogas, LM, Ragonese, H, Matson, SW, Redinbo, MR, Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proceedings of the National Academy of Science USA* **2007**, 104, 12282-12287.
52. Khatoon, H.; Iyer, R. V.; Iyer, V. N., A new filamentous bacteriophage with sex-factor specificity. *Virology*. **1972**, 48, 145-155.
53. Konarska-Kozłowska, M.; Iyer, V. N., Physical and genetic organization of the IncN-group plasmid pCU1. *Gene* **1981**, 14, (3), 195-204.
54. Paterson, E.; Iyer, V., The oriT region of the conjugative transfer system of plasmid pCU1 and specificity between it and the mob region of other N tra plasmids. *Journal of Bacteriology* **1992**, 174, (2), 499-507.
55. Hall, T., BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **1999**, 41, 95-98.
56. Thompson, J.; Higgins, D.; Gibson, T., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* **1994**.
57. Donnelly, M.; Zhou, M.; Millard, C.; Clancy, S.; Stols, L.; Eschenfeldt, W., An expression vector tailored for large-scale, highthroughput purification of recombinant proteins. *Protein Expression and Purification* **2006**, 47, 446-454.
58. Schwarzenbacher, R., The importance of alignment accuracy for molecular replacement. *Acta Crystallographica Section D Biological Crystallography* **2004**, 1229-1236.

59. Stein, N., CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *Journal of Applied Crystallography* **2008**, 41, 641-643.
60. Lo Conte, L.; Chothia, C.; Janin, J., The atomic structure of protein-protein recognition sites. *Journal of Molecular Biology* **1999**, 285, (5), 2177-2198.
61. Krissinel, E.; Henrick, K., Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **2007**, 372, 774-797.
62. Barthel, D., Hirst, JD, Blazewicz, J, Burke, EK, Krasnogor, N, ProCKSI: a decision support system for Protein (Structure) Comparison, Knowledge, Similarity and Information. *BMC Bioinformatics* **2007**, 8, 416.
63. Zhang, Y.; Skolnick, J., TM-align: A protein structure alignment algorithm based on TM-score. *Nucleic Acids Research* **2005**, 33, 2302-2309.
64. Greenfield, N., Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nature Protocols* **2006**, 1, (6), 2527-2535.
65. Moncalian, G.; de la Cruz, F., DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochimica et Biophysica Acta* **2004**, 1701, 15-23.
66. Nelson, W.; Howard, M.; Sherman, J.; Matson, S., The TraY gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid oriT. *The Journal of Biological Chemistry* **1995**, 270, 28374-28380.
67. Scherzinger, E.; Kruff, V.; Otto, S., Purification of the large mobilization protein of plasmid RSF1010 and characterization of its site-specific DNA-cleavage/DNA-joining activity. *FEBS Journal* **1993**, 217, 929-938.
68. Finney, L. A.; O'Halloran, T. V., Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **2003**, 300, (5621), 931-936.
69. Harding, M. M., Geometry of metal-ligand interactions in proteins. *Acta Crystallographica Section D Biological Crystallography* **2001**, 57, 401-411.
70. Lorenzo-Diaz, F.; Dostal, L.; Coll, M.; Schildbach, J. F.; Menendez, M.; Espinosa, M., The MobM relaxase domain of plasmid pMV158: thermal stability and activity upon Mn²⁺ and specific DNA binding. *Nucleic Acids Research* **2011**, Advance Access.

71. Singleton, M. R.; Dillingham, M. S.; Wigley, D. B., Structure and mechanism of helicases and nucleic acid translocases. *Annual Review of Biochemistry* **2007**, 76, 23-50.
72. Hall, M. C.; Matson, S. W., Helicase motifs: the engine that powers DNA unwinding. *Molecular Microbiology* **1999**, 34, (5), 867-877.
73. Matson, S. W., Dual engines moving on antiparallel tracks. *Nature Structural Biology*. **2003**, 10, (7), 499-500.
74. Grandoso, G.; Llosa, M.; Zabala, J. C.; de la Cruz, F., Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. *FEBS Journal* **1994**, 226, (2), 403-12.
75. Matson, S. W.; Sampson, J. K.; Byrd, D. R., F plasmid conjugative DNA transfer: the TraI helicase activity is essential for DNA strand transfer. *The Journal of Biological Chemistry*. **2001**, 276, (4), 2372-9.
76. Cheng, Y.; McNamara, D. E.; Miley, M. J.; Nash, R. P.; Redinbo, M. R., Functional characterization of the multi-domain F plasmid TraI relaxase-helicase. *The Journal of Biological Chemistry* **2011**, JBC Papers in Press.
77. Caruthers, J. M.; McKay, D. B., Helicase structure and mechanism. *Current Opinion in Structural Biology* **2002**, 12, 123-133.
78. Saikrishnan, K.; Griffiths, S. P.; Cook, N.; Court, R.; Wigley, D. B., DNA binding to RecD: role of the 1B domain in SF1B helicase activity. *The EMBO Journal* **2008**, 27, 2222-2229.
79. Saikrishnan, K.; Powell, B.; Cook, N. J.; Webb, M. R.; Wigley, D. B., Mechanistic basis of 5' - 3' translocation in SF1B helicases. *Cell* **2009**, 137, 849-859.
80. Lohman, T. M.; Bujalowski, W., Thermodynamic methods for model-independent determination of equilibrium binding isotherms for protein-DNA interactions: spectroscopic approaches to monitor binding. *Methods in Enzymology* **1991**, 208, (258-290).
81. Jezewska, M. J.; Marcinowicz, A.; Lucius, A. L.; Bujalowski, W., DNA polymerase X from African swine fever virus: quantitative analysis of the enzyme-ssDNA interactions and the functional structure of the complex. *Journal of Molecular Biology* **2006**, 356, 121-141.
82. Sikora, B.; Eoff, R. L.; Matson, S. W.; Raney, K. D., DNA unwinding by *Escherichia coli* DNA helicase I (TraI) provides evidence for a processive monomeric molecular motor. *The Journal of Biological Chemistry*. **2006**, 281, (47), 36110-36116.

83. Aggarwal, M.; Sommers, J. A.; Shoemaker, R. H.; Brosh, R. M., Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress. *Proceedings of the National Academy of Science USA* **2011**, 108, (4), 1525-1530.
84. Aiello, D.; Barnes, M. H.; Biswas, E. E.; Biswas, S. B.; Gu, S.; Williams, J. D.; Bowlin, T. L.; Moir, D. T., Discovery, characterization and comparison of inhibitors of *Bacillus anthracis* and *Staphylococcus aureus* replicative DNA helicases. *Bioorganic and Medicinal Chemistry* **2009**, 17, 4466-4476.
85. Chen, C.-S.; Chiou, C.-T.; Chen, G. S.; Chen, S.-C.; Hu, C.-Y.; Chi, W.-K.; Chu, Y.-D.; Hwang, L.-H.; Chen, P.-J.; Chen, D.-S.; Liaw, S.-H.; Chern, J.-W., Structure-based discovery of triphenylmethane derivatives as inhibitors of hepatitis C virus helicase. *Journal of Medicinal Chemistry* **2009**, 52, 2716-2723.
86. Chono, K.; Katsumata, K.; Kontani, T.; Kobayashi, M.; Suo, K.; Yokota, T.; Konno, K.; Shimizu, Y.; Suzuki, H., ASP2151, a novel helicase-primase inhibitor, possesses antiviral activity against varicella-zoster virus and herpes simplex virus types 1 and 2. *Journal of Antimicrobial Chemotherapy* **2010**, 65, (8), 1733-1741.
87. Crute, J. J.; Grygon, C. A.; Hargrave, K. D.; Simoneau, B.; Faucher, A.-M.; Bolger, G.; Kibler, P.; Liuzzi, M.; Cordingley, M. G., Herpes simplex virus helicase-primase inhibitors are active in animal models of human disease. *Nature Medicine* **2002**, 8, (4), 386-391.
88. Griep, M. A.; Blood, S.; Larson, M. A.; Koepsell, S. A.; Hinrichs, S. H., Myricetin inhibits *Escherichia coli* DnaB helicase but not primase. *Bioorganic and Medicinal Chemistry* **2007**, 15, 7203-7208.
89. Kandil, S.; Biondaro, S.; Vlachakis, D.; Cummins, A.-C.; Coluccia, A.; Berry, C.; Leyssen, P.; Neyts, J.; Brancale, A., Discovery of a novel HCV helicase inhibitor by a *de novo* drug design approach. *Bioorganic and Medicinal Chemistry Letters* **2009**, 19, 2935-2937.
90. Kleyman, G.; Fischer, R.; Betz, U. A. K.; Hendrix, M.; Bender, W.; Schneider, U.; Handke, G.; Eckenberg, P.; Hewlett, G.; Pevzner, V.; Baumeister, J.; Weber, O.; Henninger, K.; Keldenich, J.; Jensen, A.; Kolb, J.; Bach, U.; Popp, A.; Maben, J.; Frappa, I.; Haebich, D.; Lockhoff, O.; Rubsamen-Waigmann, H., New helicase-primase inhibitors as drug candidates for the treatment of herpes simplex disease. *Nature Medicine* **2002**, 8, (4), 392-398.
91. McKay, G. A.; Reddy, R.; Arhin, F.; Belley, A.; Lehoux, D.; Moeck, G.; Sarmiento, I.; Parr, T. R.; Gros, P.; Pelletier, J.; Far, A. R., Triaminotriazine DNA helicase inhibitors with antibacterial activity. *Bioorganic and Medicinal Chemistry Letters* **2006**, 16, 1286-1290.

92. Phoon, C. W.; Ng, P. Y.; Ting, A. E.; Yeo, S. L.; Sim, M. M., Biological evaluation of hepatitis C virus helicase inhibitors. *Bioorganic and Medicinal Chemistry Letters* **2001**, 11, 1647-1650.
93. Zhang, B.; Zhang, A.-h.; Chen, L.; Xi, X. G., Inhibition of DNA helicase, ATPase and DNA-binding activities of *E. coli* RecQ helicase by chemotherapeutic agents. *Journal of Biochemistry* **2008**, 143, 773-779.
94. Borowski, P.; Deinert, J.; Schalinski, S.; Bretner, M.; Ginalski, K.; Kulikowski, T.; Shugar, D., Halogenated benzimidazoles and benzotriazoles as inhibitors of the NTPase/helicase activities of hepatitis C and related viruses. *FEBS Journal* **2003**, 270, 1645-1653.
95. Gemma, S.; Butini, S.; Campiani, G.; Brindisi, M.; Zanolli, S.; Romano, M. P.; Tripaldi, P.; Savini, L.; Fiorini, I.; Borrelli, G.; Novellino, E.; Maga, G., Discovery of potent nucleotide-mimicking competitive inhibitors of hepatitis C virus NS3 helicase. *Bioorganic and Medicinal Chemistry Letters* **2010**, Available online.
96. Krawczyk, M.; Wasowska-Lukawska, M.; Oszczapowicz, I.; Boguszewska-Chachulska, A. M., Amidinoanthracyclines - a new group of potential anti-hepatitis C virus compounds. *Biological Chemistry* **2009**, 390, 351-360.
97. Najda-Bernatowicz, A.; Krawczyk, M.; Stankiewicz-Drogon, A.; Bretner, M.; Boguszewska-Chachulska, A. M., Studies on the anti-hepatitis C virus activity of newly synthesized tropolone derivatives: Identification of NS3 helicase inhibitors that specifically inhibit subgenomic HCV replication. *Bioorganic and Medicinal Chemistry* **2010**, 18, 5129-5136.
98. Kiianitsa, K.; Solinger, J. A.; Heyer, W.-D., NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Analytical Biochemistry* **2003**, 321, 266-271.
99. Kiianitsa, K.; Solinger, J. A.; Heyer, W.-D., NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzed enzymes with low and high specific activities. *Analytical Biochemistry* **2003**, 321, 266-271.
100. Jenkins, W. T., The pyruvate kinase-coupled assay for ATPases: a critical analysis. *Analytical Biochemistry* **1991**, 194, 136-139.
101. Charter, N. W.; Kauffman, L.; Singh, R.; Eglon, R. M., A generic, homogenous method for measuring kinase and inhibitor activity via adenosine 5'-diphosphate accumulation. *Journal of Biomolecular Screening* **2006**, 11, 390-399.

102. Zhang, J.-H.; Chung, T. D. Y.; Oldenburg, K. R., A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of Biomolecular Screening* **1999**, 4, 67-73.
103. Frost, L. S.; Koraimann, G., Regulation of bacterial conjugation: balancing opportunity with adversity. *Future Microbiology* **2010**, 5, (7), 1057-1071.
104. Gowland, P. C.; Slater, J. H., Transfer and stability of drug resistance plasmids in *Escherichia coli* K12. *Microbial Ecology* **1984**, 10, (1), 1-13.
105. Grohmann, E.; Muth, G.; Espinosa, M., Conjugative plasmid transfer in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews* **2003**, 67, (2), 277-301.
106. DeNap, J. C. B.; Thomas, J. R.; Musk, D. J.; Hergenrother, P. J., Combating drug-resistant bacteria: small molecule mimics of plasmid incompatibility as antiplasmid compounds. *Journal of the American Chemistry Society* **2004**, 126, 15402-15404.
107. Fernandez-Lopez, R.; Machon, C.; Longshaw, C.; Martin, S.; Molin, S.; Zechner, E.; Espinosa, M.; Lanka, E.; de la Cruz, F., Unsaturated fatty acids are inhibitors of bacterial conjugation. *Microbiology* **2005**, 151, 3517-3626.
108. Garcillan-Barcia, M.; Jurado, P.; Gonzalez-Perez, B.; Moncalian, G.; Fernandez, L.; de la Cruz, F., Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. *Molecular Microbiology* **2007**, 63, (2), 404-416.
109. Marraffini, L. A.; Sontheimer, E. J., CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **2008**, 322, 1843-1845.
110. Thatte, V.; Bradley, D.; Iyer, V., N conjugative transfer system of plasmid pCU1. *Journal of Bacteriology* **1985**, 163, (3), 1229-36.
111. Alonso, G.; Baptista, K.; Ngo, T.; Taylor, D. E., Transcriptional organization of the temperature-sensitive transfer system from the IncH1 plasmid R27. *Microbiology* **2005**, 151, 3563-3573.
112. Krishnan, B. R.; Iyer, V. N., Host ranges of the IncN group plasmid pCU1 and its minireplicon in Gram-negative purple bacteria. *Applied and Environmental Microbiology* **1988**, 54, (9), 2273-2276.

113. Andrup, L.; Andersen, K., A comparison of the kinetics of plasmid transfer in the conjugation system encoded by the F plasmid from *Escherichia coli* and plasmid pCF10 from *Enterococcus faecalis*. *Microbiology* **1999**, 145, 2001-2009.
114. Brasch, M. A.; Meyer, R. J., Genetic organization of plasmid R1162 DNA involved in conjugative mobilization. *Journal of Bacteriology* **1986**, 167, (2), 703-710.
115. Havekes, L.; Hoekstra, W.; Kempen, H., Relation between F, R1, R100 and R144 *Escherichia coli* K12 donor strains in mating. *Molecular and General Genetics* **1977**, 155, 185-189.
116. Sanderson, K. E.; Janzer, J.; Head, J., Influence of lipopolysaccharide and protein in the cell envelope on recipient capacity in conjugation of *Salmonella typhimurium*. *Journal of Bacteriology* **1981**, 148, (1), 283-293.
117. Sengelov, G.; Sorensen, S. J., Methods for detection of conjugative plasmid transfer in aquatic environments. *Current Microbiology* **1998**, 37, 274-280.