

# The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1

Rebekah Potts Nash<sup>1</sup>, Sohrab Habibi<sup>1</sup>, Yuan Cheng<sup>2</sup>, Scott A. Lujan<sup>2</sup> and Matthew R. Redinbo<sup>1,2,\*</sup>

<sup>1</sup>Department of Chemistry, University of North Carolina, Chapel Hill, CB 3290 and <sup>2</sup>Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, CB 7260, Chapel Hill, NC 27599, USA

Received February 9, 2010; Revised April 9, 2010; Accepted April 12, 2010

## ABSTRACT

**Bacteria expand their genetic diversity, spread antibiotic resistance genes, and obtain virulence factors through the highly coordinated process of conjugative plasmid transfer (CPT). A plasmid-encoded relaxase enzyme initiates and terminates CPT by nicking and religating the transferred plasmid in a sequence-specific manner. We solved the 2.3 Å crystal structure of the relaxase responsible for the spread of the resistance plasmid pCU1 and determined its DNA binding and nicking capabilities. The overall fold of the pCU1 relaxase is similar to that of the F plasmid and plasmid R388 relaxases. However, in the pCU1 structure, the conserved tyrosine residues (Y18,19,26,27) that are required for DNA nicking and religation were displaced up to 14 Å out of the relaxase active site, revealing a high degree of mobility in this region of the enzyme. In spite of this flexibility, the tyrosines still cleaved the nic site of the plasmid's origin of transfer, and did so in a sequence-specific, metal-dependent manner. Unexpectedly, the pCU1 relaxase lacked the sequence-specific DNA binding previously reported for the homologous F and R388 relaxase enzymes, despite its high sequence and structural similarity with both proteins. In summary, our work outlines novel structural and functional aspects of the relaxase-mediated conjugative transfer of plasmid pCU1.**

## INTRODUCTION

Antibiotics have effectively treated a broad range of bacterial infections for years. However, with the emergence of

antibiotic resistance, the treatment of these same bacterial infections is becoming increasingly difficult (1) ([http://www.cdc.gov/hicpac/mdro/mdro\\_toc.html](http://www.cdc.gov/hicpac/mdro/mdro_toc.html)). Currently, 70% of all hospital acquired infections are resistant to at least one antibiotic, and antibiotic resistance to at least one drug has been reported for every major strain of pathogenic bacteria (2). As a result, the Centers for Disease Control has declared antibiotic resistance a primary threat to human health, and multiple efforts have been made to understand and inhibit the mechanisms responsible for its spread (2–8).

Conjugative plasmid transfer (CPT) is a primary method by which bacteria acquire antibiotic resistance genes. During CPT, one bacterium (the donor) transfers genetic material, in the form of a conjugative DNA plasmid, to a neighboring cell (the recipient) via physical cell–cell contact. Each conjugative plasmid encodes the majority of proteins found within the relaxosome, a multi-protein complex necessary to promote plasmid transfer in a sequence-specific manner. The relaxosome works in concert with a type IV secretion system (T4SS), and these two complexes are linked by a type IV coupling protein (T4CP), believed to form the pore through which the T (transferred) strand of the plasmid travels into the recipient bacterium (9–13).

One component of the relaxosome, the relaxase, is responsible for the initiation and termination of T-strand transfer (14,15). A generalized mechanism describing the role of the relaxase was first determined through the study of the F (fertility) plasmid system (14–27), and further details were contributed by the study of other conjugative systems, including plasmids R388 (28–37), RP4 (38–40), R1162 (41–46) and R100 (47,48). To initiate transfer, the relaxase binds the origin of transfer (oriT) of the plasmid's T-strand and cleaves the strand at the nic site. After the T-strand has been

\*To whom correspondence should be addressed. Tel: +1 919 843 8910; Fax: +1 919 962 2388; Email: redinbo@unc.edu  
Present address:

Scott A. Lujan, Laboratories of Molecular Genetics and Structural Biology, 111 T. W. Alexander Drive, National Institute of Environmental Health Sciences, Institutes of Health, DHHS, Research Triangle Park, NC, 27709, USA.

unwound from the parent strand, the relaxase terminates transfer by religating the nicked strand (14).

The structures of the F TraI relaxase, the R388 TrwC relaxase and the R1162 MobA relaxase have been reported previously (3,23,24,33,34,44). In addition, the F TraI and R388 TrwC relaxase structures have been resolved in complex with their respective DNA substrates, revealing that the relaxase binds the T-strand oriT as a partial DNA hairpin (33,34). The 3' end of the oriT extends beyond the hairpin as a single strand and forms a U-turn around the protein N-terminus, before entering the active site of the enzyme (24,33,34). The active site is located on two strands of a central beta sheet and contains three histidines coordinating a metal cation, the identity of which has been debated (3,24,34,46). The scissile phosphate of the DNA substrate's nic site is positioned above the chelated 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 tyrosine residues (3,24,34). All three relaxase enzymes bind and cleave their specific DNA substrate with high affinity and sequence-specificity (22,33,41).

We have examined the atomic structure and the DNA binding and nicking activities of the relaxase enzyme responsible for the transfer of the resistance plasmid pCU1, originally isolated from *Salmonella typhimurium* (49–52). pCU1 is a derivative of resistance plasmid R46 and confers on its host resistance to the antibiotics ampicillin, streptomycin and spectinomycin (52,53). 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 (51,52). Our analysis of the pCU1 TraI relaxase has identified structural and functional characteristics unique to the pCU1 relaxase. These results advance our understanding of both the general mechanism of relaxase-mediated DNA binding and cleavage, as well as the unique nature of each CPT system.

## MATERIALS AND METHODS

### Relaxase constructs

The relaxase of plasmid pCU1 is located within the pCU1-encoded protein TraI (51,52). To determine the extent of the relaxase domain within pCU1 TraI, we aligned the pCU1 TraI amino acid sequence (GenBank: AAD27542) with those of R388 TrwC (GenBank: CAA44853) and F TraI (GenBank: BAA97974). Additional details describing how the alignments were performed can be found in the Supplementary Data. Residues 1 to ~300 of pCU1 TraI corresponded to the relaxase domains of F TraI and R388 TrwC. Secondary-structure predictors (Jpred3, version 2.2, <http://www.compbio.dundee.ac.uk/www-jpred/>) identified loop regions near residue 300, and the pCU1 TraI relaxase construct extending from residue 1 to 299 was chosen as the final construct. 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), 2009]. 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. Several additional relaxase constructs containing tyrosine to phenylalanine mutations were generated by Quick Change site directed mutagenesis (Stratagene) of WT\_299. Using ligation independent cloning (LIC), residues 1–299 of the relaxase domain were inserted into vector pMCSG9 (54) to create the mutant construct Nterm\_299. In vector pMCSG9, the relaxase was N-terminally fused to a maltose binding protein (MBP) affinity tag and a 6-His (MBP-HIS) affinity tag with a TEV protease cleavage site located between the protein sequence and affinity tags (54). During purification, these affinity tags were removed by the tobacco etch virus (TEV) protease, leaving three non-native residues (Ser–Asn–Ala) at the N-terminus of Nterm\_299.

### Protein expression and purification

All proteins were over-expressed in BL21 *Escherichia coli*, and the cells were harvested by centrifugation and resuspended in 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 Na<sub>2</sub>H(PO<sub>4</sub>) pH 8.0, 10% glycerol, 5 mM imidazole, 0–1 mM EDTA, 0.01% azide]. Resuspended cells were stored at –80°C. Additional details can be found in the Supplementary Data.

Prior to protein purification, cells were thawed in the presence of lysozyme and a cocktail of protease inhibitors, lysed on ice using a Sonic Dismembrator, Model 500 (Fisher Scientific), and centrifuged to isolate the soluble fraction. WT\_299 was purified on Chitin Resin (NEB) using a batch bind method followed by an extended wash step, 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 for 16 h. WT\_299 was then eluted from the Chitin Resin. Buffer C was used throughout the purification. Nterm\_299 was purified over a 5 ml HisTrap HP column (GE Healthcare) on an Aktaexpress FPLC (GE Healthcare). Fusion protein was eluted with a high imidazole buffer B [500 mM NaCl, 20 mM Na<sub>2</sub>H(PO<sub>4</sub>) pH 8.0, 10% glycerol, 500 mM imidazole, 0–1 mM EDTA, 0.01% azide] after washing the column to baseline 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 and further purified over two sequential 5 ml HisTrap HP gravity columns. The cleaved protein that flowed through the final column was collected. All proteins were visualized by SDS–PAGE and found to be >95% pure (Supplementary Figure S1).

For crystallization, WT<sub>299</sub> and Nterm<sub>299</sub> were dialyzed into 250 mM ammonium acetate while concentrating to 5–6 mg/ml (Nterm<sub>299</sub>) or ~14 mg/ml (WT<sub>299</sub>), then used immediately for crystallization. For analysis by circular dichroism (CD) spectroscopy, WT<sub>299</sub> and Nterm<sub>299</sub> were dialyzed into buffer CD (100 mM NaF, 20 mM Tris–HCl pH 7.5, 5% glycerol, 0.01% azide) and then used immediately. For DNA binding, nicking and cross-over assays with WT<sub>299</sub>, an additional size exclusion chromatography purification step was performed. During this step, WT<sub>299</sub> was loaded onto a HiLoad 16/60 Superdex 200 column (GE Healthcare) pre-equilibrated with buffer S (100 mM NaCl, 20 mM Tris–HCl pH 7.5, 5% glycerol, 0–5 mM EDTA, 0.01% azide) on an Atkxpress FPLC (GE Healthcare). The protein was eluted from the column in buffer S. When necessary, any EDTA present was removed by dialysis of WT<sub>299</sub> and Nterm<sub>299</sub> into buffer D (100 mM NaCl, 20 mM Tris–HCl pH 7.5, 5% glycerol, 0.01% azide). For inductively coupled plasma mass spectrometry (ICP-MS) analysis, a separate preparation of WT<sub>299</sub> and Nterm<sub>299</sub> was performed in order to prevent contaminating the protein sample with metals from outside sources. All water was obtained from the laboratory Barnstead E-pure water filtration system, at >17 megaohm-cm (ddH<sub>2</sub>O). Following purification by size-exclusion chromatography, the sample was then dialyzed into buffer ICP (20 mM NaCl, 10 mM Tris–HCl pH 7.5, 0.01% azide).

### Protein crystallization

Crystals of WT<sub>299</sub> and Nterm<sub>299</sub> were grown at 20°C by the hanging-drop vapor diffusion method. Equal volumes of purified protein and well solution (250 mM triNa citrate, 22% PEG 3350) were mixed and crystals grew over a course of 4–5 weeks. Crystals were initially small and hexagonal, but following micro-seeding and the addition of 5 mM MgCl<sub>2</sub> 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.

### Data collection, structure determination and refinement

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 either the Southeast Regional Collaborative Access Team (SER-CAT) Sector 22-ID beam-line at the Advanced Photon Source (APS) at Argonne National Laboratory or the Stanford Synchrotron Radiation Laboratory (SSRL) using remote access of beam-line 9-2 ID. Data from the best diffracting WT<sub>299</sub> crystals were indexed and scaled with X-ray Detector Software (XDS, MPI for Medical Research, Heidelberg, 2009, <http://xds.mpimf-heidelberg.mpg.de/>). Data from the best diffracting Nterm<sub>299</sub> crystals were indexed and scaled with HKL2000 (HKL Research Inc., 2005, <http://www.hkl-xray.com>). Initial phases were determined by molecular replacement in Phaser ([\[.cam.ac.uk/phaser\]\(http://www-structmed.cimr.cam.ac.uk/phaser\)\), 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 \(PDB code 1omh\). Data were refined using remlac \(<http://www.ysbl.york.ac.uk/~garib/remlac/>\), 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 \(<http://biop.ox.ac.uk/coot>\). 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 \(<http://davapc1.bioch.dundee.ac.uk/prodrgr/>\). Data were verified using MolProbity \(<http://molprobity.biochem.duke.edu>\). Figures were constructed in PyMOL \(DeLano Scientific LLC, San Carlos, CA, USA, <http://www.pymol.org>, 2009\). AREAIMOL, 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, \[http://www.ebi.ac.uk/msd-srv/prot\\\_int/pistart.html\]\(http://www.ebi.ac.uk/msd-srv/prot\_int/pistart.html\)\). Structural alignments were performed with TM-align \(<http://zhanglab.ccmb.med.umich.edu/TM-align>\), using the ProCKSI-Server \(version procksi-8.7, \[www.procksi.net\]\(http://www.procksi.net\)\), to generate the root mean square deviation \(RMSD\) of C \$\alpha\$  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<sub>299</sub> and Nterm<sub>299</sub> have been deposited in the PDB \(accession codes 3L57, 3L6T\).](http://www-structmed.cimr</a></p>
</div>
<div data-bbox=)

### Inductively coupled plasma mass spectrometry (ICP-MS)

WT<sub>299</sub> and Nterm<sub>299</sub> were analyzed by ICP-MS to determine the identity and concentration of metals present in the samples. 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<sub>3</sub>. Each sample was then diluted in 10 ml of 1.4% HNO<sub>3</sub>. A control consisting of each buffer used during protein purification was analyzed to determine the concentration of metal in the buffers used. Each experimental sample, as well as the buffer control, was monitored for Mg24, Mn55, Ca44, Fe57, Ni60 and Zn66. Data were collected as part per billion (ppb) and converted to  $\mu$ M. The original concentration of each metal in each sample was then determined in Excel 2007 (Microsoft, 2006). Data were then presented as fold increase in metal concentration over background levels, where background values represented the signal due to 1.4% HNO<sub>3</sub> alone.

### DNA binding assays

The affinity of WT<sub>299</sub> and Nterm<sub>299</sub> for a panel of DNA substrates was determined by fluorescence anisotropy-based DNA binding assays. 5' fluorescein-labeled DNA substrates were commercially synthesized



and HPLC purified (IDT). Table 1 provides the nucleic acid sequence of each substrate investigated. Each substrate was resuspended in buffer R (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.05 mM EDTA), heated to 95°C for 10 min, and then allowed to slowly cool to room temperature. Each stock was then diluted to 0.1 μM in ddH<sub>2</sub>O prior to use. Additional details on the assay protocol and data analysis can be found in the Supplementary Data.

### DNA nicking and DNA cross-over assays

The DNA nicking and religation activity of WT<sub>299</sub> and Nterm<sub>299</sub> for a panel of DNA substrates was determined by DNA nicking and DNA cross-over assays. 5' fluorescein-labeled and unlabeled DNA substrates were commercially synthesized (IDT); 5' fluorescein-labeled DNA substrates were HPLC purified. Table 2 provides the nucleic acid sequence of each substrate investigated. Each substrate was resuspended in buffer R, heated to 95°C for 10 min, and then either allowed to cool passively to room temperature or snap cooled on ice. Both preparations yielded similar results. Each stock was then diluted to 10 μM in ddH<sub>2</sub>O prior to use. Additional details on the assay protocol can be found in the Supplementary Data.

To test the metal dependence of the pCU1 relaxase, WT<sub>299</sub> (10 μM) was treated with 10 mM EDTA for 24 h. The protein was then diluted either 2- or 10-fold upon the addition of 1 μM 35/7oriT-hairpin, and 0–15 mM metal was added to each reaction to generate an excess of 0–10 mM metal. Stocks (10×) of all metals investigated were made by dissolving the corresponding metal salt [MgCl<sub>2</sub>, MnCl<sub>2</sub>, CaCl<sub>2</sub>, NiCl<sub>2</sub>, Cu(SO<sub>4</sub>), Zn(acetate)<sub>2</sub> and ZnCl<sub>2</sub>] in 1 M Tris-HCl pH 7.5.

DNA nicking activity and DNA religation activity were reported as percent product generated (the intensity of the product band divided by the sum of product and substrates band intensities).

Each data point represents the average of three or more reactions and error bars represent the standard error of these data points. 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).

## RESULTS

### Crystal structure of the pCU1 TraI relaxase domain

The 2.3 Å structure of WT<sub>299</sub> was determined by X-ray crystallography. Crystals were grown by hanging-drop vapor diffusion, and the resulting structure was determined to 2.3 Å by molecular replacement using the R388 TrwC relaxase [PDB code 1omh (33)] as the search model (Table 3). The protein crystallized as a dimer in the asymmetric unit (asu) in the space group P2<sub>1</sub>. The two monomers were determined to have a combined 20 600 Å<sup>2</sup> of solvent accessible surface area of which 1670 Å<sup>2</sup> was buried upon dimerization. While this analysis could suggest a physiologically relevant dimer (55), only 18 potential hydrogen bonds and four 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 packing (56). Indeed, the protein was determined to be a monomer by size exclusion chromatography (Supplementary Figure S2). The electron density of monomer A was consistently higher in quality than that of monomer B. As a result, out of the protein's 299 residues, 223 were confidently modeled into the electron density of monomer A, while only 203 were placed into the electron density of the B monomer (Table 3).

The overall fold of the pCU1 TraI relaxase domain (Figure 1A) resembles that of a human hand, as seen in previously solved relaxase structures (3,23,24,33,34,44). The palm consists of five central anti-parallel β-strands, and 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<sub>299</sub> are disordered, but are predicted to be α-helical (Figure 2).

**Table 1.** DNA binding assay substrates

Substrate	Nucleic Acid Sequence	K <sub>D</sub> (nM)	SE	R <sub>2</sub>
35/7oriT-hairpin	T G T G A T <b>A G C G T G</b> A T T T A T <b>C G C G C T</b> G C G T T A G G T G T <b>A T A G C A G</b>	840	50	0.978
35oriT-hairpin	T G T G A T <b>A G C G T G</b> A T T T A T <b>C G C G C T</b> G C G T T A G G T G T	890	60	0.976
35oriT-no_hairpin	<b>C T G T A G T A G T A T G A C T A T</b> <b>C G C G C T</b> G C G T T A G G T G T	810	50	0.994
20oriT-half_hairpin	T A T <b>C G C G C T</b> G C G T T A G G T G T	960	50	0.997
35mer-hairpin	<b>C T A G C T</b> <b>C C G A G C</b> A T A A G A <b>G C T C G G</b> A C T A C C 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	<b>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</b>	1180	70	0.995
24mer	<b>C G A G C C A G C G A G C G A G C G A A C G C G</b>	850	70	0.991
19mer	<b>C A G C G A G C G A G C G A A C G C G</b>	750	50	0.995
10mer	<b>A G C G A A C G C G</b>	1900	110	0.996

The nucleic acid sequence, secondary structure and pCU1 TraI relaxase binding affinity, as determined by fluorescence anisotropy-based binding assays, for each substrate are provided. The names of substrates encoding a wild-type nic site and minimal oriT include the phrase 'oriT'. The names of mutant substrates include the phrase 'mer'. Mutated bases are in red font. Inverted repeats predicted to form a hairpin are boxed in black. The predicted nic site for each substrate (if present) is indicated by vertical double black lines. The DNA binding affinity of WT<sub>299</sub> for each substrate is presented as the dissociation constant (K<sub>D</sub>), as calculated by Equation 1, and the error represents the standard error of the K<sub>D</sub> when calculated by Equation 1. Additional details can be found in the Materials and Methods section and the Supplementary Data.

**Table 2.** DNA nicking and DNA cross-over assay substrates

	Nicking Substrate	Nucleic Acid Sequence		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 G T A T A G C A G		23.07	3.46
	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 G T A T A G C A G		45.98	1.69
	20/7oriT-half_hairpin	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		32.56	3.33
	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
	Cross-over Substrate	Nucleic Acid Sequence	Mis-Nicking	Activity (%)	S.E.
Hairpin Mutants	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		28.76	3.69
	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 G T A T A G C A G		29.08	1.84
	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 A G G C T T G C G T T A G G T G T A T A G G A G		38.50	0.7
Hairpin + TAG Mutants	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 T G T A T A G 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 G T A T A G G A G		31.49	0.33
	F_plasmid_oriT-G1'A	A A A A T C A G C A A A A A C T T G J T T T T G C G T G G G G T G T A T A G C T T		73.50	2.92
TAG Mutants	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 G T A T A G C A G		25.70	1.17
Pentanucleotide Mutants	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 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 T G T T A G G A G		1.44	1.3
	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 T G A T A G C 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 T G T A T A G C A G		46.03	2.08
	35/7oriT-A1'T	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 T A G C A G		30.23	3.06
Post Nic Mutants	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
Multiple Regions Mutated	35/7oriT-4_region_mutant	C T A G C A T C C G A G C A T A A G A G C T C G G A G T A G 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 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 A C T T G T T T T T G C G T G G G G T G T G G T 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 A G 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 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 T G G T G T G T A T G T A T A	yes	25.07	1.16	

The nucleic acid sequence, secondary structure and nicking or religation activity of WT\_299, as measured by DNA nicking and cross-over assays, for each substrate are provided. Cross-over assay substrates are listed according to the region of the pCU1 oriT which is mutated (see Figure 6 for a definition of these regions). Mutated bases are in red font. Light orange shading emphasizes the key areas of each mutant substrate relative to the wild-type substrate. Inverted repeats 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 for each substrate is indicated by vertical double black lines, and non-canonical nic sites, as observed in DNA cross-over assays, are indicated by vertical double red lines. To determine the DNA nicking or religation activity of the wild-type pCU1 TraI relaxase, 5 μM enzyme was incubated with 1 μM labeled substrate and, for cross-over assays, 1 μM unlabeled substrate for 1 h at 37°C; products were then separated on denaturing polyacrylamide gels. ‘Misnicking’ activity is reported for each substrate that generated a product during 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. Enzyme activity is represented as percent product generated. Each data point is the average of at least three experiments, and the error is the standard error of these replicates. Additional details can be found in the Materials and Methods section and the Supplementary Data.

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 (24,34).

The active site of WT\_299 is found within the palm of the hand, and consists of three histidines, two on β7 (His160, 162) and one on β6 (His149) (Figure 1B). Together, they chelate a divalent metal cation, occupying three of the bound metal’s coordination sites. The remaining coordination sites surrounding the divalent cation can be satisfied by a variety of compounds, such as water molecules or components from the cryoprotectant or well solution. In the structure of WT\_299, these remaining sites are 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 chelated by the histidines. Mn<sup>2+</sup>, Fe<sup>2+</sup>, Ni<sup>2+</sup>, Cu<sup>2+</sup> and Zn<sup>2+</sup> all satisfied the electron density as assessed by difference density and

R-factors. Mn<sup>2+</sup> was later determined to be the most likely metal in the active site based on a series of experiments described below.

The four tyrosines of WT\_299 implicated in DNA nicking (Tyr18,19,26,27) are located on loop A, adjacent to the active site (Figure 1B). In monomer B, only the first few residues of the loop were sufficiently ordered to be modeled into the structure. In monomer A, however, the entire loop was observed, and it was located in a unique orientation relative to previous relaxase structures. Tyr18 and Tyr19 were found displaced from the active site, flipped out and pointing away from the bound metal ion (Figure 1C). In the F TraI relaxase structure, the equivalent tyrosine residues are oriented towards the active site, with Tyr16 (equivalent to pCU1 relaxase Tyr18) positioned to attack the scissile phosphate of the DNA substrate (3,23,24). The equivalent tyrosines (Tyr18,19) in the R388 TrwC relaxase are also directed towards the active site, with Tyr18 shifted slightly behind the scissile phosphate, relative to Tyr16 of the F

**Table 3.** Crystallographic statistics

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 <sub>1</sub>	P2 <sub>1</sub>
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 <sub>sym</sub> <sup>a</sup> (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 (2.29-2.40)	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 <sub>work</sub> <sup>b</sup> (highest shell)	0.173 (0.203)	0.168 (0.235)
R <sub>free</sub> <sup>c</sup> (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
RMDSs		
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

<sup>a</sup>R<sub>sym</sub> =  $\sum |I - I_{\text{mean}}| / \sum I$ , where  $I$  is the observed intensity and  $I_{\text{mean}}$  is the average intensity of several symmetry-related observations.

<sup>b</sup>R<sub>work</sub> =  $\sum |F_o - F_c| / \sum F_o$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factors, respectively.

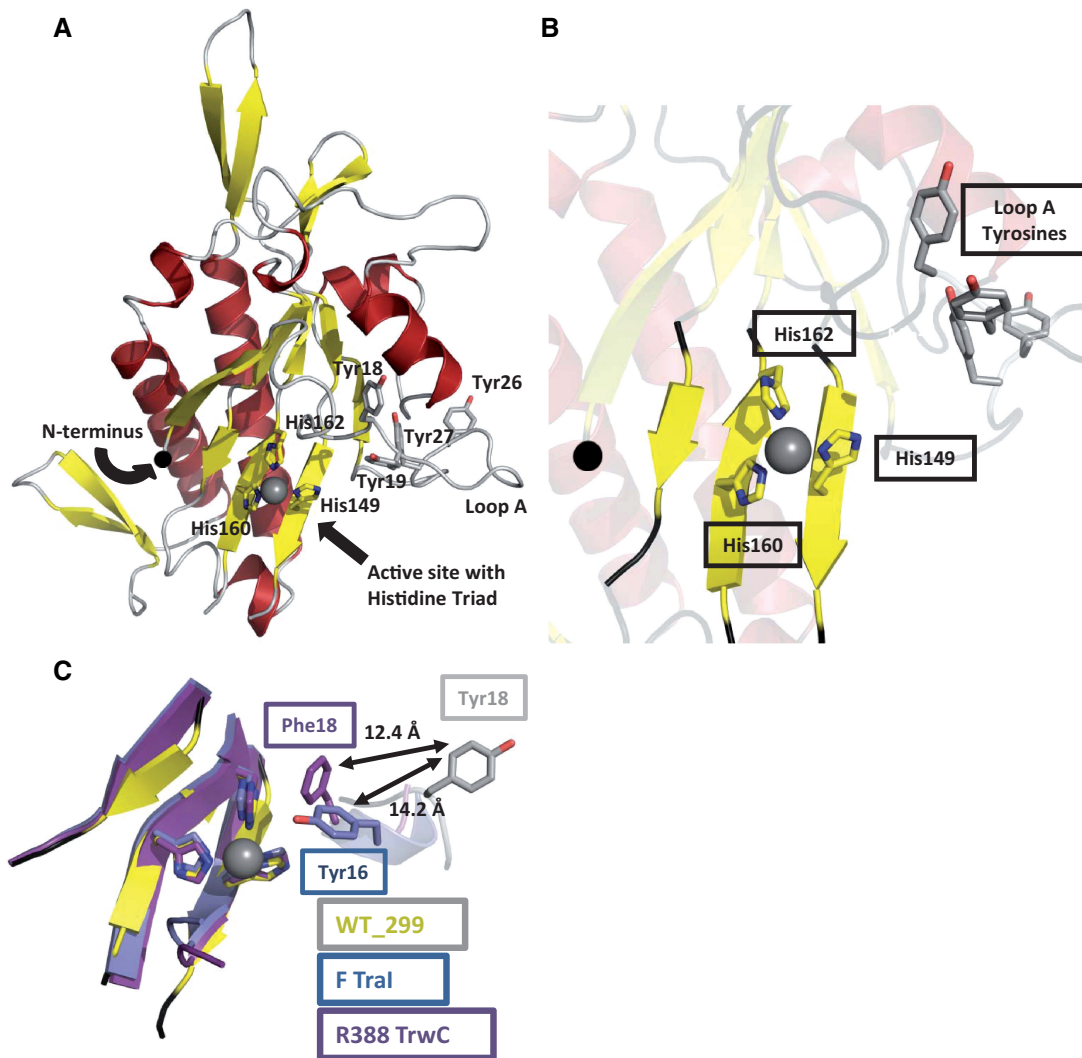
<sup>c</sup>R<sub>free</sub> = Calculated as above for 5% of data not used in any step of refinement.

TraI (33,34). When the structure of WT\_299 monomer A was overlaid with the R388 TrwC relaxase [PDB ID 2cdm (34)] and F TraI relaxase [PDB ID 1p4d (23)] (Figure 1C), Tyr18 of WT\_299 was displaced 14.2 Å away from the histidine triad when compared to Tyr16 of F TraI and 12.4 Å when compared to Tyr18 of R388 TrwC.

#### DNA binding by the pCU1 TraI relaxase

The relaxase binds the oriT of its DNA plasmid's T-strand prior to cleaving the T-strand's nic site (22,25,33). We

determined the binding affinity and specificity of the pCU1 TraI relaxase for a panel of DNA substrates (Table 1) using fluorescence anisotropy-based DNA binding assays. The effect of DNA substrate length, secondary structure and sequence on relaxase binding affinity was investigated. For all assays, a hyperbolic binding curve was generated by measuring the change in fluorescence anisotropy of the labeled DNA substrate as a function of enzyme concentration. By fitting these curves with Equation 1 (see Supplementary Data) using



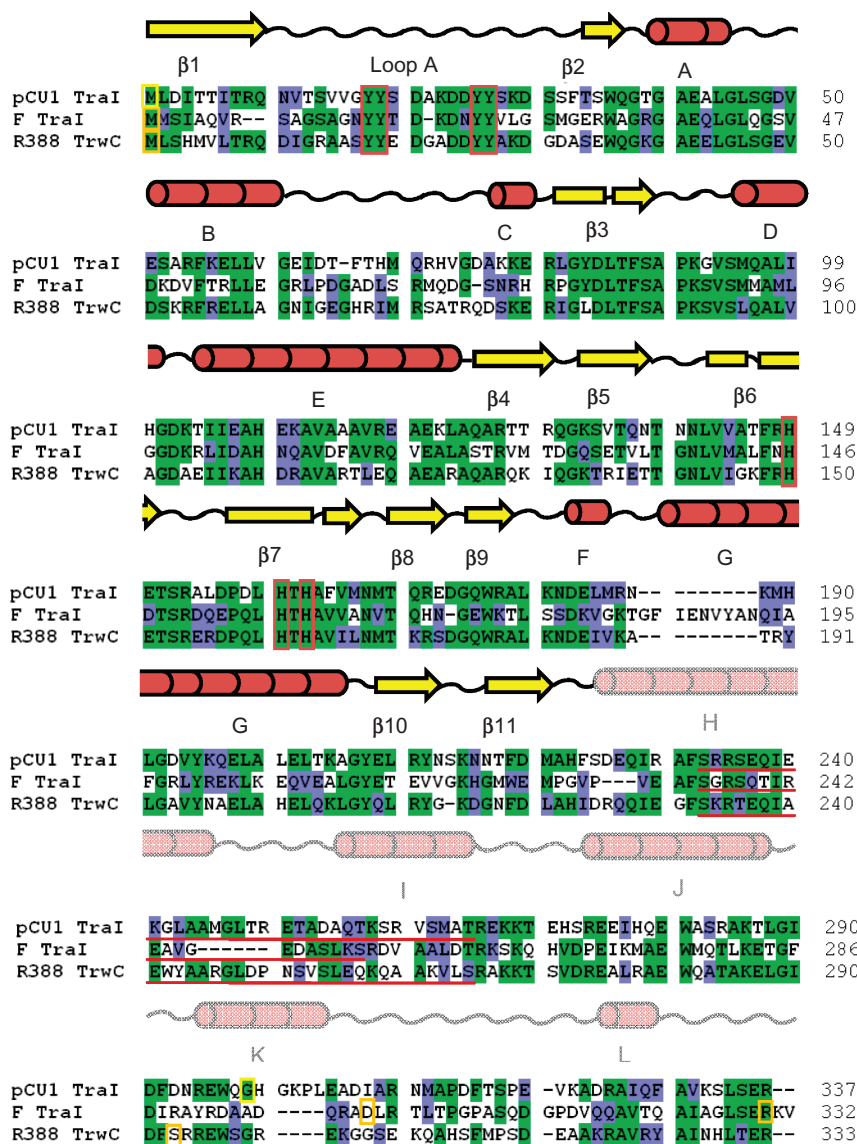
**Figure 1.** The 2.3 Å resolution crystal structure of the pCU1 TraI relaxase. (A) The overall structure of the pCU1 relaxase (WT\_299) with secondary-structures colored as represented in Figure 2 ( $\beta$ -strands are represented by yellow arrows,  $\alpha$ -helices by red helices and loop regions by gray loops). The histidine triad of the active site and the tyrosines implicated in DNA nicking are shown as sticks. The manganese ion bound by the histidine triad and bound manganese ion is shown as a grey sphere. (B) The active site of the pCU1 relaxase (WT\_299) with secondary structures represented as in (A). At the center is the histidine triad and bound manganese ion. In the upper right is loop A and the four tyrosines implicated in DNA nicking, and to the left is the N-terminus, represented as a black sphere. (C) The active site of the pCU1 relaxase [WT\_299, colored as described in (A)] aligned with the active sites of the F plasmid TraI relaxase (in blue) and the relaxase of plasmid R388 TrwC (in magenta). The histidine triad and the primary DNA nicking tyrosine of each protein are shown as sticks. The distances between the tyrosines are indicated. Note that in the structure of the plasmid R388 TrwC, Tyr18 was mutated to Phe18.

non-linear regression, we calculated dissociation constants ( $K_D$ ) for each experiment and discovered that both WT\_299 and Nterm\_299 bound DNA substrates in an oriT- and hairpin-independent manner. The binding affinity of both WT\_299 and Nterm\_299 was similar for all substrates investigated, and varied between 700 nM and 1.2  $\mu$ M (Table 1, 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 (22,25,33,41). WT\_299 was confirmed to be folded and stable in solution by circular dichroism spectroscopy (Supplementary Figures S3 and S4), and the design of the WT\_299 construct incorporated the full extent of the relaxase domain of pCU1 TraI (Figure 2).

Therefore, the design of the DNA binding assay was modified to determine if the experimental conditions for WT\_299 DNA binding could be further optimized. When incubating the Y18,19,26,27F mutant of the pCU1 relaxase with a DNA substrate that extended beyond the nic site (35/7oriT-hairpin), no improvement in DNA binding affinity or selectivity was observed, and when supplementing the reaction conditions with 4 mM  $MnCl_2$  there was only a 2- to 3-fold increase in DNA binding affinity (Figure 3A and B). 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





**Figure 2.** Amino acid sequence alignment and secondary structure of the pCU1 TraI relaxase. The initial 337 amino acids of pCU1 TraI (GenBank ID AAD27542) 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 and yellow 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 (3,23), therefore the terminal residues of both are boxed. Red boxes indicate the location of the conserved tyrosine residues and the active site histidine triad. Red lines underline the residues forming the '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.  $\beta$ -strands are represented by yellow arrows;  $\alpha$ -helices by red cylinders; and loop regions by black wavy lines. The secondary structure of residues 1–225 reflects that observed in the crystal structure of WT<sub>299</sub>; the secondary structure of residues 226–337 reflects that predicted by Jpred 3.

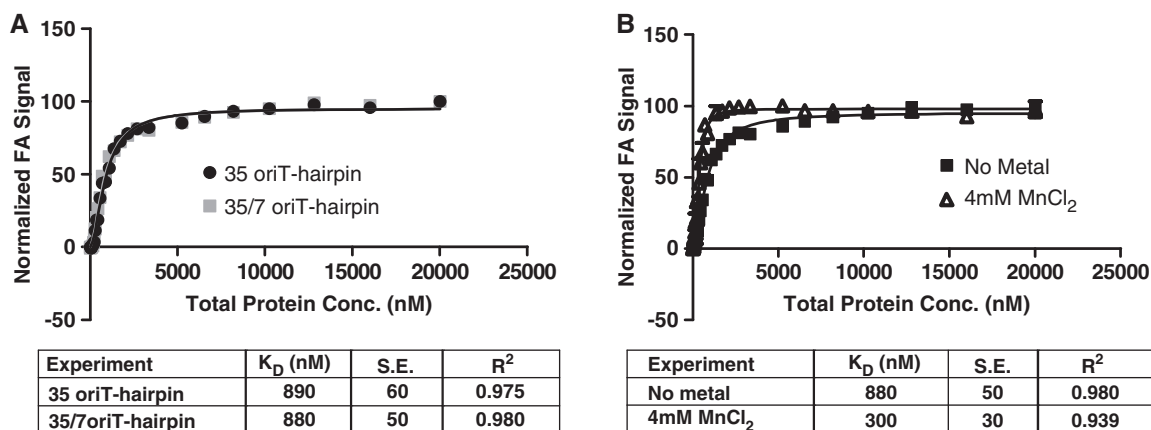
fitting the data with the Hill parameter, the calculated  $K_D$  did not improve, while the goodness-of-fit parameters (dependency and  $R^2$ ) showed slight or no improvement. These results did not justify the use of an additional parameter. Therefore, the data indicate that WT<sub>299</sub> binds DNA in a weak, sequence-independent manner, and may require an additional protein partner to initiate sequence-specific DNA binding.

#### Tyrosine-dependent DNA nicking by the pCU1 TraI relaxase

In the crystal structure of WT<sub>299</sub>, the tyrosine residues predicted to nick DNA (Y18,19,26,27) were displaced up

to 14 Å from the active site (Figure 1C). To determine if this significant displacement impacted the activity of the tyrosines when nicking DNA, we assessed the enzyme's DNA nicking activity by incubating WT<sub>299</sub> with 5' fluorescently labeled synthetic oligonucleotides and then monitored the amount of nicked product generated (Figure 4A). WT<sub>299</sub> efficiently nicked DNA containing the pCU1 oriT sequence and nic site. We then cloned, expressed and purified mutants of WT<sub>299</sub> in which one (Y18F and Y26F), two (Y18,26F) or four (Y18,19,26,27F) of the tyrosine residues were mutated to the catalytically inactive phenylalanine residue. The nicking activity of each mutant enzyme was then compared to that of





**Figure 3.** DNA binding by the pCU1 TraI relaxase. DNA binding curves were generated by monitoring the change in fluorescence anisotropy (FA) of a fluorescein-labeled DNA substrate as a function of protein concentration and then plotting the normalized FA as a function of total protein concentration. Each data point is the average of at least three replicates, with error bars representing the standard error of these replicates. The DNA binding affinity ( $K_D$ , in nM) was calculated by fitting the data with Equation 1 using non-linear regression. The error represents the standard error of the  $K_D$  when calculated by Equation 1. (A) Binding of the substrates 35oriT-hairpin and 35/7oriT-hairpin by WT\_299 and the pCU1 relaxase mutant Y18,19,26,27F\_299, respectively. The substrate 35/7oriT-hairpin extends 7 nt beyond the oriT nic site, while 35oriT-hairpin is truncated at the nic site. (B) Binding of the substrate 35/7oriT-hairpin by Y18,19,26,27F\_299 in the absence or presence of 4 mM  $MnCl_2$ .

WT\_299. Single and double tyrosine mutants suffered an ~50% decrease in activity, while the mutation of all four tyrosines effectively eliminated all nicking activity (Figure 4A). Thus, the catalytic tyrosines of the pCU1 relaxase were shown to be structurally mobile and functionally redundant.

#### Metal-dependent DNA nicking by the pCU1 TraI relaxase

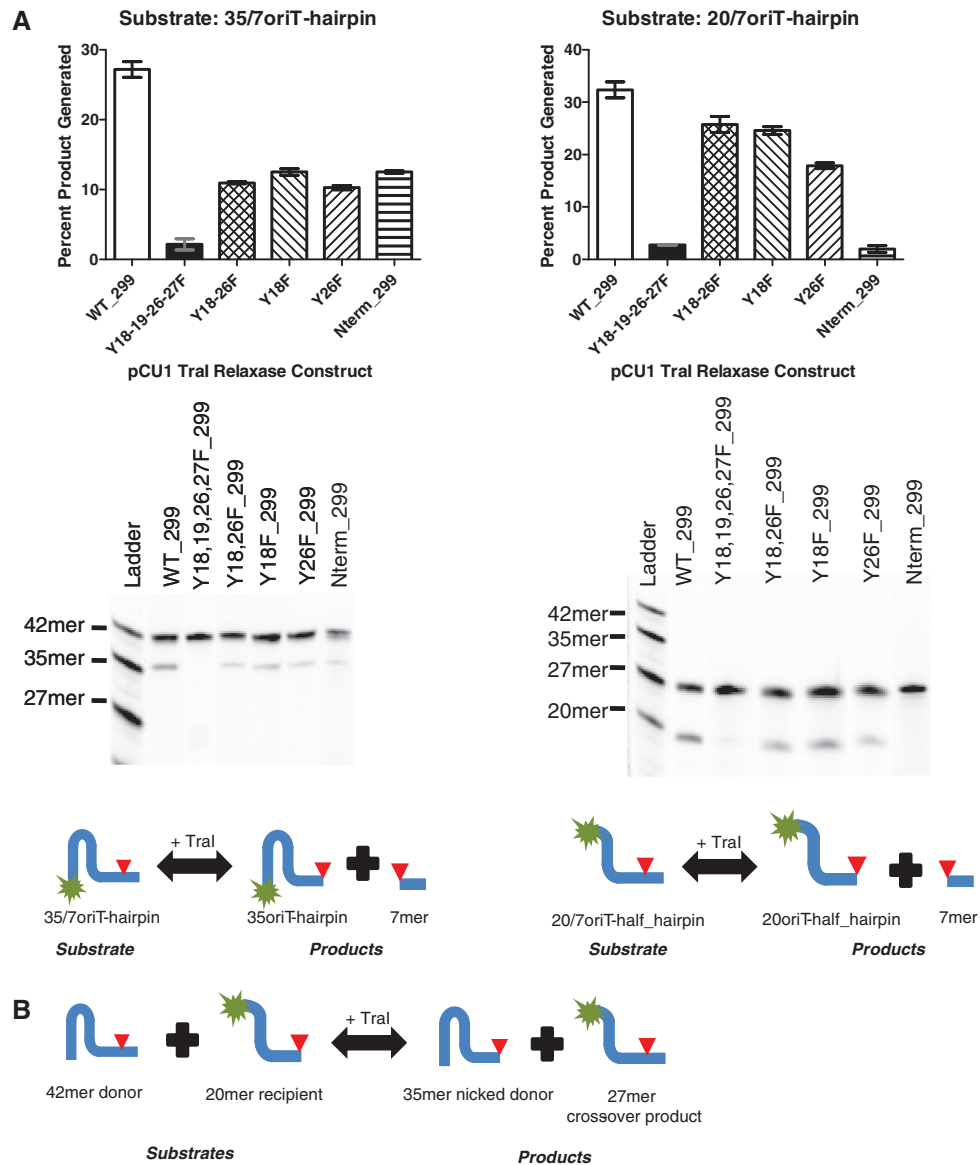
Both the F TraI and R388 TrwC relaxase enzymes exhibit metal-dependent DNA nicking, though the optimal metal for each is under debate (3,24,26,34,46). To determine the metals most likely to be used by WT\_299 during DNA nicking, the nicking activity of the enzyme was measured in the presence of EDTA and several metals ( $Mg^{2+}$ ,  $Mn^{2+}$ ,  $Ca^{2+}$ ,  $Ni^{2+}$ ,  $Cu^{2+}$  and  $Zn^{2+}$ ) (Figure 5). In the presence of EDTA, the DNA nicking activity of WT\_299 was inhibited, but this activity could be rescued by then supplementing the reaction with an increasing concentration of a variety of metals.  $Mg^{2+}$  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.* (34), the relaxase maintained residual DNA nicking activity in the presence of EDTA. Despite controlling the pH of  $Cu(SO_4)$ ,  $ZnCl_2$  and  $Zn(acetate)_2$  solutions, 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 insolubility within the pH range of the assay. Using inductively coupled plasma mass spectrometry (ICP-MS), we then determined the concentrations of Mg24, Mn55, Ca44, Fe57, Ni60 and Zn66 in samples of WT\_299 (Table 4). When ranked according to fold increase over background levels, Zn66 was

by far the most prevalent metal in the WT\_299 sample, followed by Ni60 and Mn55. Taken together, these data indicate that while the relaxase can function in the presence of many metals, its nicking activity is best supported by  $Mg^{2+}$ ,  $Mn^{2+}$ ,  $Ca^{2+}$  or  $Ni^{2+}$ , despite the overabundance of Zn66 bound by the protein sample.

#### Sequence-dependent DNA nicking by the pCU1 TraI relaxase

DNA binding assays demonstrated the sequence-independent DNA binding of WT\_299. To determine if WT\_299 also nicked and religated DNA in a sequence-independent manner, the DNA nicking and religation activities of WT\_299 were tested using a series of DNA nicking and cross-over assays. First, to examine the effect of DNA secondary structure on pCU1 TraI relaxase DNA nicking activity, 5' fluorescently labeled substrates containing the pCU1 oriT sequence, nic site and either a full hairpin-forming inverted repeat (35/7oriT-hairpin), a mutated inverted repeat unable to form a hairpin (35/7-no\_hairpin), or a truncated hairpin (20/7oriT-half\_hairpin) were incubated with WT\_299 (Table 2). Each substrate was cleaved, but the percent product generated increased upon truncation of the inverted repeat. A similar trend was observed when studying the nicking activity of both the F TraI and R388 TrwC relaxase enzymes (25,37).

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. WT\_299 only cleaved the oriT-containing substrate (Table 2). In the same manner, during DNA cross-over assays, WT\_299 could only successfully religate oriT-containing substrates. In fact, both donor and recipient substrates had to contain the pCU1 oriT sequence and nic site. If either donor or recipient did not encode the



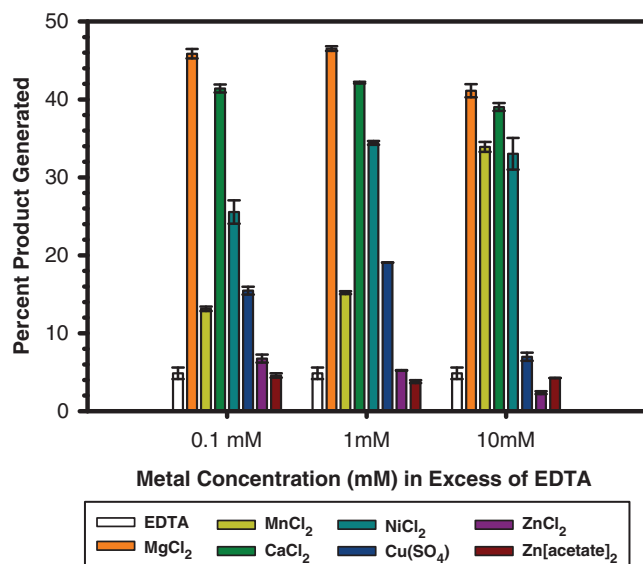
**Figure 4.** DNA nicking and religation by the pCU1 TraI relaxase domain. (A) To determine the DNA nicking activity of the wild-type pCU1 TraI relaxase, as well as several relaxase mutants, 5  $\mu$ M enzyme was incubated with 1  $\mu$ M substrate (35/7oriT- or 20/7oriT-hairpin) for 1 h at 37°C, and products were then separated on denaturing polyacrylamide gels. Nicking and religation activity is represented as percent product generated. Each experiment was performed in triplicate; the error is the standard error of these replicates. For each experiment, a representative gel is provided. Below each gel, a reaction scheme is provided, where a fluorescein label is indicated by a green star and the substrate nic site by a red triangle. Additional experimental details can be found in the Materials and Methods section and the Supplementary Data. (B) A reaction scheme illustrating a DNA cross-over experiment, where a fluorescein label is indicated by a green star and the substrate nic site by a red triangle.

oriT, or were truncated prior to the nic site, no cross-over products were observed (Supplementary Figure S5).

To determine the specific bases within the pCU1 oriT sequence required for relaxase-mediated DNA nicking, a series of 42 nt-long unlabeled DNA substrates were designed in which mutations were made within one or multiple regions of the wild-type pCU1 oriT sequence (Table 2, Figure 6). These regions were designated 'hairpin', 'TAG', 'pentanucleotide', or 'post nic', and corresponded to sections of the oriT previously shown by work with the F and R388 plasmids to influence DNA nicking, binding, or plasmid transfer (22,25,31,37). With these data in mind, the TAG and pentanucleotide regions

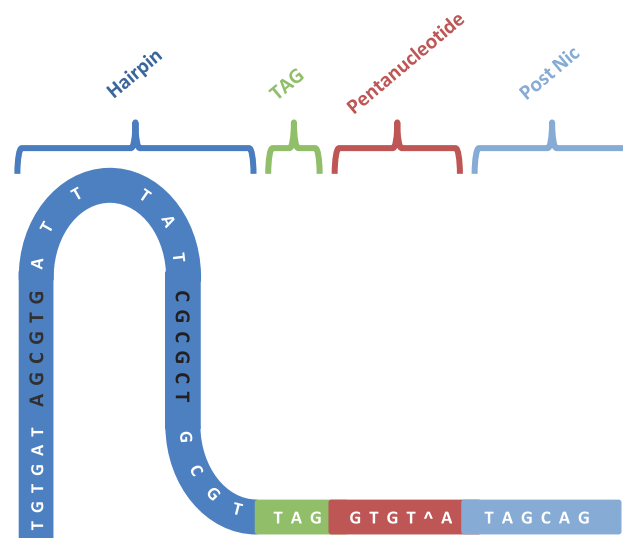
were predicted to strongly influence DNA nicking, while the hairpin and post nic regions were expected to have only a limited effect. WT<sub>299</sub> was then incubated with each of these unlabeled mutant DNA substrates (the donor substrate) and a fluorescently labeled 20 nt-long recipient substrate (20-oriT\_half-hairpin) that encoded the pCU1 oriT and terminated at the nic site. An experiment was thus performed to measure the ability of the relaxase to nick the donor substrate and generate a 27 nt-long cross-over product (Figure 4B). 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, post nic 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 2).



**Figure 5.** DNA nicking by the pCU1 TraI relaxase in the presence of metals. 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  $\mu$ M enzyme was incubated with 1  $\mu$ M substrate (35/7oriT-hairpin) for 1 h at 37°C, and products were then separated on denaturing polyacrylamide gels. Nicking and religation activity is represented as percent product generated. Each experiment was performed in triplicate; the error is the standard error of these replicates. Additional details can be found in the Materials and Methods section and the Supplementary Data.

A number of the substrates were apparently nicked at erroneous nic sites by WT\_299 prior to religation onto 20oriT-half\_hairpin, thus creating products either longer or shorter than the expected 27mer (Table 2). In particular, these misnicked 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 misnicked 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,



**Figure 6.** Nucleic acid sequence and sub-divisions of the pCU1 oriT. The nucleic acid sequence of the plasmid pCU1 oriT 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 indicated.

**Table 4.** Summary of metal data

Metal	Mg	Mn	Ca	Fe	Ni	Cu	Zn
Conc. in Bacterial Cells ( $\mu$ M) <sup>a</sup>	>10,000	10	100	100	<b>Low, n/a</b>	10	100
Fold increase in conc. over background <sup>b</sup> (Protein: WT_299)	1.16	1.50	1.10	1.15	7.75	Not tested	45.12
Percent product generated <sup>b,c</sup> (Substrate: 35/7oriT-hairpin)	46.5	33.9	42.1	N/A	34.4	19.1	6.7
Satisfies electron density? <sup>b</sup>	<b>No</b>	Yes	<b>No</b>	Yes	Yes	Yes	Yes
Preferred coordination <sup>d</sup>	6	5,6	6	5,6 <sup>e</sup>	6	<b>3,4</b>	<b>4</b>

<sup>a</sup>Source: Finney *et al.* (59).

<sup>b</sup>Source: This article.

<sup>c</sup>DNA nicking data is the average of three replicates and the activity reported is the percent product generated at the optimal concentration for each metal; Cu<sup>2+</sup> becomes inhibitory as its concentration increases.

<sup>d</sup>Source: Harding *et al.* (61).

<sup>e</sup>This analysis data includes both ferric and ferrous groups.

absence, or location of the hairpin forming inverted repeat.

The pentanucleotide region of the pCU1 oriT 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 (27,37). To illustrate the importance of the interaction between the pentanucleotide region and the pCU1 TraI N-terminus, we compared the structure and function of WT\_299 to that of a mutant pCU1 TraI relaxase, Nterm\_299, in which three residues (Ser-Asn-Ala) were appended to the protein N-terminus. We measured the DNA nicking activity of Nterm\_299 and determined its structure to 1.9 Å by X-ray crystallography. The nicking activity of Nterm\_299 was in fact inhibited relative to WT\_299 (Figure 4A). The structures of WT\_299, Nterm\_299 and the DNA-bound relaxase domain of the R388 TrwC relaxase [PDB code 2cdm (34)] were overlaid (Supplementary Figure S6). In the TrwC relaxase structure, its DNA substrate, a 27mer oligonucleotide containing the R388 oriT, is bound by the relaxase as a partial hairpin. The overall folds of the three proteins are similar (maximum rmsd over equivalent C $\alpha$  positions = 2.05 Å) (57,58), and the histidine triad-dominated active sites are practically identical (averaged rmsd of these residues between each of the three structures = 0.076 Å). However, the three non-native N-terminal residues of Nterm\_299 clearly clash with the bound DNA substrate of TrwC. Therefore, Nterm\_299 should be incapable of nicking oriT DNA without either physically moving its N-terminus out of the way, or by binding the DNA in a novel orientation. Both options destroy all sequence-specific protein-DNA interactions formed between the enzyme N-terminus and the oriT DNA. Thus, this structural clash clearly explains the functional inhibition observed during DNA nicking assays.

## DISCUSSION

### The role of tyrosine mobility during pCU1 relaxase-mediated DNA nicking

CPT with concomitant replication is proposed to proceed through two sequential DNA nicking steps that are accomplished by catalytic tyrosine residues contributed by relaxase enzymes (3,32,35). These reactions can be completed using either single tyrosines on two relaxases acting sequentially, or by one multi-tyrosine relaxase, through a ‘flip-flop’ mechanism of DNA nicking and religation described previously (3,14,32,35). While the ‘flip-flop’ mechanism requires significant rearrangement of tyrosine residues within the active site of the multi-tyrosine relaxase, it abrogates the need for multiple enzymes to accomplish the transfer of one conjugative plasmid.

Our structural and function data have emphasized the flexible nature of the tyrosines within multi-tyrosine systems. In particular, the crystal structure of WT\_299 demonstrated the mobility of loop A and its accompanying four catalytic tyrosines. As compared to homologous structures, loop A of WT\_299 monomer A is rotated away

from the active site and its first two catalytic tyrosines are displaced from the metal-bound histidine triad (Figure 1C). In monomer B, loop A was too mobile to be modeled into the structure. Our DNA nicking data also highlighted the flexibility of the four tyrosines. Mutation of up to two tyrosine residues only eliminated ~50% of nicking activity, indicating that tyrosines other than the predicted primary DNA nicking residue, Tyr18, were also able to enter the active site in a productive orientation (Figure 4A). These data support the conclusion that the active site residues of the pCU1 relaxase, and perhaps additional multi-tyrosine systems, are sufficiently mobile to allow for the rearrangement necessary to accommodate the ‘flip-flop’ mechanism of plasmid transfer.

### Metal ions in the pCU1 relaxase active site

Productive CPT requires that the relaxase nick the T-strand DNA within the enzyme’s metal-chelating active site (3,32,35). However, the identity of this essential metal ion is debated (3,26,34,46). Therefore, we used enzyme activity assays to establish that Mg $^{2+}$ , Mn $^{2+}$ , Ca $^{2+}$  and Ni $^{2+}$  all supported pCU1 relaxase-mediated DNA nicking (Figure 5). In contrast, both Zn $^{2+}$  and Cu $^{2+}$  supported only limited DNA nicking by WT\_299 at low metal concentrations and were actually inhibitory as the concentration of metal increased. We then measured concentration of each metal in a purified sample of WT\_299 using ICP mass spectrometry (Table 4). Based on these experimental data, as well as the low concentration of Ni $^{2+}$  within the bacterial cytoplasm (59,60) 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 relaxase most likely employed Mn $^{2+}$  during DNA nicking and religation reactions.

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 (44,46). Larkin *et al.* demonstrated by isothermal titration calorimetry (ITC) that the affinity of F TraI relaxase was ~25 times stronger for Mn $^{2+}$  than for Ca $^{2+}$  and Mg $^{2+}$  (24), 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 $^{2+}$  concentrations (34).

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 a trigonal bipyramidal or octahedral arrangement, Zn $^{2+}$  and Cu $^{2+}$  prefer three or four coordinating ligands (61). 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.* (3) and Datta *et al.* (23) both observed an octahedral arrangement



of ligands surrounding the bound  $Mg^{2+}$  in the active site of the F TraI relaxase, and Lujan and colleagues noted that this octahedral arrangement would allow for the 'flip-flop' mechanism proposed for the relaxase (3). In summary, while the pCU1 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.

#### Low affinity and sequence-independent DNA binding by the pCU1 relaxase

In contrast to the high affinity, sequence-specific DNA binding by the F TraI and R388 TrwC relaxase enzymes (22,25,33), the pCU1 TraI relaxase bound DNA in a non-specific manner and with low affinity (Table 1, Figure 3). We propose that the pCU1 enzyme 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 deletion (51). Homologs 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 (13,21,40,62). In particular, studies of RP4 have shown that its TraK homolog, TraJ, enhances sequence-specific binding and nicking by the RP4 TraI enzyme (13,40). In addition, the F plasmid-encoded TraK homolog, TraY, has been shown necessary for relaxase binding and nicking *in vivo* (13). Therefore, pCU1 sequence-specific binding may be accomplished by a two-component system. pCU1 TraK could bind selectively to the pCU1 oriT, to then promote high affinity DNA binding by TraI.

#### Sequence-specific DNA cleavage by the pCU1 relaxase

In order to cleave the T-strand DNA at the scissile phosphate of the plasmid nic site, the relaxase must position the oriT DNA within its active site such that the nic site is located directly above the chelated metal ion (24). However, due to the pCU1 relaxase's low affinity and non-specific DNA binding, it was unclear whether or not this enzyme would be able to nick DNA in the anticipated sequence-specific manner. In crystal structures of the DNA-bound F TraI and R388 TrwC relaxase enzymes, the DNA substrate is observed to form a U-turn around the enzyme N-terminus immediately prior to entering the active site. As a result, the DNA bases immediately upstream of the substrate's nic site form several contacts with the relaxase N-terminus that have been shown to regulate relaxase-mediated DNA nicking activity (24,27,33,34). Therefore, we determined the DNA nicking activity of the pCU1 relaxase against a library of DNA substrates designed to test the sequence selectivity of the enzyme for the pCU1 oriT, with a particular emphasis on the regions of the pCU1 oriT corresponding to the bases of the F and R388 oriT known to form this U-turn structure (Table 2, Figure 6).

The results of the DNA nicking assays indicated that the bases of the pCU1 oriT, in particular those within the pentanucleotide region, cooperate with the N-terminus of

WT\_299 to position the nic site within the enzyme active site to promote sequence-specific DNA nicking (Table 2). Depending on the location of the mutation(s) within the pentanucleotide region, WT\_299 failed to cleave the substrate's nic site entirely, cleaved the site at a reduced level, or cleaved the substrate at non-canonical nic sites. In particular, by shifting the location of the pentanucleotide region relative to the nic site, the location of DNA nicking could be shifted correspondingly. DNA nicking was also inhibited by mutation of the enzyme N-terminus, as illustrated by the failure of mutant enzyme Nterm\_299 to nick oriT-containing DNA substrates (Figure 4A). Of particular interest, the bases within the pentanucleotide region had previously been shown to be necessary for pCU1 plasmid transfer during cell-based plasmid transfer assays (52). This correlation emphasizes the pivotal role of the relaxase enzyme within the relaxosome during sequence-specific CPT.

#### ACCESSION NUMBERS

PDB accession code 3L57, 3L6T.

#### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

#### ACKNOWLEDGEMENTS

We thank Suzanne Paterson for the generous gift of the TraI gene. We thank Jeff Keenan for assistance with the initial cloning of the TraI gene. We thank Sarah Kennedy for collecting data sets leading to PDB accession code 3L57. We thank Eric Ortlund, Bret Wallace, Jillian Orans and Laurie Betts for guidance when analyzing the crystallographic data leading to PDB accession codes 3L57 and 3L6T. We thank Joseph V. Lomino for assistance when analyzing CD data. We thank the Matson laboratory for instruction when performing non-denaturing DNA gels. We thank Dan McNamara, Andrew Hemmert and members of the Redinbo lab for assistance with manuscript preparation. Use of the Advanced Photon Source (APS) was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract No W-31-109-Eng-38. We thank SER-CAT for access to beam-line 22-ID at the APS. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. We thank SSRL for access to beam-line BL9-2.

**FUNDING**

National Institutes of Health [grant numbers AI078924 to M.R.R., ES016488 to R.P.N., GM008719 to UNC-CH]. Funding for open access charge: National Institutes of Health (grant number AI078924).

*Conflict of interest statement.* None declared.

**REFERENCES**

- Siegel, J.D., Rhinehart, E., Jackson, M. and Chiarello, L. (2006) Management of Multidrug-Resistant Organisms in Healthcare Settings. *The Centers for Disease Control and Prevention. Healthcare Infection Control Practices Advisory Committee.*
- Clatworthy, A.E., Pierson, E. and Hung, D.T. (2007) Targeting virulence: a new paradigm for antimicrobial therapy. *Nat. Chem. Biol.*, **3**, 541–548.
- Lujan, S.A., Guogas, L.M., Ragonese, H., Matson, S.W. and Redinbo, M.R. (2007) Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proc. Natl Acad. Sci. USA*, **104**, 12282–12287.
- Fernandez-Lopez, R., Machon, C., Longshaw, C.M., Martin, S., Molin, S., Zechner, E.L., Espinosa, M., Lanka, E. and de la Cruz, F. (2005) Unsaturated fatty acids are inhibitors of bacterial conjugation. *Microbiology*, **151**, 3517–3626.
- Garcillan-Barcia, M.P., Jurado, P., Gonzalez-Perez, B., Moncalian, G., Fernandez, L.A. and de la Cruz, F. (2007) Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells with intrabodies. *Mol. Microbiol.*, **63**, 404–416.
- Williams, J.J. and Hergenrother, P.J. (2008) Exposing plasmids as the Achilles' heel of drug-resistant bacteria. *Curr. Opin. Chem. Biol.*, **12**, 389–399.
- Qiu, X.-Q., Wang, H., Lu, X.-F., Zhang, J., Li, S.-F., Cheng, G., Wan, L., Yang, L., Zuo, J.-Y., Zhou, Y.-Q. *et al.* (2003) An engineered multidomain bactericidal peptide as a model for targeted antibiotics against specific bacteria. *Nat. Biotechnol.*, **21**, 1480–1485.
- Potts, R.G., Lujan, S.A. and Redinbo, M.R. (2008) Winning the asymmetric war: new strategies for combating antibacterial resistance. *Future Microbiol.*, **3**, 119–123.
- Willets, N. and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.*, **48**, 24–41.
- Lanka, E. and Wilkins, B.M. (1995) DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.*, **64**, 141–169.
- de la Cruz, F. and Davies, J. (2000) Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol.*, **8**, 128–133.
- Llosa, M., Gomis-Ruth, F.X., Coll, M. and de la Cruz, F. (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.*, **45**, 1–8.
- de la Cruz, F., Frost, L.S., Meyer, R.J. and Zechner, E.L. (2009) Conjugative DNA Metabolism in Gram-negative bacteria. *FEMS Microbiol. Rev.*, **34**, 18–30.
- Byrd, D.R.N. and Matson, S.W. (1997) Nicking by transesterification: the reaction catalysed by a relaxase. *Mol. Microbiol.*, **25**, 1011–1022.
- Matson, S.W., Sampson, J.K. and Byrd, D.R.N. (2001) F plasmid conjugative DNA transfer: the TraI helicase activity is essential for DNA strand transfer. *J. Biol. Chem.*, **276**, 2372–2379.
- Tatum, E.L. and Lederberg, J. (1947) Gene Recombination in the bacterium *Escherichia coli*. *J. Bacteriol.*, **53**, 673–684.
- Matson, S.W., Nelson, W.C. and Morton, B.S. (1993) Characterization of the reaction product of the oriT nicking reaction catalyzed by *Escherichia coli* DNA helicase I. *J. Bacteriol.*, **175**, 2599–2606.
- Sherman, J.A. and Matson, S.W. (1994) *Escherichia coli* DNA helicase I catalyzes a sequence-specific cleavage/ligation reaction at the F plasmid origin of transfer. *J. Biol. Chem.*, **269**, 26220–26226.
- Howard, M.T., Nelson, W.C. and Matson, S.W. (1995) Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J. Biol. Chem.*, **270**, 28381–28386.
- Matson, S.W. and Ragonese, H. (2005) The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. *J. Bacteriol.*, **187**, 697–706.
- Nelson, W.C., Howard, M.T., Sherman, J.A. and Matson, S.W. (1995) The *traY* gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid oriT. *J. Biol. Chem.*, **270**, 28374–28380.
- Stern, J.C. and Schildbach, J.F. (2001) DNA recognition by F factor TraI36: highly sequence-specific binding of single-stranded DNA. *Biochemistry*, **40**, 11586–11595.
- Datta, S., Larkin, C. and Schildbach, J.F. (2003) Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure*, **11**, 1369–1379.
- Larkin, C., Datta, S., Harley, M.J., Anderson, B.J., Ebie, A., Hargreaves, V. and Schildbach, J.F. (2005) Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure*, **13**, 1533–1544.
- Williams, S.L. and Schildbach, J.F. (2006) Examination of an inverted repeat within the F factor origin of transfer: context dependence of F TraI relaxase DNA specificity. *Nucleic Acids Res.*, **34**, 426–435.
- Larkin, C., Haft, R.J.F., Harley, M.J., Traxler, B. and Schildbach, J.F. (2007) Roles of active site residues and the HUH motif of the F plasmid TraI relaxase. *J. Biol. Chem.*, **282**, 33707–33713.
- Hekman, K., Guja, K., Larkin, C. and Schildbach, J.F. (2008) 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 Res.*, **36**, 4565–4572.
- Llosa, M., Bolland, S. and de la Cruz, F. (1991) 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. *Mol. General Genet.*, **226**, 473–483.
- Grandoso, G., Llosa, M., Zabala, J.C. and de la Cruz, F. (1994) Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. *Eur. J. Biochem.*, **226**, 403–412.
- Llosa, M., Grandoso, G. and de la Cruz, F. (1995) Nicking activity of TrwC directed against the origin of transfer of the IncW plasmid R388. *J. Mol. Biol.*, **246**, 54–62.
- Llosa, M., Grandoso, G., Hernando, M.A. and de la Cruz, F. (1996) Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J. Mol. Biol.*, **264**, 56–67.
- Grandoso, G., Avila, P., Cayon, A., Hernando, M.A., Llosa, M. and de la Cruz, F. (2000) Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J. Mol. Biol.*, **295**, 1163–1172.
- Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F.X., de la Cruz, F. and Coll, M. (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nature Struct. Biol.*, **10**, 1002–1010.
- Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A.G., Perez-Luque, R., Coll, M. and de la Cruz, F. (2006) Unveiling the molecular mechanism of a conjugative relaxase: the structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J. Mol. Biol.*, **358**, 857–869.
- Gonzalez-Perez, B., Lucas, M., Cooke, L.A., Vyle, J.S., de la Cruz, F. and Moncalian, G. (2007) Analysis of DNA processing reactions in bacterial conjugation by using suicide oligonucleotides. *EMBO J.*, **26**, 3847–3857.
- Gonzalez-Perez, B., Carballeira, J.D., Moncalian, G. and de la Cruz, F. (2008) Changing the recognition site of a conjugative relaxase by rational design. *Biotechnol. J.*, **4**, 554–557.
- Lucas, M., Gonzalez-Perez, B., Cabezas, M., Moncalian, G., Rivas, G. and de la Cruz, F. (2010) Relaxase DNA binding and cleavage are two distinguishable steps in conjugative DNA processing that involve different sequence elements of the *nic* site. *J. Biol. Chem.*, doi: 10.1074/jbc.M109.057539.

38. Pansegrau, W., Schroder, W. and Lanka, E. (1993) Relaxase (TraI) of IncPa plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl Acad. Sci. USA*, **90**, 2925–2929.
39. Balzer, D., Pansegrau, W. and Lanka, E. (1994) Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. *J. Bacteriol.*, **176**, 4285–4295.
40. Pansegrau, W. and Lanka, E. (1996) Mechanisms of initiation and termination reactions in conjugative DNA processing. *J. Biol. Chem.*, **271**, 13068–13076.
41. Bhattacharjee, M. and Meyer, R.J. (1993) 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 Res.*, **21**, 4563–4568.
42. Becker, E.C. and Meyer, R.J. (2000) Recognition of oriT for DNA processing at termination of a round of conjugal transfer. *J. Mol. Biol.*, **300**, 1067–1077.
43. Becker, E.C. and Meyer, R.J. (2003) Relaxed specificity of the R1162 nickase: a model for evolution of a system for conjugative mobilization of plasmids. *J. Bacteriol.*, **185**, 3538–3546.
44. Monzingo, A.F., Ozburn, A., Xia, S., Meyer, R.J. and Robertus, J.D. (2007) The structure of the minimal relaxase domain of MobA at 2.1 Å resolution. *J. Mol. Biol.*, **366**, 165–178.
45. Meyer, R.J. (2008) The R1162 Mob proteins can promote conjugative transfer from cryptic origins in the bacterial chromosome. *J. Bacteriol.*, **191**, 1574–1580.
46. Xia, S. and Robertus, J.D. (2009) Effect of divalent ions on the minimal relaxase domain of MobA. *Arch. Biochem. Biophys.*, **488**, 42–47.
47. Harley, M.J. and Schildbach, J.F. (2003) Swapping single-stranded DNA sequence specificities of relaxases from conjugative plasmids F and R100. *Proc. Natl Acad. Sci. USA*, **100**, 11243–11248.
48. Fekete, R.A. and Frost, L.S. (2000) Mobilization of chimeric oriT plasmids by F and R100-I: role of relaxosome formation in defining plasmid specificity. *J. Bacteriol.*, **182**, 4022–4027.
49. Khatoon, H., Iyer, R.V. and Iyer, V.N. (1972) A new filamentous bacteriophage with sex-factor specificity. *Virology*, **48**, 145–155.
50. Konarska-Kozłowska, M. and Iyer, V.N. (1981) Physical and genetic organization of the IncN-group plasmid pCU1. *Gene*, **14**, 195–204.
51. Paterson, E.S., More, M.I., Pillay, G., Cellini, C., Woodgate, R., Walker, G.C., Iyer, V.N. and Winans, S.C. (1999) Genetic analysis of the mobilization and leading regions of the IncN plasmids pKM101 and pCU1. *J. Bacteriol.*, **181**, 2572–2583.
52. Paterson, E.S. and Iyer, V.N. (1997) Localization of the nic site of IncN conjugative plasmid pCU1 through formation of a hybrid oriT. *J. Bacteriol.*, **179**, 5768–5776.
53. Paterson, E.S. and Iyer, V.N. (1992) The oriT region of the conjugative transfer system of plasmid pCU1 and specificity between it and the mob region of other N tra plasmids. *J. Bacteriol.*, **174**, 499–507.
54. Donnelly, M.I., Zhou, M., Millard, C.S., Clancy, S., Stols, L., Eschenfeldt, W.H., Collart, F.R. and Joachimiak, A. (2006) An expression vector tailored for large-scale, high throughput purification of recombinant proteins. *Protein Expression Purif.*, **47**, 446–454.
55. Lo Conte, L., Chothia, C. and Janin, J. (1999) The atomic structure of protein-protein recognition sites. *J. Mol. Biol.*, **285**, 2177–2198.
56. Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.*, **372**, 774–797.
57. Barthel, D., Hirst, J., Blazewicz, J., Burke, E. and Krasnogor, N. (2007) ProCKSI: a decision support system for protein (structure) comparison, knowledge, similarity and information. *BMC Bioinformatics*, **8**, 416.
58. Zhang, Y. and Skolnick, J. (2005) TM-align: a protein structure alignment algorithm based on TM-score. *Nucleic Acids Res.*, **33**, 2302–2309.
59. Finney, L.A. and O'Halloran, T.V. (2003) Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science*, **300**, 931–936.
60. Outten, C. and O'Halloran, T. (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science*, **292**, 2488–2492.
61. Harding, M.M. (2001) Geometry of metal-ligand interactions in proteins. *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **57**, 401–411.
62. Moncalian, G. and de la Cruz, F. (2004) DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim. Biophys. Acta*, **1701**, 15–23.