

An intrastrand three-DNA-base interaction is a key specificity determinant of F transfer initiation and of F Tral relaxase DNA recognition and cleavage

Katherine Hekman, Kip Guja, Chris Larkin and Joel F. Schildbach*

Department of Biology, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA

Received January 14, 2008; Revised June 9, 2008; Accepted June 18, 2008

ABSTRACT

Bacterial conjugation, transfer of a single conjugative plasmid strand between bacteria, diversifies prokaryotic genomes and disseminates antibiotic resistance genes. As a prerequisite for transfer, plasmid-encoded relaxases bind to and cleave the transferred plasmid strand with sequence specificity. The crystal structure of the F Tral relaxase domain with bound single-stranded DNA suggests binding specificity is partly determined by an intrastrand three-way base-pairing interaction. We showed previously that single substitutions for the three interacting bases could significantly reduce binding. Here we examine the effect of single and double base substitutions at these positions on plasmid mobilization. Many substitutions reduce transfer, although the detrimental effects of some substitutions can be partially overcome by substitutions at a second site. We measured the affinity of the F Tral relaxase domain for several DNA sequence variants. While reduced transfer generally correlates with reduced binding affinity, some *oriT* variants transfer with an efficiency different than expected from their binding affinities, indicating ssDNA binding and cleavage do not correlate absolutely. Oligonucleotide cleavage assay results suggest the essential function of the three-base interaction may be to position the scissile phosphate for cleavage, rather than to directly contribute to binding affinity.

INTRODUCTION

Bacterial conjugation is a DNA transfer process in which a single strand of a conjugative plasmid is transferred from donor to recipient. Conjugative plasmid transfer

assists in diversifying bacterial genomes, contributes to an accelerated evolution rate in pathogenic *Escherichia coli* lineages (1), and may directly participate in pathogenic processes (2). Numerous conjugative plasmids have been identified and studied, and while the plasmids display enormous diversity, the transfer process is analogous for most plasmids (3).

Conjugative transfer is highly regulated in numerous ways. For example, in one type of temporal regulation, a stable intercellular contact, or mating pair, is usually formed between donor and recipient prior to transfer. On the molecular level, the site at which the plasmid is cleaved, and from which transfer initiates, is also highly regulated. This regulation is mediated at least partially through the sequence specificity of the plasmid-encoded relaxase protein (4–8). Relaxases, or nickases, are single-strand DNA (ssDNA) nucleases that cleave one strand of the conjugative plasmid at the *nic* site within the plasmid *oriT* (origin of transfer). As part of the reaction, the protein forms a long-lived covalent linkage with the DNA (9,10). The relaxase may serve as a pilot protein, guiding the attached DNA into the recipient cell and joining the plasmid ends together to circularize the plasmid to terminate transfer (11).

For conjugative plasmid F factor, the bifunctional TraI protein both serves as the relaxase and possesses an essential helicase activity. Using *in vitro* binding assays, we previously demonstrated that TraI36, the 36-kDa relaxase domain of F TraI, recognizes its binding site within the *ForiT* with subnanomolar K_D and remarkable sequence specificity (8). The crystal structure of F TraI36 in complex with its ssDNA *oriT* binding site revealed that the protein binds the ssDNA in a cleft in the protein, and additional contacts are provided by a flap of the protein that folds over the bound ssDNA (12). TraI36 interacts with ssDNA both through nonspecific electrostatic interactions with backbone phosphates and more specific interactions with the bases, including a knobs-into-holes interaction between two DNA bases and pockets

*To whom correspondence should be addressed. Tel: (410) 516 0176; Fax: (410) 516 5213; Email: joel@jhu.edu

Present address:

Katherine Hekman, Pritzker School of Medicine, The University of Chicago, 924 East 57th Street, Chicago, IL 60637, USA
Chris Larkin, Laboratory of Molecular Biology, NIDDK, NIH, Bethesda, MD 20892, USA

within the binding cleft. We also observed an intrastrand three-way base-pairing interaction in the bound ssDNA (12). As shown in Figure 1, the base immediately 5' to *nic*, T141' [numbered according to Frost and colleagues (13), with the prime indicating the position on the 'bottom' strand] forms hydrogen bonds with both G144' and G142' (the fourth and second bases 5' to *nic*, respectively). An intrastrand base-pairing interaction also occurs in the crystal structure of the complex of ssDNA with the TrwC relaxase domain (14). For the F TraI relaxase, our examination of a limited panel of variants demonstrated that single base substitutions for the three bases involved in the interaction can dramatically reduce *in vitro* binding affinity and *in vivo* transfer efficiency (8). We wondered, however, whether the wild-type trio of bases was the only combination conducive to transfer. To examine this question, we generated a set of plasmids containing variant *ForiT* regions having all possible individual substitutions at positions 141', 143' and 144', and having pairwise substitutions at position 141' and at position 142' or 144'. We then measured the efficiency with which these plasmids were transferred. In addition, we examined the *in vitro* binding of a subset of these variant sequences to the F TraI relaxase domain to determine how well transfer efficiency correlates with ssDNA binding affinity. We found that nearly all substitutions for T141', G144' and G142' reduced transfer efficiency, although substitutions at a second position could at least partially compensate for some of the single substitutions. We also found that while reduced affinity generally correlated with reduced transfer, there were exceptions, suggesting that requirements for binding and cleavage are to some extent distinct.

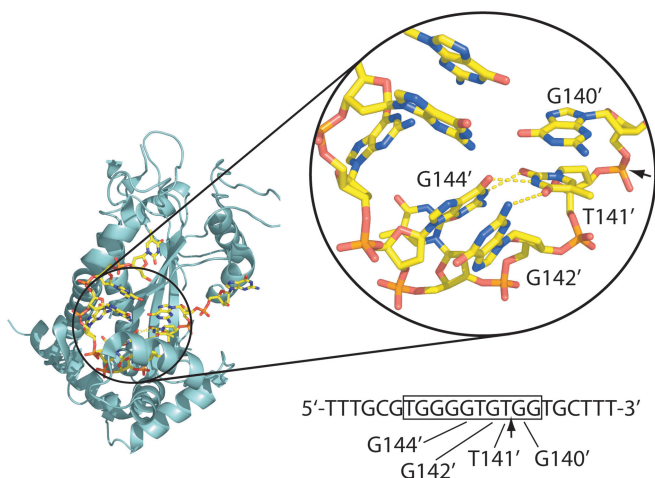


Figure 1. The intramolecular base pairing in the F TraI36:ssDNA crystal structure [PDB ID 2A0I; (12)]. TraI36 is shown as a turquoise ribbon while DNA bases are shown as sticks, with oxygen in red, phosphorus in orange, carbon in yellow and nitrogen in blue. Labeled DNA bases are those for which substitutions were made. Arrows indicate the position of *nic*, the scissile phosphate, both in the structure and the sequence underneath. The sequence shown is that of the 22-base oligonucleotide used in binding studies. Bases visible in the structure are boxed. The figure was generated using the program PyMOL (<http://pymol.sourceforge.net>).

MATERIALS AND METHODS

Strains, plasmids and proteins

E. coli strain ER2738 (F' proA⁺B⁺ lacI^q Δ (lacZ)M15 zzz::Tn10(Tet^R)/fhuA2 glnV Δ (lac-proAB) thi-1 Δ (hdsS-mcrB)5 and TB1 (F^- ara Δ (lac-proAB) [ϕ 80dlac Δ (lacZ)M15] rpsL(Str^R) thi hsdR)), and plasmid pACYC177, were obtained from New England BioLabs, Ipswich, MA, USA. Plasmid pACYC177-*ForiT* was constructed as described (15) and includes bp 1–530 [numbering according to Frost and colleagues (13)]. The region contains *nic*, *sbyA*, *sbyC*, *sbmC*, *sbmB*, *sbmA* and IHF-binding sites A and B.

Variant *oriT* segments were engineered in pACYC177-*ForiT* using the QuikChange Mutagenesis kit (Stratagene, Cedar Creek, TX, USA). Wild-type TraI36 and variant proteins were expressed and purified as described (16). The 3' carboxytetramethylrhodamine (TAMRA)-labeled FTAM oligonucleotide (5'-TTTGCGTGGGGTGT[^]GGTGCTT-3') and the 3'-TAMRA-labeled T141'G variant of FTAM (5'-TTTGCGTGGGGTGG[^]GGTGCTT-3', with the base substitution underscored), were purchased from Integrated DNA Technologies, Coralville, IA, USA and purified as described (8). Variant *oriT* oligonucleotides used in competition binding assays were purchased from Integrated DNA Technologies and used without further purification.

DNA-binding affinity measurements

Affinities of TraI36 for wild-type and variant single-stranded *oriT* DNA sequences were measured in an affinity assay by determining the effect of the presence of unlabeled inhibitor oligonucleotide on binding affinity for the labeled FTAM oligonucleotide. Competitor oligonucleotide (at final concentrations as low as 5 nM for wild-type *oriT* and as high as 400 nM for low affinity binders) was mixed with 4 nM FTAM oligonucleotide in binding buffer (100 mM NaCl, 50 mM Tris pH 7.5, 0.1 mM EDTA) and TraI36 was titrated into the mixture. The resulting increase in fluorescence emission intensity of FTAM was measured in an AVIV (Lakewood, NJ, USA) ATF-105 automatic titrating fluorometer. K_I values (K_D values for the interaction of wild-type TraI36 with an unlabeled inhibitor oligonucleotide) were estimated by fitting volume corrected fluorescence emission data with the model described by Wang (17) using KaleidaGraph (Synergy Software, Reading, PA, USA). In these fits, the concentration of ligand was set at 4 nM and the K_D for the interaction between TraI36 and wild-type FTAM oligonucleotide was set at 0.6 nM. The method works well with high affinity interactions as assessed by the excellent agreement between our K_I value for the wild-type *ForiT* sequence (Table 1) and our previous K_D values for the same interaction measured by direct binding (8,18–20). Because of the nature of the inhibition assay, oligonucleotides that bind to TraI36 with K_I values greater than $\sim 3 \mu\text{M}$ do not affect the binding curve significantly and therefore their affinities cannot be reliably estimated. The relative K_I of these oligonucleotides is listed as $>3500 \text{ nM}$.

Table 1. Inhibition constants of *oriT* variant oligonucleotides

<i>oriT</i> variant	$K_1 \pm \text{SD}$ (nM)	n^a	Relative K_1^b	Relative transfer efficiency ^c
Wild-type <i>oriT</i>	0.44 ± 0.22	5	1.0	1.0
G140'A	11 ± 3.6	3	25	1.0
G140'C	74 ± 24	5	170	0.73
G140'T	40 ± 23	9	91	0.27
T141'A	>3500 ^d	3	>8000	<0.030
T141'C	43 ± 22	5	98	0.35
T141'G	~1000 ^e	3	~2300	<0.0024
G142'A	>3500 ^d	8	>8000	0.14
G142'C	800 ± 220	2	1800	0.16
G142'T	2400 ± 2000	6	5500	0.27
G144'A	3100 ± 2300	3	7000	0.12
G144'C	>3500 ^d	4	>8000	<0.0016
G144'T	945 ± 500	2	2100	0.24
T141'G/G140'A	370 ± 28	4	840	5.9
T141'C/G140'A	71 ± 37	5	160	0.44
T141'C/G140'C	410 ± 130	4	930	0.60
T141'C/G140'T	380 ± 170	4	860	0.19
G144'T/T141'A	435 ± 35	2	990	0.32
G144'T/T141'G	>3500 ^d	6	>8000	0.068

^aNumber of binding measurements.

^b K_1 value relative to K_1 for wild-type *oriT* sequence (variant K_1 /wild-type K_1).

^cTransfer efficiency of pACYC177-*ForiT* containing the listed sequence, relative to transfer efficiency of wild-type pACYC177-*ForiT*.

^dEstimated from absence of observed effect of 400 nM unlabeled variant oligonucleotide on binding of labeled wild-type oligonucleotide.

^eEstimated from binding assays using 4 nM labeled variant oligonucleotide performed because of unusual variability in measurements using inhibition assays.

One variant, T141'G, yielded highly variable K_1 values using the competition binding method described above. The affinity of TraI36 for the T141'G sequence was instead estimated from a direct binding assay using a 3'-TAMRA-labeled T141'G oligonucleotide as described (21).

Plasmid transfer assays

Plasmid mobilization assays were performed as described (8) with minor changes. Donors (ER2738 with wild-type or variant pACYC177-*ForiT*) and recipients (TB1) growing at log phase were washed in sterile PBS, mixed at a ratio of 1 donor to 9 recipients (final volume of 1 ml) and incubated for 1 min at ambient temperature. Transfer was stopped by vortexing cells and placing them on ice. Cells were serially diluted in PBS and plated on LB agar plates with tetracycline and ampicillin to detect donors; and streptomycin and ampicillin to detect transconjugants. Plasmid pACYC177 is a low copy number plasmid, present at approximately 15 copies per cell.

Oligonucleotide cleavage assays

The ability of TraI36 to cleave wild-type and variant single-stranded oligonucleotides was assessed as described (8) except that either 20 mM MgCl₂ or 2 mM MnCl₂ was added to the reaction.

Ligation assays

The ability of TraI36 to ligate single-stranded oligonucleotides was tested using wild-type or variant 17-base *oriT*

3'-Cy5-labeled oligonucleotides (Integrated DNA Technologies). A 20 μl reaction containing 1 μM TraI36, 4 nM 3'-Cy5-labeled wild-type (5'-TTTGCGTGGGGTGT^GGT-3', where ^ denotes the *nic* cleavage site) or G140'A (5'-TTTGCGTGGGGTGT^AGT-3', where the base substitution is underscored) *oriT* oligonucleotide, and either 20 mM MgCl₂ or 2 mM MnCl₂ in reaction buffer (100 mM NaCl and 20 mM Tris-HCl at pH 7.5) was incubated at 37°C for 60 min. Unlabeled wild-type (5'-CTTGTTTTTTCGTTGGGGTGT-3') or variant oligonucleotides consisting of the 20 bases 5' to the *oriT* nick site were then added to the reactions to final concentrations ranging from 20 nM to 20 μM. Samples were incubated at 37°C for 3 h, then the reactions were stopped by adding sodium dodecyl sulfate (0.1% final concentration) and incubating for 10 min at 37°C. Samples were applied to a 16% urea denaturing polyacrylamide gel (National Diagnostics, Atlanta, GA, USA) and results were visualized using a Typhoon 9410 Variable Mode Imager (GE Healthcare, Piscataway, NJ, USA).

RESULTS

Transfer efficiency of *oriT* variants

To examine the importance of the bases involved in the intrastrand base-pairing interaction seen in the TraI36:ssDNA crystal structure, base substitutions for T141', G142' and G144' were engineered into an *oriT* region cloned into pACYC177. Because G140' stacks with T141' in the complex and could potentially affect the position of T141' or the stability of interactions involving T141', we tested substitutions for G140' as well. We then measured the efficiency with which plasmids with variant *oriT* sequences having substitutions at one or two of these positions were mobilized to a recipient. The results are summarized in Figure 2 as fractions of the wild-type transfer efficiency. For some variants with reduced transfer efficiency, transfer was not observed in some or any of the experiments. For these, efficiencies were calculated as if a single transconjugant was observed and efficiencies are reported as being less than the calculated value. Those variants showing no transfer in some assays are marked with two asterisks, and those showing no transfer in any assay are marked with three asterisks. In our analysis, we assume that variants for which transfer was not detected in any assay have a lower transfer efficiency than those variants for which transconjugants were observed in at least some experiments.

Of the four positions tested, 140' was the most accommodating of substitutions. G140'T, the substitution causing the greatest reduction, reduced relative transfer efficiency to 0.27 ± 0.13 (27% of wild-type). This decrease is not statistically significant when significance is defined as a *P*-value < 0.05 by a two-sided student's *t*-test. Position 142' was less accommodating than 140', with G142'C (0.16 ± 0.06) and G142'A (0.14 ± 0.06) causing relatively small but significant reductions in transfer efficiency, and G142'T (0.27 ± 0.09) producing a smaller reduction.

Some single substitutions for G144' and T141' caused more profound effects than any substitution at the other

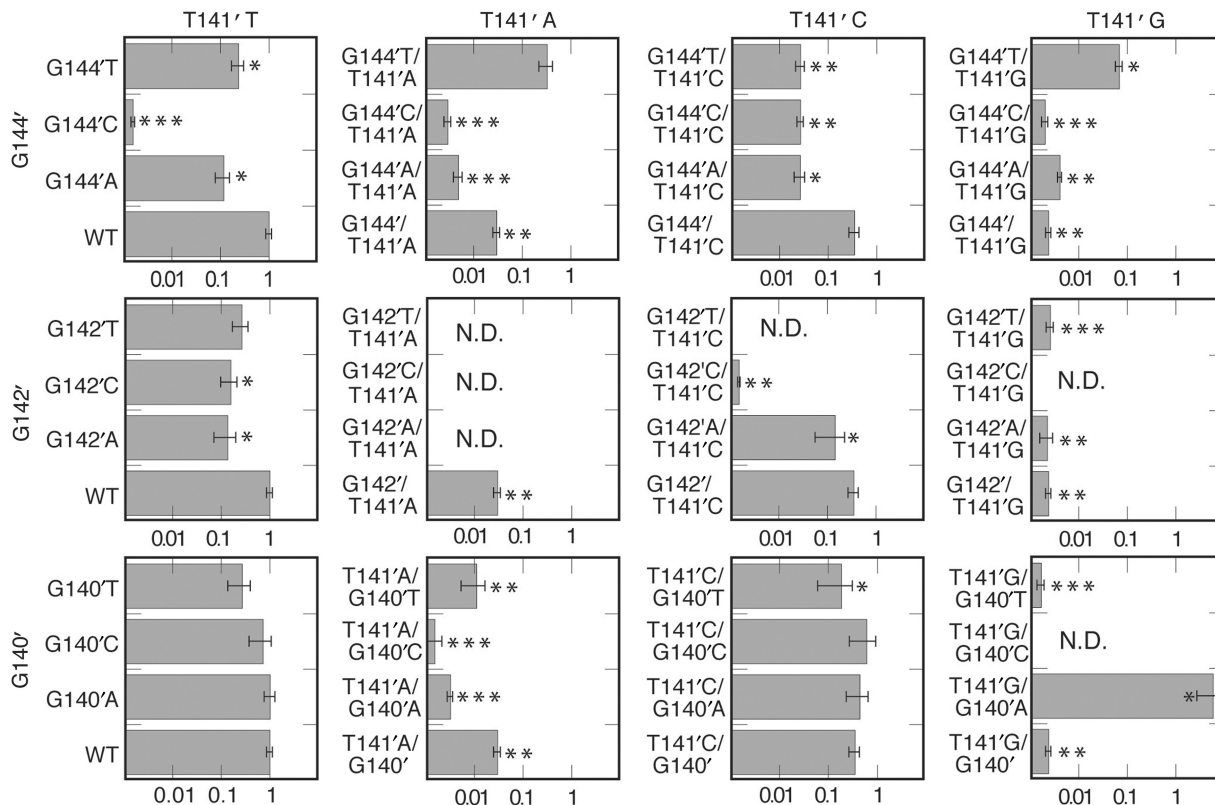


Figure 2. Transfer efficiencies of plasmids containing *oriT* variants. The results are grouped into columns based on the DNA base at position 141', and into rows based on the second position (144', 142' or 140') that was varied. Transfer efficiencies from 3 to 33 assays were averaged and the standard deviation of each measurement is shown as an error bar. Constructs with transfer reduced significantly relative to wild-type (where significance is defined as a *P*-value < 0.05 by a two-sided student's *t*-test) are marked with asterisks. Two asterisks indicate constructs for which one or more assay yielded no detectable transfer. Three asterisks indicate constructs that yielded no detectable transfer in any assay. For assays in which no transfer was observed, the upper limit of transfer was estimated by calculating efficiencies as if a single transconjugant was observed and these values are reported. ND means assay not done because construct not generated.

two sites. G144'C reduced transfer to undetectable levels (<0.2% of wild-type), and G144'T (0.24 ± 0.07) and G144'A (0.12 ± 0.04) also reduced transfer to a statistically significant degree. T141'G reduced transfer to <0.2% of wild-type (< 0.002 ± 0.0001 ; transfer not observed in some assays), and T141'A also significantly affected transfer (< 0.03 ± 0.005 ; transfer not observed in some assays) while T141'C had a relatively minor effect (0.35 ± 0.08).

Reasoning that a base substitution might be compensated for by a substitution at a second site, we generated and tested the transfer efficiency of plasmids containing most of the possible combinations of pairs of bases with substitutions at 141' and at 144', 142' or 140'. While many of the second substitutions had little or no effect, we observed improved transfer in four cases. The reduced transfer resulting from a purine at position 141' (T141'A < 0.03 ± 0.005 ; T141'G < 0.002 ± 0.0003) could be partially compensated for by a G144'T substitution (G144'T/T141'A = 0.32 ± 0.10 ; G144'T/T141'G = 0.07 ± 0.01). In another case involving these two positions, transfer of G144'C (< 0.002 ± 0.0001) is improved by T141'C (G144'C/T141'C < 0.03 ± 0.004 ; transfer not observed in some assays). Finally, in the largest improvement observed, the T141'G substitution (< 0.002 ± 0.0003) is

more than compensated for by a G140'A substitution (T141'G/G140'A = 5.8 ± 3.1).

Other second site substitutions decreased binding relative to single substitutions. Combining T141'A (< 0.03 ± 0.005) with G144'C (G144'C/T141'A < 0.003 ± 0.0005 ; transfer undetected), G144'A (< 0.005 ± 0.001 ; transfer undetected), G140'A (< 0.003 ± 0.0004 ; transfer undetected) or G140'C (< 0.002 ± 0.0006 ; transfer undetected) reduced transfer. Transfer also decreases when T141'C (0.35 ± 0.08) is combined with G144'T (< 0.03 ± 0.005 ; transfer not detected in some assays), G144'C (< 0.03 ± 0.004 ; transfer not detected in some assays), G144'A (0.03 ± 0.006) or G142'C (< 0.002 ± 0.0009 ; transfer undetected).

Binding affinity of *oriT* variants

The reduced transfer of the variants could be due to a reduced binding affinity of TraI for their *oriT* sequences. To test this, we measured the *in vitro* binding of a subset of the variant *oriT* sequences, as single-stranded oligonucleotides, to TraI36, the F TraI relaxase domain. The K_1 values for the sequences were estimated by measuring binding of TraI36 to a TAMRA-labeled wild-type *oriT* sequence in the presence of an unlabeled variant

oriT oligonucleotide. We use this method rather than using increasing concentrations of unlabeled variant oligonucleotide to compete with binding of the labeled oligonucleotide because we discovered that at high ($\sim 1 \mu\text{M}$) concentrations some unlabeled oligonucleotides could affect the fluorescence of the labeled oligonucleotide, perhaps due to interactions between the oligonucleotides. Using the assay described here, we obtain a K_I value for the wild-type *oriT* oligonucleotide (0.4 nM) within error of the measurement of the K_D measured by direct binding (0.6 nM). In contrast, the competition assay that we used before (20) yielded a somewhat higher K_I value for the same oligonucleotide (2.7 nM). Given the potential problems caused by use of high concentrations of oligonucleotides (noted above), the highest concentration of inhibitor we used was 400 nM. At this concentration of inhibitor, we estimate that the K_I value for an oligonucleotide would have to be lower than 3–4 μM for the binding of the labeled oligonucleotide to be measurably affected and the K_I value estimated.

The results from the binding assays are listed in Table 1 and representative curves and fits are shown in Supplementary Figure 1. The data and their relationship to the transfer efficiency of plasmids bearing the sequences are also depicted in Figure 3. In general, *oriT* sequences that facilitated transfer at a lower rate also were bound with lower affinities, consistent with a role for reduced TraI binding in reduced transfer efficiency. A reduction in affinity of 1000- to 10000-fold usually correlated with a reduction in transfer efficiency to 10–30% of wild-type.

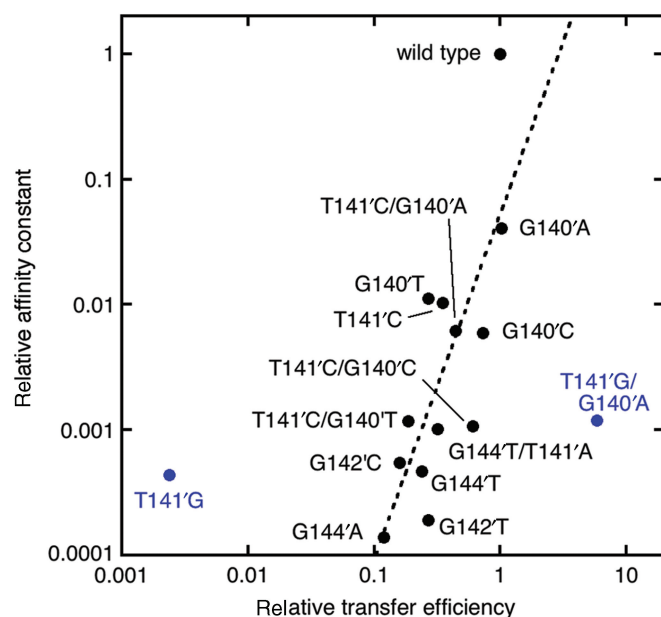


Figure 3. Correlation of reduced *in vitro* binding affinity of an oligonucleotide containing a variant *oriT* sequence and reduced transfer efficiency of a plasmid containing the variant sequence. The data for the variants are in black, except two outliers (T141'G and T141'G/G140'A) that are marked in blue. Note that T141'G is included even though its transfer efficiency could not be measured in all assays. The dashed line represents a power law fit to the data excluding the outliers performed using KaleidaGraph. The equation for the fit is $y = 0.05186 \cdot x^{(2.7535)}$ ($R = 0.6451$).

The relationship between transfer efficiency and affinity of TraI36 for the sequence is not simple, however. For example, plasmids containing wild-type or the G140'A *oriT* variant transfer with similar efficiencies despite a 25-fold difference in TraI36 affinity for the sequences. Plasmids containing G140'T and G142'T yield similar mobilization efficiencies despite a 60-fold difference in affinities. T141'G showed greater than expected binding ($K_I \sim 1 \mu\text{M}$) given that its transfer efficiency was too low to measure in some of the assays. In contrast, the T141'G/G140'A variant transfers with a 6-fold greater efficiency than wild-type, but the sequence is bound with an affinity that is reduced 370-fold relative to wild-type.

Comparison of the affinities and transfer efficiencies of the G140'A, T141'G and T141'G/G140'A variants indicates that it is possible for a second *oriT* substitution to dramatically enhance plasmid transfer without a large increase, and even a significant decrease, in ssDNA affinity. Clearly, while a minimum affinity of TraI for *oriT* is required, conjugative transfer has requirements beyond simple binding of the *oriT*.

Cleavage and ligation of *oriT* variants

We showed previously that the wild-type *oriT* sequence is cleaved efficiently *in vitro* by TraI36 while G144'C is both cleaved and bound poorly by TraI36 (8,15). In this case, the *in vitro* binding and cleavage correlates with the poor transfer of the G144'C *oriT* variant. It is conceivable though, that two variants having similar affinities for TraI36 but different transfer efficiencies could be explained if the sequences were cleaved or ligated by TraI with different efficiencies. We tested a subset of variant *oriT* sequences for their ability to be cleaved by TraI36 (Figure 4). Cleavage assays were performed with either 20 mM MgCl_2 or 2 mM MnCl_2 in the reaction mixture. TraI36 binds Mn^{2+} with higher affinity than Mg^{2+} and has greater apparent activity with Mn^{2+} (12). Using both metals allows us to examine a greater range of activity than would be possible using a single metal.

As shown in Figure 4, TraI36 cleaves a single-stranded wild-type *oriT* oligonucleotide readily, while the T141'G variant oligonucleotide is cut poorly in the presence of either Mg^{2+} or Mn^{2+} . The G144'T/T141'G variant is not detectably cleaved in the presence of Mg^{2+} , but is cleaved in the presence of Mn^{2+} , albeit with reduced efficiency relative to the wild-type sequence. The results suggest that the G144'T/T141'G variant, despite being bound by TraI36 under assay conditions with an affinity too low to reliably measure, can adopt a conformation conducive to cleavage by TraI36 while the T141'G cannot. In contrast, TraI36 cleaves the T141'G/G140'A variant oligonucleotide poorly even in the presence of Mn^{2+} despite binding to TraI36 with a higher affinity than either T141'G or G144'T/T141'G. The enhanced transfer of the T141'G/G140'A variant, relative to wild-type and the T141'G variant, cannot therefore be explained by either enhanced binding affinity or enhanced cleavage.

Results of ligation assays (Supplementary Figure 2) were similar to the results of the cleavage reactions. TraI36 generated the G144'T/T141'G ligation product

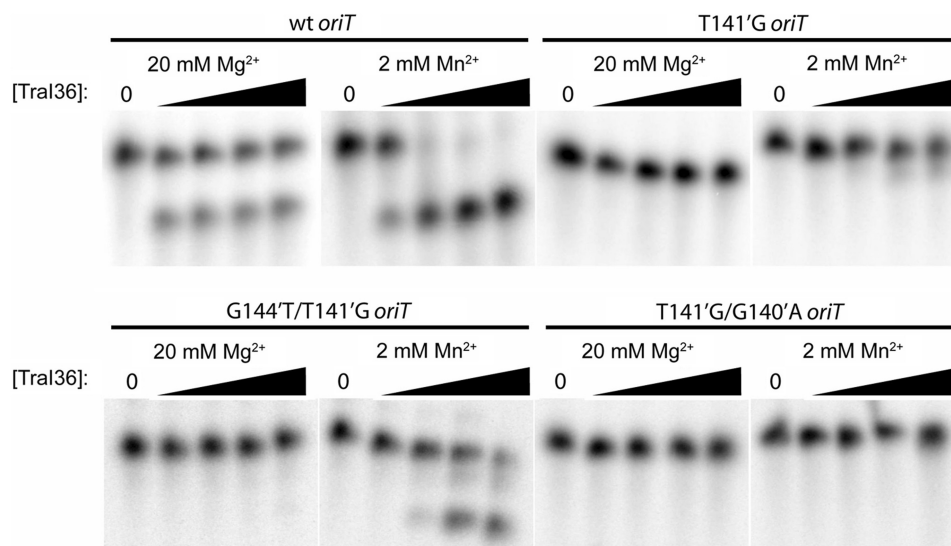


Figure 4. Cleavage activity of TraI36 against selected *oriT* variant oligonucleotides. 32 P-5'-end-labeled 22-base oligonucleotides were incubated with 0, 1, 10, 100 or 1000 nM TraI36 in the presence of either 20 mM $MgCl_2$ or 2 mM $MnCl_2$. Oligonucleotide cleavage is indicated by appearance of the faster migrating 14-base band.

markedly more efficiently than either the T141'G or the T141'G/G140'A product.

DISCUSSION

To better understand the functional role of the intrastrand base-pairing interactions observed in the TraI36:ssDNA crystal structure, we studied the effect of *oriT* base substitutions on conjugative transfer. In the base-pairing interaction (Figure 1), T141' forms hydrogen bonds with G144'. Not surprisingly, some substitutions in these positions were poorly tolerated. For example, single and double substitution variants that contained purine substitutions for T141' generally showed reduced transfer. All variants with purines at both 141' and 144' had significantly reduced transfer efficiency. A simple explanation for these results is that the volume of the purines is difficult to accommodate within the tight confines of the binding site, and the resulting steric clashes reduce affinity. In addition, the scissile phosphate is located between bases 141' and 140'. Even relatively small structural adjustments made to accommodate the purines could move the scissile phosphate away from the catalytic Tyr, reducing the efficiency of cleavage.

Other observations are more difficult to explain. The G144'C substitution interferes with transfer, with no combination of G144'C and any base at 141' transferring with better than 3% of wild-type efficiency. The structural basis for the inefficient transfer and poor *in vitro* binding of this variant is not apparent when examining the TraI36:ssDNA crystal structure. There are no obvious steric clashes that would result from the substitution. In addition, we cannot identify any intermolecular or intramolecular ssDNA interactions that might arise from the substitution that would interfere with binding by TraI.

We also examined the effect of substitutions on *in vitro* oligonucleotide binding, cleavage and ligation, and

compared these results to the effects of the substitutions on conjugative transfer. Based on the results presented here, we can draw several conclusions. First, as shown in Figure 3, a reduced *in vitro* binding affinity of TraI36 for a single-stranded variant oligonucleotide generally correlates with a reduced transfer efficiency of a plasmid bearing the same variant sequence. Some previous results from this lab on a small set of variants suggested this same correlation (8). The result is not surprising in the sense that DNA cleavage by TraI is essential for F conjugative transfer, and TraI, of course, must bind *oriT* DNA before it can cleave it. We do not know, however, how closely the fluorescently labeled single-stranded oligonucleotide used in the binding assay resembles the actual *in vivo* substrate. The data presented here suggest that the resemblance is sufficient for the *in vitro* binding affinity to reflect on the *in vitro* transfer efficiency.

The second conclusion is that when *in vitro* oligonucleotide binding and transfer efficiency do not correlate, the discrepancy may reflect a bound conformation of the oligonucleotide that is not conducive to cleavage and ligation. The G144'T/T141'G variant transfers with significantly higher efficiency than T141'G, yet the G144'T/T141'G sequence is bound with an affinity too low for us to measure ($K_I > 3500$ nM) and T141'G binds with an ~ 1000 nM K_D . The G144'T/T141'G variant oligonucleotide, though, is cleaved and ligated with a greater efficiency than is the T141'G oligonucleotide. We believe that although the G144'T/T141'G sequence is bound with lower affinity than the T141'G sequence, the conformation of the G144'T/T141'G oligonucleotide better positions the scissile phosphate for cleavage by the active site Tyr. Although even a subtle structural change could render an oligonucleotide a poor relaxase cleavage substrate, work with the telomere end-binding protein from *Oxytrichia nova* (*OnTEBP*) has shown that, at least for this protein, cognate and non-cognate ssDNA sequences

can bind with similar affinities but with some dramatic conformational differences. Theobald and Schultz examined the affinity of *OnTEBP* for 10 non-cognate ssDNA sequences and determined the structures of the protein:ssDNA complexes (22). Based on the range of conformational differences between the bound cognate and non-cognate sequences, the authors conclude that it would be difficult to predict the structural and functional effects of the base substitutions starting from the structure of *OnTEBP* with its cognate-binding site. We suspect that the TraI-binding site, with its combination of a binding groove and a flap that folds over the bound ssDNA, may be somewhat less forgiving of base substitutions in its ssDNA substrate than is the *OnTEBP* binding cleft. The flexibility of the ssDNA substrate, however, certainly complicates predicting the effect of base substitutions on TraI function, just as it does for *OnTEBP*.

The third conclusion is that the requirements for efficient transfer extend beyond simple ssDNA sequence recognition and cleavage. The T141'G/G140'A variant transfers with an efficiency several-fold higher than wild-type, yet the T141'G/G140'A variant oligonucleotide is bound with 100-fold lower affinity than wild-type and is cleaved and ligated inefficiently. While we have no convincing explanation for the enhanced transfer efficiency of T141'G/G140'A, there are several ways in which altering the *oriT* sequence could alter the efficiency of transfer. Changing the base composition near *nic* could change the accessibility of the *oriT* sequence bound by TraI. Presumably TraI recognizes its site when the region around *nic* is in a melted or bubble conformation, and base changes that alter the melting temperature or local conformation of the region could affect TraI access and therefore cleavage and transfer. Another possibility is that the complex of TraI and a variant *oriT* sequence, following cleavage and concurrent formation of the covalent linkage between TraI and the transferred strand, may have an altered stability. Altered stability of this complex could mean that an appropriate TraI-ssDNA conjugate would not be available to initiate transfer when a stable mating pair was formed. Alternatively, altered stability of the complex could prevent the TraI-ssDNA conjugate from converting from its role as a stable component of the relaxosome to its role as an active participant in transfer once the stable mating pair was formed. Finally, proteins such as F TraM, F TraY and IHF participate in the cleavage of *oriT* and initiation of F conjugal transfer. Some substitutions near *nic* could have unanticipated effects on the binding of these other proteins, or could affect their participation in *oriT* cleavage or in transfer initiation. These effects could in turn lead to reduced transfer.

Given the apparent importance of the intramolecular DNA interactions to binding and cleavage by TraI, we might suspect that intrastrand base pairing is a common feature of relaxase:ssDNA interactions. Although there is not yet published work to support the importance of such interactions for other relaxases, a similar intramolecular interaction is observed in one structure of the distantly related R388 TrwC relaxase domain with bound ssDNA (14). In addition, the three interacting bases in the F *oriT*

are conserved among a number of F-like and other plasmids (3,23), suggesting that their relaxases might use a similar type of recognition. The recently determined structure of the relaxase domain of R1162 MobA shows that it shares a similar structure with the relaxase domains of R388 TrwC and F TraI, despite having little obvious sequence identity (24). Unfortunately, the MobA structure has no bound DNA, preventing comparison of the conformations of the bound DNA. Structures of some other ssDNA binding proteins show that intrastrand base interactions are formed in their complexes. In the best example, the *Schizosaccharomyces pombe* Pot1 (protection of telomeres 1) protein uses intrastrand base-pairing interactions to stabilize a compact ssDNA structure essential for high affinity binding by this protein (25). While this intrastrand base-pairing may not be used by many proteins, the evidence from *S. pombe* Pot1 and F TraI demonstrate that it can be a valuable mechanism to determine binding specificity of a ssDNA-binding protein.

In sum, the results we present suggest that reducing the affinity of a sequence for TraI or impairing its cleavage reaction can reduce its transfer efficiency. Of these two requirements, the cleavage efficiency is probably the more stringent of the two. Our results also underscore the complexity of the role of both relaxase and *oriT* in transfer, and point out that there are aspects of these roles that are both important to the transfer process and currently not fully appreciated.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Christine DeGennaro for engineering the *oriT* plasmid used in this work and Lubomir Dostal for helpful discussions of cleavage and ligation reactions. This manuscript is based on work supported by National Institutes of Health grant GM61017 to J.F.S. Funding to pay the Open Access publication charges for this article was provided by National Institutes of Health grant GM61017 to J.F.S.

Conflict of interest statement. None declared.

REFERENCES

1. Wirth,T., Falush,D., Lan,R., Colles,F., Mensa,P., Wieler,L.H., Karch,H., Reeves,P.R., Maiden,M.C. *et al.* (2006) Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol. Microbiol.*, **60**, 1136–1151.
2. Hill,V.T., Townsend,S.M., Arias,R.S., Jenabi,J.M., Gomez-Gonzalez,I., Shimada,H. and Badger,J.L. (2004) TraJ-dependent *Escherichia coli* K1 interactions with professional phagocytes are important for early systemic dissemination of infection in the neonatal rat. *Infect. Immun.*, **72**, 478–488.
3. Zechner,E.L., de la Cruz,F., Eisenbrandt,R., Grahn,A.M., Koraimann,G., Lanka,E., Muth,G., Pansegrau,W., Thomas,C.M. *et al.* (2000) Conjugative-DNA transfer processes. In Thomas,C.M. (ed), *The Horizontal Gene Pool*, Harwood Academic Publishers, Amsterdam, pp. 87–174.

4. Waters, V.L., Hirata, K.H., Pansegrau, W., Lanka, E. and Guiney, D.G. (1991) Sequence identity in the nick regions of IncP plasmid transfer origins and T-DNA borders of *Agrobacterium* Ti plasmids [published erratum appears in *Proc Natl Acad Sci USA* 1991 Jul 15;88(14):6388]. *Proc. Natl Acad. Sci. USA*, **88**, 1456–1460.
5. Willetts, N. and Maule, J. (1979) Investigations of the F conjugation gene *traI:traI* mutants and *lambda*traI transducing phages. *Mol. Gen. Genet.*, **169**, 325–336.
6. Fekete, R.A. and Frost, L.S. (2000) Mobilization of chimeric *oriT* plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. *J. Bacteriol.*, **182**, 4022–4027.
7. Willetts, N. and Maule, J. (1986) Specificities of IncF plasmid conjugation genes. *Genet. Res.*, **47**, 1–11.
8. 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.
9. 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.
10. Pansegrau, W., Ziegelin, G. and Lanka, E. (1990) Covalent association of the *traI* gene product of plasmid RP4 with the 5'-terminal nucleotide at the relaxation nick site. *J. Biol. Chem.*, **265**, 10637–10644.
11. Draper, O., Cesar, C.E., Machon, C., de la Cruz, F. and Llosa, M. (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc. Natl Acad. Sci. USA*, **102**, 16385–16390.
12. 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 (Camb)*, **13**, 1533–1544.
13. Frost, L.S., Ippen-Ihler, K. and Skurray, R.A. (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.*, **58**, 162–210.
14. 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.
15. 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.
16. Street, L.M., Harley, M.J., Stern, J.C., Larkin, C., Williams, S.L., Miller, D.L., Dohm, J.A., Rodgers, M.E. and Schildbach, J.F. (2003) Subdomain organization and catalytic residues of the F factor TraI relaxase domain. *Biochim. Biophys. Acta*, **1646**, 86–99.
17. Wang, Z.X. (1995) An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS Lett.*, **360**, 111–114.
18. 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.
19. Harley, M.J., Toptygin, D., Troxler, T. and Schildbach, J.F. (2002) R150A mutant of F TraI relaxase domain: reduced affinity and specificity for single-stranded DNA and altered fluorescence anisotropy of a bound labeled oligonucleotide. *Biochemistry*, **41**, 6460–6468.
20. Stern, J.C., Anderson, B.J., Owens, T.J. and Schildbach, J.F. (2004) Energetics of the sequence-specific binding of single-stranded DNA by the F factor relaxase domain. *J. Biol. Chem.*, **279**, 29155–29159.
21. Larkin, C., Haft, R.J., 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.
22. Theobald, D.L. and Schultz, S.C. (2003) Nucleotide shuffling and ssDNA recognition in *Oxytricha nova* telomere end-binding protein complexes. *EMBO J.*, **22**, 4314–4324.
23. 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. *Nat. Struct. Biol.*, **10**, 1002–1010.
24. 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.
25. Lei, M., Podell, E.R., Baumann, P. and Cech, T.R. (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature*, **426**, 198–203.