Structural and functional characterization of TraI from pKM101 reveals basis for DNA processing

Annika Breidenstein[1,2], Josy ter Beek[1,2, #] and Ronnie P-A Berntsson[1,2, #]

[1] Department of Medical Biochemistry and Biophysics, Umeå University, SE-90187 Umeå, Sweden

[2] Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden

[#] Correspondence should be addressed to J.t.B (email: josy.beek@umu.se) or R.P-A.B. (email: ronnie.berntsson@umu.se)

Abstract

To support conjugation via Type 4 Secretion Systems, relaxases perform the initial processing of the substrate DNA. TraI from the well-studied conjugative plasmid pKM101 is one such relaxase, which has not yet been characterized. Here, we report the crystal structure of the trans-esterase domain of TraI in complex with its substrate *oriT* DNA, highlighting the conserved DNA binding mechanism. Additionally, we present an apo structure of the trans-esterase domain of TraI, which allows us to visualize the conformational changes involved in DNA binding. We also show a complete characterization of the DNA binding, nicking and religation of the trans-esterase domain, helicase domain and full-length TraI. Furthermore, TraI is shown to be a tetramer. These results reveal that the trans-esterase domain behaves in a similar way as its homologs TraI from the F plasmid and TrwC from R388, but the tetramerization of the full-length protein highlights a significant difference that affects the function of TraI.

Introduction

Antibiotic resistance is one of the most pressing health challenges in today's world. An important contribution to this problem is the ability of pathogens to spread resistance genes using horizontal gene transfer. This is a process that is often facilitated by conjugative Type 4 Secretion Systems (T4SSs). These are highly versatile systems that transport DNA and proteins from a bacterial donor cell into a recipient cell. In recent years, researchers have generated structural and functional insight into these systems, especially from Gram-negative (G-) bacteria, such as the T4SSs from the *E. coli* F-plasmid and the pKM101 plasmid, as well as the *Legionella pneumophilia* Dot/Icm T4SS[1–3]. The T4SS of pKM101 belongs to the class of minimal T4SS that consist of twelve proteins homologous to the paradigmatic VirB/VirD4 T4SS from *Agrobacterium tumefaciens*.

For all known T4SSs that transfer conjugative plasmids, the plasmid DNA must be processed before it can be transported. This is done via the relaxosome complex, which consists of a relaxase, accessory factor proteins and the DNA. To form this complex, one or several accessory proteins bind to the origin of transfer (*oriT*) on the DNA and locally melt the double stranded DNA to promote relaxase binding to a defined sequence, often forming a hairpin, close to the nicking site[4]. The relaxase binds this single stranded *oriT* DNA via its N-terminal trans-esterase domain[5]. This domain reacts with the DNA via a transesterification reaction at the specific *nic*-site, which generates the transfer intermediate consisting of the relaxase covalently bound to the 5' end of the cleaved transfer strand (T-strand)[6]. Many relaxases have a second functional domain in the more variable C-terminal part. This is often a helicase domain, which is responsible for unwinding the DNA to allow for transport of the single stranded transfer DNA[5]. The relaxase-T-strand complex (T-complex) is recruited to the T4SS by the type IV coupling protein (T4CP) and is transported through the T4SS channel into the recipient cell[7]. Once present in the recipient cell, the trans-esterase domain religates the DNA to regenerate the circularized double-stranded plasmid.

Relaxases have been phylogenetically classified into eight MOB families: MOB_F, MOB_H, MOB_Q, MOB_C, MOB_P, MOB_V, MOB_T and MOB_B. Of these, pKM101-encoded TraI ($TraI_{pKM101}$) belongs to the MOB_{F11} subclade together with its closest relative TrwC from plasmid R388 ($TrwC_{R388}$) and TraI from the sister-plasmid pCU1[5,8,9]. While structural information is available for the trans-esterase domains of several MOB_F relaxases, structural

data is available only for the substrate DNA bound states of $TrwC_{R388}$ and F-plasmid TraI ($TraI_F$), which belongs to a different subclade MOB_{F12}[10,11]. Furthermore, there is only very limited data on the full-length relaxases, with most information available from $TraI_F$. $TraI_F$ consists of one trans-esterase domain, a vestigial helicase domain and an active helicase domain (Fig. 1A). $TraI_F$ binds *oriT* as a heterogenous dimer. One $TraI_F$ subunit binds *oriT* DNA at a hairpin positioned 5' of the *nic*-site with the trans-esterase domain in an open conformation, while the other subunit binds ssDNA at the helicase domains in a closed conformation. These two states are incompatible with each other in a single subunit[12]. A cryo-EM structure of this second "helicase" state is available and shows how the single stranded DNA is almost entirely surrounded by the helicase domains. In this structure, all of the protein except the very flexible C-terminal domain can be seen[12]. It is not clear to what extent this information applies to $TraI_{pKM101}$, due to important differences between the proteins. $TraI_{pKM101}$ is shorter with only 1078 residues compared to $TraI_F$ which consists of 1756 residues. This is a result of a shorter C-terminal domain in $TraI_{pKM101}$ as well as the absence of the second, vestigial helicase domain of $TraI_F$ (Fig. 1A).

In this study we present the crystal structure of the trans-esterase domain of TraI from pKM101 bound to its substrate *oriT* DNA. This structure confirms the highly conserved DNA binding mode of the trans-esterase domain observed in other MOB_F family relaxases. We also present the apo-structure with the flexible thumb-domain of the protein visible, showing a large movement between the apo and substrate-bound structures. TraI is characterized via electrophoretic mobility shift assays, nicking- and religation assays, and we present data that shows that $TraI_{pKM101}$ acts as a tetramer.

Materials and Methods

Plasmids

The full-length *traI* gene, and shorter constructs for the trans-esterase domain (TraI-TE) (residues 1-299) and the helicase domain (TraI-H) (residues 429-910) were PCR amplified and cloned into p7XC3H and p7XNH3 via the FX cloning system[13] under the T7 promotor, with an N- or C-terminal 10-HIS tag and a 3C protease sequence (see Table S1).

Protein expression and purification

All TraI variants were produced in *E. coli* BL21 (DE3). The cells were grown in TB medium at 37°C using a LEX (Large-scale Expression) bioreactor (Epiphyte3). Once the cultures reached an OD_{600} between 1.0 and 1.5, the temperature was reduced to 18°C and expression was induced by adding 0.4 mM IPTG for ~16 hours. Cells were spun down at 6000 g and pellets were resuspended in resuspension buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 15 mM Imidazole, 0.2 mM AEBSF, and ca 0.02 mg/mL DNase I). Resuspended cells were lysed in a Cell Disruptor (Constant Systems) at 25 kPsi at 4°C, followed by centrifugation at 30,000 g for 30 min at 4°C. The cell lysate was then incubated for 1h at 4°C with Ni-NTA resin (~2 mL/L culture for TraI and TraI-H or ~4 mL/L culture for TraI-TE) before transfer to a gravity flow column. The resin was subsequently washed with 10 column volumes (CV) of wash buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 50 mM Imidazole, 0.2 mM AEBSF), followed by 10 CV wash buffer with 2 M LiCl and another 10 CV wash buffer. TraI (full-length) was eluted with 5 mL/CV elution buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 500 mM Imidazole).

For TraI-TE and TraI-H, 3C PreScission protease was added, while the protein was still bound to the IMAC resin, in an estimated ratio of 1:50 to 1:100 to remove the HIS-tag and incubated ~16 hours at 4°C. The cleaved protein was then recovered by collecting the flow through.

All proteins were 3 x diluted with dilution buffer (25 mM HEPES pH 7.0, 50 mM NaCl) after the Ni-NTA purification and loaded on a HiTraP Heparin HP (5 mL) column, equilibrated with Buffer A (25 mM HEPES pH 7.0, 150 mM NaCl) with a peristaltic pump. After loading the protein, the column was connected to an ÄKTA Pure (Cytiva) and washed with Buffer A for ~2 CV until absorbance 280 was stable. All TraI variants were subsequently eluted via a salt gradient to 100 % Buffer B (25 mM HEPES pH 7.0, 1.0 M NaCl) over 80 mL. Peak fractions with pure protein, as judged by the 260/280 ratio, were concentrated using Amicon Ultra centrifugal filters with molecular weight cutoffs at 50 kDa (TraI), 30 kDa (TraI-H) or 10 kDa (TraI-TE). Size exclusion chromatography was performed in 25 mM HEPES pH 7.0, 300 mM NaCl on a Superdex 200 10/300 GL Increase column (TraI-TE and TraI-H) or a Superose 6 10/300 GL Increase column (TraI).

Crystallization and structure determination

Protein crystals were obtained with the sitting drop vapor diffusion method at 20°C. Substrate DNA, *oriT*11 (Fig. 1B & Table S1) was added to TraI-TE, to yield a final concentration of 9 mg/mL TraI-TE and an equimolar amount of the DNA. This mixture was used for co-

crystallization experiments, which yielded both the apo structure of TraI-TE and the DNA-bound structure. The apo structure crystallized in 10 % w/v PEG 4000, 20 % v/v glycerol, 0.03 M $MgCl_2$, 0.03 M $CaCl_2$ and 0.1 M MOPS/HEPES-Na pH 7.5. The DNA-bound structure crystallized in 10 % w/v PEG 20000, 20 % v/v PEG MME 550, 0.03 M sodium nitrate, 0.03 M disodium hydrogen phosphate, 0.03M ammonium sulfate and 0.1 M MES/imidazole pH 6.5. Crystals were flash-frozen in liquid nitrogen, without the addition of extra cryo-protectant. X-ray diffraction data was collected on ID30B, ESRF, France. The data was processed using XDS[14,15]. Both crystals had the space group P212121 and contained a single copy of the protein in the asymmetric unit. The phase-problem was solved using molecular replacement with the trans-esterase domain of $TraI_{pCU1}$ (PDB: 3L6T) in PHENIX phaser[16]. The structures were built in Coot[17] and refined at 2.1 Å (DNA bound structure) and 1.7 Å (apo structure) in PHENIX refine[18], to R_{work}/R_{free} values of 22.5/26.9 % and 17.1/19.1 % respectively. Further refinement statistics can be found in Table S2. Atomic coordinates and structure factors of both the apo and DNA-bound structure of TraI have been deposited with the Protein Data Bank (PDB: 8A1B and 8A1C).

Size-exclusion chromatography coupled to multiangle light scattering (SEC-MALS)

250 µl of protein sample, at a minimum concentration of 1 mg/mL, were run over either a Superose 6 10/300 GL Increase column or a Superdex200 10/300 GL Increase column. Experiments were performed in 25 mM HEPES pH 7.0, 300 mM NaCl, on an ÄKTA Pure (Cytiva) coupled to a light scattering (Wyatt Treas II) and refractive index (Wyatt Optilab T-Rex) detector. Data was collected and analyzed using Astra software (Wyatt Technology; version 7.2.2) as described by Some *et al.* 2019[19]. The molecular weight of the protein samples were calculated as an average from a minimum of 3 measurements and reported with the standard deviation.

Native gels for oligomerization determination

20 µl of TraI SEC fractions were diluted with 2xNative Tris-Glycine Sample buffer (Novex) and loaded on precast native gradient gels (Novex™ WedgeWell™ 4-20 % Tris-Glycine, Thermo Fisher Scientific) and electrophoresed in 1x Tris-Glycine Native Running Buffer (Novex) at 4°C for 90 min at 150 V. The gels were stained with Coomassie blue and visualized using a ChemiDoc Imaging system (Bio Rad).

DNA oligomers used for binding assays and crystallization

Single stranded DNA oligomers were purchased from Eurofins Genomics, with and without the fluorescent label fluorescein isothiocyanate (FITC). The position of the label at the 5' or 3' end is indicated in Table S1 and a schematic of the different *oriT* DNA used in this study can be found in Fig. 1B. Stock solutions were dissolved in MQH_2O, but all further dilutions were done in 25 mM HEPES pH 7.0, 300 mM NaCl. Oligomers were heated at 95°C for 5 min and cooled down to room temperature over at least 20 minutes prior to use.

Electrophoretic mobility shift assay

50 nM fluorescently labeled DNA (Table S1) was incubated for 15 min with 0 to 1600 nM of protein (see results) in 25 mM HEPES pH 7.0, 300 mM NaCl at room temperature before adding 6x native loading dye (3x TBE, 30% glycerol, 0.125% bromophenol blue). The samples were resolved on a native gel (5% polyacrylamide, 0.75x TBE) in 0.75x TBE running buffer at 50 V for 90 min at 4°C. Gels were imaged on an Amersham Typhoon 5 scanner with excitation at 488 nm, using the Cy2 emission filter (515-535 nm).

Nicking and religation assays

Nicking and religation assays were performed in 25 mM HEPES pH 7.0, 300 mM NaCl, 5 mM $MgCl_2$. Nicking assays contained 100 nM F-*oriT*57, religation assays (version 1) contained 100 nM F-*oriT*57 and 100 nM F-*oriT*20, religation assays (version 2) contained 100 nM F-*oriT*35 and 100 nM non-fluorescent *oriT*57 (Fig. 1B & Table S1). DNA oligomers were incubated with increasing concentrations of TraI in a heatblock at 37°C, rotating at 300 rpm, for 1h. The reactions were stopped by adding 2X stop solution (96 % formamide, 20 mM EDTA, 0.1 % bromophenol blue). The samples were boiled for 5 min at 95°C before loading on a denaturing gel (16 % polyacrylamide, 7 M Urea, 1X TBE). The gel was run at 50 V for 15 min followed by 100 V for 1h at room temperature and then imaged on an Amersham Typhoon 5 scanner with excitation at 488 nm, using the Cy2 emission filter (515-535 nm).

Results

Crystal structure of TraI trans-esterase domain

The TraI trans-esterase domain (TraI-TE) (33 kDa) was expressed in *E. coli* with a C-terminal HIS-tag and purified to homogeneity in a three-step process during which the tag was removed. This purified TraI-TE was co-crystallized with single stranded 11-mer *oriT* DNA (*oriT*11) (Fig. 1B). The crystals grew in space group P212121 and contained one molecule in the asymmetric unit. The crystallographic phase problem was solved with molecular replacement using the previously solved structure of the trans-esterase domain of pCU1 (PDB: 3L6T) as a search model, and the crystal structure was refined at a resolution of 2.1 Å. Virtually all residues were visible in the electron density, with the exception of five residues in flexible loop regions and at the C-terminus, and we were able to place ten of the eleven bases of the bound DNA.

The overall architecture of TraI-TE can be described as a hand, with the bound DNA located between the palm and thumb domain (Fig. 2A). The structure is similar to the previously described homologs $TrwC_{R388}$, and $TraI_F$ which are the only previously solved structures of MOB_F family relaxases with bound DNA (Fig. 2B)[10,11,20]. As in the homologs, a histidine triad coordinates a divalent metal that is required for the trans-esterase reaction. This triad in TraI-TE consists of H149, H160 and H162 (Fig. 2C). Electron density was observed in the middle of this triad and was best fitted with a manganese ion. Like $TrwC_{R388}$, the conserved aspartate, D84 in $TraI_{pKM101}$, is situated within H-bonding distance of both H162 and Y18 (2.8 Å and 2.7 Å respectively). It is therefore in the position to activate the catalytic tyrosine Y18, the first tyrosine of the conserved YY-$X_{(5-6)}$-YY motif[4,5], during the trans-esterification reaction with the scissile phosphate at the *oriT nic*-site.

DNA binding

The DNA substrate used for co-crystallization was *oriT*11, representing the 11 bases directly 5' of the *nic*-site. Of these, the first base at the 5' end, T1, did not have any defined electron density, likely because it was not properly stabilized due to the missing hair-pin that would follow in a longer *oriT* structure. While the first visible base, C2, does not show any specific contacts with TraI-TE, the following guanine G3 is forming 4 H-bonds with R80 and D182 using both the Hoogsteen and Watson-Crick edge, and thymine T4, is forming 2 H-bonds with N181. The following two bases T5 and A6 are stacked on top of each other perpendicular to the orientation of the previous three bases and forming H-bonds with K188 (T5) and Q256

(A6). These initial bases are oriented in a position that is pointing away from the active site. The following bases place the DNA in a conformation resembling a U-turn that allows for the scissile phosphate to be placed close to the catalytic tyrosine. This conformation is stabilized in several ways; with hydrophobic and pi-pi interactions between the stacked bases G8 and G10 as well as multiple H-bonds. G7 forms two H-bonds with L2 at its Watson-Crick edge and an additional one with K196 from its Hoogsteen edge. G8 forms two Watson-Crick H-bonds with D3 and G10 forms two Hoogsten H-bonds with R250. The phosphates of the three most 3' bases are stabilized with several H-bonds (S236 for T9, S236 and K91 for G10, and R235 and R153 for T11) (Fig. 2D). Taken together, these DNA-DNA and DNA-protein interactions illustrate a binding mechanism reliant on DNA secondary structure and sequence specificity.

Apo structure reveals rearrangement of the thumb domain upon DNA binding

An important feature of DNA binding is the thumb domain, consisting of two α-helices connected by a loop region. This domain lies on top of the 3' part of the DNA and forms H-bonds with the *oriT* DNA with R235, R250 and Q256 (Fig. 2A & D). The thumb domain is highly flexible and has previously only been observed in relaxase structures with bound DNA. However, our co-crystallization experiment resulted in an additional crystal with the space group P212121 that resolved to 1.7 Å. This structure had a different unit cell, which had the fortuitous outcome of supporting an apo state of TraI-TE, in which we were able to build a large part of the thumb domain in an open conformation (Fig. 3A). The electron density corresponding to residues 261-271 was of insufficient quality to allow for building those residues in the model, but inspection of the density strongly suggests a continued α-helix towards the base of the thumb. Between the apo and the DNA bound state, the thumb moves up to 40 Å, and undergoes important structural changes. The conformational differences include a reorganization of α-helix 8 and the extension of α-helix 9 (numbered based on the DNA-bound structure, one helix disappears in the apo structure) to transition from the apo to the DNA bound structure, as illustrated by the color scheme of the thumb in Fig. 3A. These rearrangements reveal the structural changes that are necessary for DNA binding and highlight the importance of the thumb domain. The C-terminal helix at the base of the thumb domain (α-helix 10) interacts with the loop at the end of α-helix 1 via an H-bond between R274 and D25 in both the DNA-bound and apo-structures (Fig. S1A). An additional interaction is made between H278 and Y26, tyrosine three of the YY-$X_{(5-6)}$-YY motif, but only in the DNA bound structure.

Full-length TraI forms a concentration dependent tetramer

While the TraI trans-esterase and helicase domains both elute as single peaks from the size-exclusion column (SEC), several peaks were visible in the elution profile of the full-length protein, likely representing different oligomeric species of TraI. This oligomerization is concentration dependent, as the ratio of the earliest peak (largest oligomer) is increased at higher protein concentrations. To investigate this oligomerization further, we determined the molecular weight of the main peaks of TraI using SEC-MALS[19]. The higher oligomer had a molecular weight of 485 +/- 10 kDa (Fig. 4A), which matches well to the expected molecular weight of a tetramer (492 kDa). The 2nd main peak from the SEC had a molecular weight of 123 +/- 5 kDa, which corresponds to the weight of a TraI monomer (Fig. S2A). The molecular weight was in the same manner also determined of the TraI trans-esterase domain and the helicase domain, both of which were confirmed to be monomeric with 33 +/- 1 kDa and 49 +/- 2 kDa respectively (Fig. S2B & C). To further investigate the oligomeric state, TraI from both the tetrameric and monomeric fraction were run on native gels. Both fractions migrated at the expected size of the tetramer (Fig. 4B). This indicates that the tetramer is the preferred oligomerization state of TraI.

TraI is a functional relaxase capable of binding, nicking and religating *oriT*-DNA

We performed several activity assays to biochemically characterize TraI. To estimate differences in binding affinity and specificity of the different TraI domains we used electrophoretic mobility shift assays (EMSAs). We investigated DNA binding to a fluorescent 22-mer with the post-*nic* DNA sequence from the *oriT* of pMK101 (*oriT*22-F), a fluorescent 35-mer of the pre-*nic* DNA containing the hairpin structure (F-*oriT*35) and a fluorescent full-length 57-mer with both the pre- and post-*nic* sequence (F-*oriT*57), as well as random 57- and 35-mers (Fig. 1B & Table S1).

We found that the TraI trans-esterase domain binds both F-*oriT*57 and F-*oriT*35 DNA with high affinity (apparent K_D is below 200 nM), while this domain only interacts with post-*nic* or random DNA at the protein concentrations above 800 or 1600 nM (Fig. 5A & S3A), indicating that the trans-esterase domain binds sequence specific to the pre-*nic* region. In contrast, the TraI helicase domain shows only low affinity binding to both *oriT* and random DNA, showing a significant shift only at 400 nM protein and not reaching 100 % shifted DNA at the highest concentration 1600 nM. This domain thus appears to have no sequence specificity but seems

to have a stronger interaction with the longer DNA constructs (Fig. 5B & S3B). As full-length TraI includes both the helicase and the trans-esterase domain, we expected to find a combination of their DNA binding capabilities. TraI indeed binds both *oriT* and random DNA, but with an intriguing difference; the random DNA, which mainly the helicase domain binds, shows a significant mobility shift with protein concentrations of 200 nM and above (Fig. 5C). This indicates a significant increase in the apparent affinity compared to the helicase domain in isolation. Both F-*oriT*57 and F-*oriT*35, show a complete mobility shift with protein concentrations of 200 nM and above. At higher concentrations additional supershifts are also observed. At lower protein concentrations part of the *oriT* DNA remains stuck in the well. This complex can only be observed up to 400 nM TraI and disappears as the protein concentration is increased further. This result has been consistently observed in mobility shift assays with *oriT* DNA, but not with random DNA (Fig. 5C & S3C).

Finally, we examined the capacity of TraI to nick and religate its substrate *oriT* DNA. TraI was incubated with the fluorescent full-length F-*oriT*57, containing the recognition hairpin, the *nic*-site and a post-*nic* stretch of 22 bases, or a random DNA sequence of the same length. DNA products were consequently separated on a denaturing polyacrylamide gel. Our results demonstrate that both TraI and TraI-TE can nick the tested *oriT* DNA substrate to a similar degree (about 50% of the DNA, Fig. 6A & S4A), while they are unable to nick a random DNA control (Fig. S4C & S4D). We did not detect a higher degree of cleaved F-*oriT*57 DNA, even at higher protein concentrations. We predicted that this was due to the religation ability of TraI. To test this, we added a shorter version of the fluorescent pre-*nic* DNA (F-*oriT*20) in addition to F-*oriT*57 to the incubation mixture of TraI. In this experiment F-*oriT*57 is expected to be (partly) cleaved by TraI into a fluorescent 35-mer and a non-fluorescent covalently bound 22-mer. If religation occurs, the 22-mer would be religated either to the fluorescent 35-mer to re-form the original F-*oriT*57 or the F-*oriT*20 to form a 42-mer. This 42-mer was indeed seen on the denaturing gel (Fig. 6B). A second way of testing for nicking activity was also used. F-*oriT*35 and non-fluorescent *oriT*57 were incubated with TraI. The appearance of F-*oriT*57 shows that non-fluorescent *oriT*57 was nicked and recombined with F-*oriT*35 (Fig. 6B).

Discussion

In this study, we structurally and biochemically characterized the relaxase TraI from the MOB_F family plasmid pKM101 to better understand the diversity of a protein family of key importance to type IV secretion. We solved the structure of the trans-esterase domain of TraI with and without substrate DNA. In contrast to findings reported for the identical domain from the pCU1-encoded relaxase ($TraI_{pCU1}$)[21], our structure highlights the conserved active site of these trans-esterases and reveals the importance of the flexible thumb for DNA binding. We also observed tetramerization of the full-length protein, which offers important insights into the variability of relaxase quaternary structures as a whole.

The DNA bound structure of the $TraI_{pKM101}$ trans-esterase domain (Fig. 2A) shows a protein fold that can be described as a palm, at the center of which is the active site and a thumb domain that is involved in holding the substrate DNA in place. This structure is similar to the previously solved trans-esterase domains of its homologs[10,11,20]. The highest structural homology is found with $TrwC_{R388}$, which is also a member of the MOB_{F11} subclade (PDB: 2CDM, 50 % identity, root mean square deviation (RMSD) = 1.7 Å over 276 residues), but the structure is also similar to that of $TraI_F$, a MOB_{F12} relaxase, (PDB: 2A0I, 37 % identity, RMSD = 2.3 Å over 269 residues)[22]. The DNA binding site is highly conserved between these proteins and the bound ssDNA makes a U-turn that is characteristic for *oriT* specific DNA binding[23].

The electron density of the ion, coordinated by the histidine triad at the active site can best be modeled by a manganese ion in both structures. This Mn^{2+} is likely retained during purification since no metals were added at any point during purification or crystallization of the DNA-bound structure. Our activity assays were performed with magnesium, and no other divalent metals were tested as TraI homologs were shown to be able to actively nick in the presence of various metals[11,21].

While the DNA bound structures of the MOB_F family relaxases are very similar, there is more variability between the apo structures, which probably reflects the greater flexibility of the structures in the absence of ligand. The apo structure of $TraI_{pKM101}$ presented here in Fig. 3A shows the thumb domain in an open conformation. Previously, this thumb domain has only been observed in crystal structures with bound DNA[10,11,20]. The RMSD between our apo and DNA-bound structure is 3.3 Å over 277 residues, which is indicative of the large molecular

movement of up to 40 Å visualized in Fig. 3A. It is likely that the open conformation is important for *oriT* DNA to access and bind to the active site.

An additional apo trans-esterase domain structure of interest is the TraI-TE of the plasmid pCU1. pCU1 is an antibiotic resistance plasmid that shares a common ancestor with pKM101, R46[9]. pKM101 and pCU1 have an identical *oriT* region and identical relaxases. An apo structure of this $TraI_{pCU1}$ (PDB: 3L6T) shows large differences to both structures of $TraI_{pKM101}$ despite of its 100 % sequence identity (Fig. 3B). The overall fold is similar, with a RMSD of 1.4 Å over 208 residues, but the entire C-terminal region of $TraI_{pCU1}$ is missing including the thumb domain, and the helix at the base of the thumb (α-helices 7-10 of the $TraI_{pKM101}$ DNA bound structure). Additionally, α-helix 1, which contains the beginning of the conserved YY-$X_{(5-6)}$-YY motive that provides the catalytic tyrosine, mostly appears as a loop and points far away from the active site, resulting in a >10 Å displacement compared to the positions in the homologs[21]. Possibly, the displacement of α-helix 1 is related to the destabilization of the thumb domain of the protein. In contrast, both structures presented here, as well as other relaxase structures ($TrwC_{R388}$ PDB: 1S6M, $TraI_F$ PDB: 2Q7T), show various interactions between the base of the thumb domain (α-helix 10) and the loop region following α-helix 1 (R274 to D25 in $TraI_{pKM101}$) (Fig. S1). In the DNA bound structure of $TraI_{pKM101}$ an additional interaction was found between H278 (α-helix 10) and Y26 (the third Y of the YY-$X_{(5-6)}$-YY motif), which might further contribute to this stabilization. It is also noteworthy that TraI-TE_{pCU1} was reported to bind DNA in a weak and sequence independent manner, with K_D values for binding to *oriT* and non-*oriT* oligomers ranging from 0.7 to 1.2 μM[21]. These observations are not in line with our data nor with that of other relaxases showing that relaxases bind *oriT* DNA in a sequence specific manner, with high affinity[20,23,24]. While the mobility shift assays presented here (Fig. 5 & S3) are not suitable for precise K_D determinations, the apparent K_D of TraI-TE_{pKM101} for *oriT* DNA is <200 nM, and interactions with random DNA were observed only at much higher apparent K_D values (protein concentrations above 800 nM, Fig. 5A). Our data clearly indicate that the trans-esterase domain of $TraI_{pKM101}$ has a high affinity and specificity for its cognate *oriT* DNA. We conclude that $TraI_{pKM101}$ - and identical $TraI_{pCU1}$ - thus behave like other characterized relaxases.

Additional DNA binding experiments supplied further evidence that TraI functions like its homologs. The TraI helicase domain (TraI-H) bound equally well to both *oriT* and random

DNA, consistent with expectations that the helicase acts to unwind the entire plasmid. Full-length TraI, with both the helicase and trans-esterase domains, also bound both *oriT* and random DNA. Finally, our nicking and religation assays showed that TraI as well as TraI-TE are equally capable of nicking and religating the *oriT* DNA, but not random DNA, confirming that the trans-esterase is active and *oriT* specific.

Although the trans esterase domain of $TraI_{pKM101}$ structurally and functionally resembles other MOB_F homologs, we also gained evidence that $TraI_{pKM101}$ uniquely oligomerizes as tetramers. This was first indicated by size exclusion analysis of purified TraI, and further confirmed by SEC-MALS showing that the protein in the first elution peak has a molecular weight corresponding to a tetramer. The tetramer appears to be relatively stable, but in an equilibrium with the monomer, as deduced from our findings that TraI tetramer fractions predominantly remained tetrameric after re-running the same sample over a second SEC experiment. On native gels, SEC fractions also invariably migrated at a position corresponding to a tetramer. Finally, we found that TraI induces multiple supershifts in our EMSAs, presumptively through binding of distinct oligomeric states including the tetrameric form to substrate DNA (Fig. 5C). Thus, while $TraI_{pKM101}$ was previously reported to self-interact by use of a two-hybrid screen[27], our present findings strongly favor a proposal that $TraI_{pKM101}$ functions as a tetramer. Both TraI-TE and TraI-H domains individually are monomeric in solution (Fig. S2), with only the full-length $TraI_{pKM101}$ forming tetramers. We therefore hypothesize that the CTD of $TraI_{pKM101}$ is involved in tetramerization, which is in line with prior findings that a part of the CTD domain of $TraI_F$ tetramerizes upon crystallization[25].

The functional role of tetramerization of $TraI_{pKM101}$ remains to be explored, but there are several ways in which it could be important. For $TraI_{pKM101}$s closest relative, $TrwC_{R388}$, it has been proposed that the second nicking reaction, necessary after rolling-circle replication, is mediated by the third tyrosine (Y26) of the conserved YY-X(5-6)-YY motif in the trans-esterase domain[26]. However, there is another model, based on studies of the relaxase from the F-plasmid, in which the second nicking is performed by another relaxase molecule[27]. While it is characteristic for the MOB_F family relaxases to have this second pair of conserved tyrosines, other classes of relaxases only have a single catalytic tyrosine and have to use a different mechanism for this second reaction[5]. The oligomerization that we observed for TraI, could position a second TraI molecule for the termination nicking reaction. But why a tetramer

instead of a dimer? Here we can only speculate. It is however likely that the answer lies not with the nicking function of the relaxase, considering that both nicking and religation reactions were found to be performed at similar rates by the monomeric trans-esterase domain and the full-length protein, which was found to be in a monomer/tetramer equilibrium. It is also possible that instead of needing several trans-esterase domains, it is more advantageous to have multiple copies of the TraI-helicase, since $TraI_{pKM101}$ does not have a 2nd helicase domain like $TraI_F$. The extra vestigial helicase domain in $TraI_F$ has been suggested to function like a single stranded DNA binding protein (SSB)[28], which is a role that potentially could be performed by a helicase domain on another subunit of the tetrameric TraI. It is also possible that tetramerization increases the processivity of the helicase which is considered high for MOB_F relaxases[29,30]. Alternatively, tetramerization might be important for interaction with other components of the relaxosome or the T4SS.

To conclude, we have shown that $TraI_{pKM101}$ has a high affinity and specificity to its cognate *oriT*, in contrast to previous findings for the identical trans-esterase domain of pCU1. We have visualized the large movement that the thumb domain undergoes upon DNA binding in the trans-esterase domain. We have shown that $TraI_{pKM101}$ can both nick and religate *oriT* in a similar manner as other MOB_F family relaxases. Finally, in striking contrast to other characterized relaxases, $TraI_{pKM101}$ is a tetramer in solution and appears to function as a tetramer as demonstrated by our EMSAs. Exploring the biological functions of tetrameric TraI *in vivo* remains an exciting challenge for future research.

Acknowledgments

A pET15b plasmid containing the pKM101 TraI gene that we used for further cloning in *E. coli* was generously donated by Prof. Peter J Christie, whom we also thank for fruitful discussions regarding the project. We acknowledge MAX IV Laboratory for time on Beamline BioMax under Proposal 20180236. Research conducted at MAX IV, a Swedish national user facility, is supported by the Swedish Research council under contract 2018-07152, the Swedish Governmental Agency for Innovation Systems under contract 2018-04969, and Formas under contract 2019-02496. We also acknowledge the synchrotrons Swiss Light Source (Paul Scherrer Institute, Switzerland) for time at beamline PX1 and the ESRF (France) for time at beamlines ID23 and ID30. This work was supported by grants from the Swedish Research Council (2016-03599), Knut and Alice Wallenberg Foundation and Kempestiftelserna (SMK-1762 & SMK-1869) to R.P-A.B.

CRediT statement

Annika Breidenstein: Conceptualization, Investigation, Writing - Original Draft, Writing – Revision. Josy ter Beek: Conceptualization, Writing - Original Draft, Writing – Revision, Supervision. Ronnie Berntsson: Conceptualization, Writing - Original Draft, Writing – Revision, Supervision, Funding acquisition.

References

1. Khara, P., Song, L., Christie, P. J. & Hu, B. *In Situ* Visualization of the pKM101-Encoded Type IV Secretion System Reveals a Highly Symmetric ATPase Energy Center. *mBio* e02465-21 (2021) doi:10.1128/mBio.02465-21.
2. Hu, B., Khara, P. & Christie, P. J. Structural bases for F plasmid conjugation and F pilus biogenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **116**, 14222–14227 (2019).
3. Chetrit, D., Hu, B., Christie, P. J., Roy, C. R. & Liu, J. A unique cytoplasmic ATPase complex defines the Legionella pneumophila type IV secretion channel. *Nat. Microbiol.* **3**, 678–686 (2018).
4. Zechner, E. L., Moncalián, G. & de la Cruz, F. Relaxases and Plasmid Transfer in Gram-Negative Bacteria. in *Type IV Secretion in Gram-Negative and Gram-Positive Bacteria* (eds. Backert, S. & Grohmann, E.) vol. 413 93–113 (Springer International Publishing, 2017).
5. Garcillán-Barcia, M. P., Francia, M. V. & de La Cruz, F. The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* **33**, 657–687 (2009).
6. Byrd, D. R. & Matson, S. W. Nicking by transesterification: the reaction catalysed by a relaxase. *Mol. Microbiol.* **25**, 1011–1022 (1997).
7. Alvarez-Martinez, C. E. & Christie, P. J. Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiol. Mol. Biol. Rev.* **73**, 775–808 (2009).
8. Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F. & Rocha, E. P. C. The Repertoire of ICE in Prokaryotes Underscores the Unity, Diversity, and Ubiquity of Conjugation. *PLoS Genet.* **7**, e1002222 (2011).
9. Paterson, E. S. *et al.* Genetic Analysis of the Mobilization and Leading Regions of the IncN plasmids pKM101 and pCU1. *J. Bacteriol.* **181**, 2572 (1999).

10. Larkin, C. *et al.* Inter- and Intramolecular Determinants of the Specificity of Single-Stranded DNA Binding and Cleavage by the F Factor Relaxase. *Structure* **13**, 1533–1544 (2005).
11. Boer, R. *et al.* 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 (2006).
12. Ilangovan, A. *et al.* Cryo-EM Structure of a Relaxase Reveals the Molecular Basis of DNA Unwinding during Bacterial Conjugation. *Cell* **169**, 708-721.e12 (2017).
13. Geertsma, E. R. & Dutzler, R. A Versatile and Efficient High-Throughput Cloning Tool for Structural Biology. *Biochemistry* **50**, 3272–3278 (2011).
14. Kabsch, W. *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
15. Monaco, S. *et al.* Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. *J. Appl. Crystallogr.* **46**, 804–810 (2013).
16. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
17. Emsley, P. & Cowtan, K. *Coot* : model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
18. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
19. Some, D., Amartely, H., Tsadok, A. & Lebendiker, M. Characterization of Proteins by Size-Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *J. Vis. Exp.* 59615 (2019) doi:10.3791/59615.
20. Guasch, A. *et al.* Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat. Struct. Mol. Biol.* **10**, 1002–1010 (2003).

21. Nash, R. P., Habibi, S., Cheng, Y., Lujan, S. A. & Redinbo, M. R. The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res.* **38**, 5929–5943 (2010).

22. Holm, L. Using Dali for Protein Structure Comparison. in *Structural Bioinformatics* (ed. Gáspári, Z.) vol. 2112 29–42 (Springer US, 2020).

23. Carballeira, J. D., González-Pérez, B., Moncalián, G. & la Cruz, F. de. A high security double lock and key mechanism in HUH relaxases controls oriT-processing for plasmid conjugation. *Nucleic Acids Res.* **42**, 10632–10643 (2014).

24. Stern, J. C. & Schildbach, J. F. DNA Recognition by F Factor TraI36: Highly Sequence-Specific Binding of Single-Stranded DNA. *Biochemistry* **40**, 11586–11595 (2001).

25. Guogas, L. M., Kennedy, S. A., Lee, J.-H. & Redinbo, M. R. A Novel Fold in the TraI Relaxase–Helicase C-Terminal Domain Is Essential for Conjugative DNA Transfer. *J. Mol. Biol.* **386**, 554–568 (2009).

26. Grandoso, G. *et al.* 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 (2000).

27. Dostál, L., Shao, S. & Schildbach, J. F. Tracking F plasmid TraI relaxase processing reactions provides insight into F plasmid transfer. *Nucleic Acids Res.* **39**, 2658–2670 (2011).

28. Dostál, L. & Schildbach, J. F. Single-Stranded DNA Binding by F TraI Relaxase and Helicase Domains Is Coordinately Regulated. *J. Bacteriol.* **192**, 3620–3628 (2010).

29. Sikora, B., Eoff, R. L., Matson, S. W. & Raney, K. D. DNA Unwinding by Escherichia coli DNA Helicase I (TraI) Provides Evidence for a Processive Monomeric Molecular Motor. *J. Biol. Chem.* **281**, 36110–36116 (2006).

30. Lahue, E. E. & Matson, S. W. Escherichia coli DNA helicase I catalyzes a unidirectional and highly processive unwinding reaction. *J. Biol. Chem.* **263**, 3208–3215 (1988).

Figures

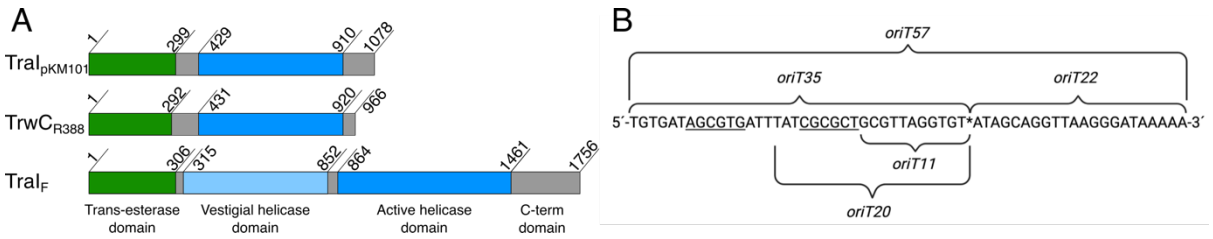

Figure 1. Schematic overviews of domain organization of relaxases and the cognate *oriT* of pKM101. (A) Domain structure of three MOB_F family relaxases $TraI_{pKM101}$, $TrwC_{R388}$ and $TraI_F$[4,12]. Trans-esterase domains are shown in green, vestigial helicase domain in light blue, active helicase domains in dark blue and the C-terminal domains and linker regions are shown in grey. (B) *oriT* sequence of pKM101 and an overview of the different *oriT* nucleotides used in this study. * indicates the *nic*-site, underlined regions of the sequence indicate the bases participating in hairpin formation.

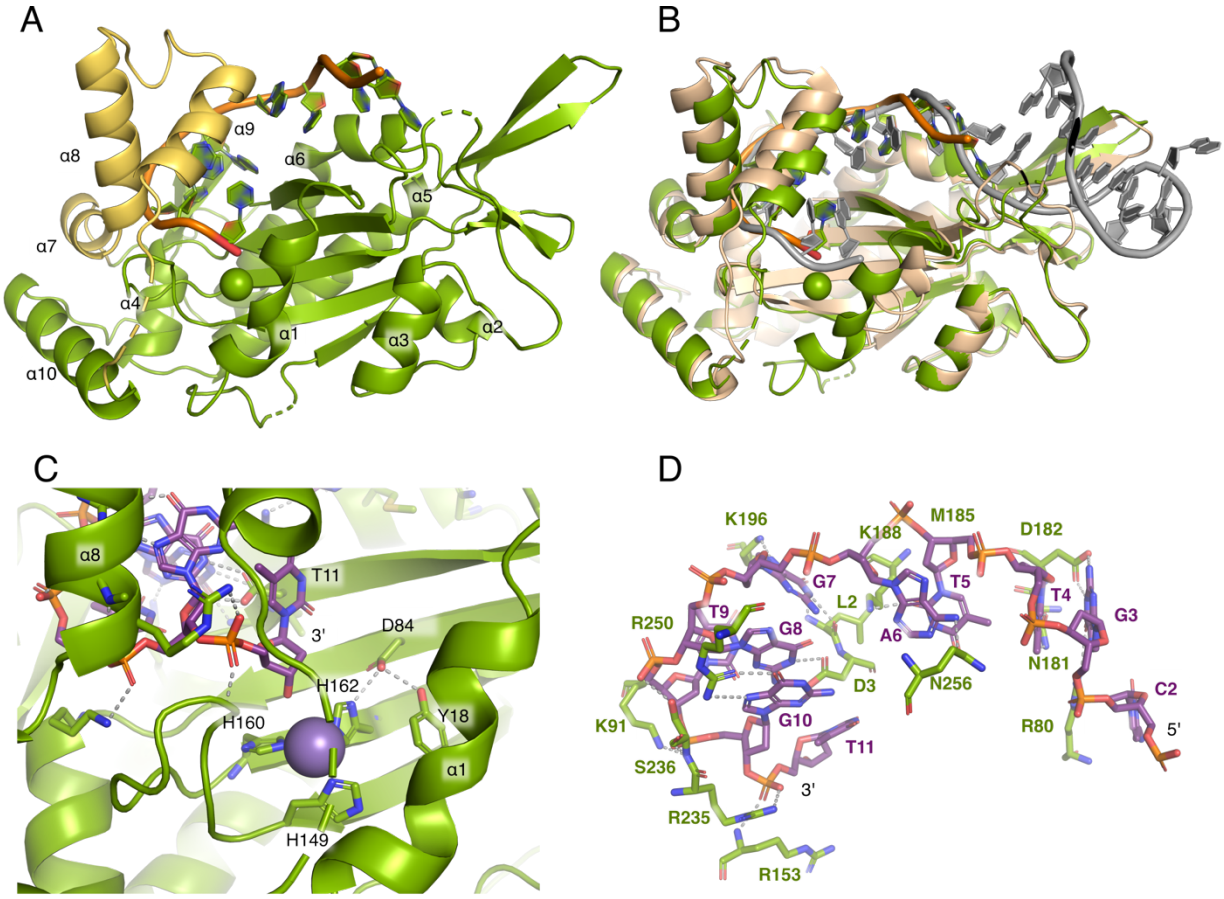

Figure 2. Crystal structure of DNA-bound trans-esterase domain of $TraI_{pKM101}$ (green). (A) Overview with thumb domain shown in yellow. (B) Superimposition of $TraI_{pKM101}$ and $TrwC_{R388}$ (protein shown in wheat, DNA in grey, PDB: 2CDM). (C) Active site of TraI-TE_{pKM101} consisting of the histidine triad coordinating a Mn^{2+} (H149, H160 and H162), conserved aspartate (D84) and catalytic tyrosine (Y18). (D) The bound *oriT* DNA (purple) and its interactions with TraI-TE. Residues that are forming H-bonds with DNA are represented as sticks.

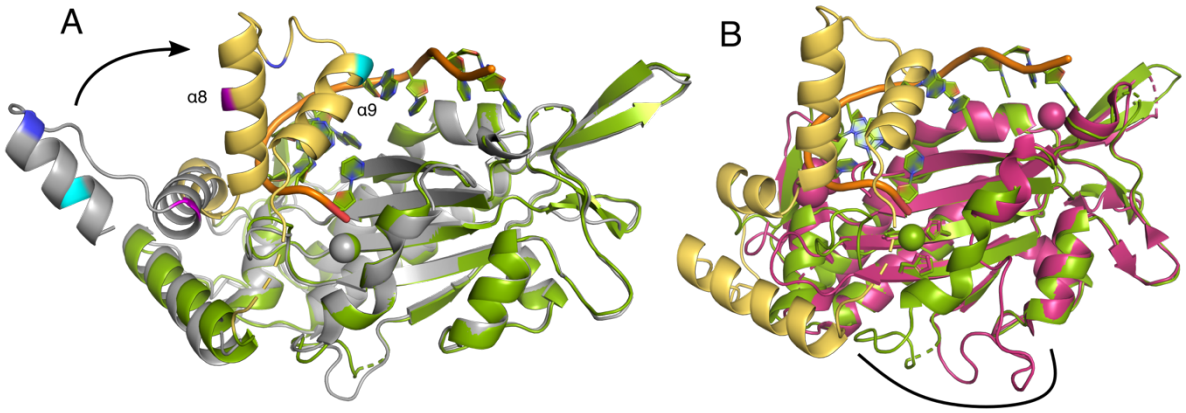

Figure 3. Apo structures of TraI. (A) DNA bound structure of $TraI_{pKM101}$ (green) with the thumb domain in yellow, superimposed with the apo structure (grey). To provide a visual aid in understanding the conformational changes that occur in the thumb domain upon DNA binding, residue K241 (pink), R251 (blue) and Q256 (cyan) are highlighted in both structures. (B) DNA bound structure of $TraI_{pKM101}$ (green) with thumb domain and C-terminus in yellow, superimposed with $TraI_{pCU1}$ (red, PDB: 3L6T). The black line highlights the different conformation of α-helix 1, which forms part of the active site, in the two different structures.

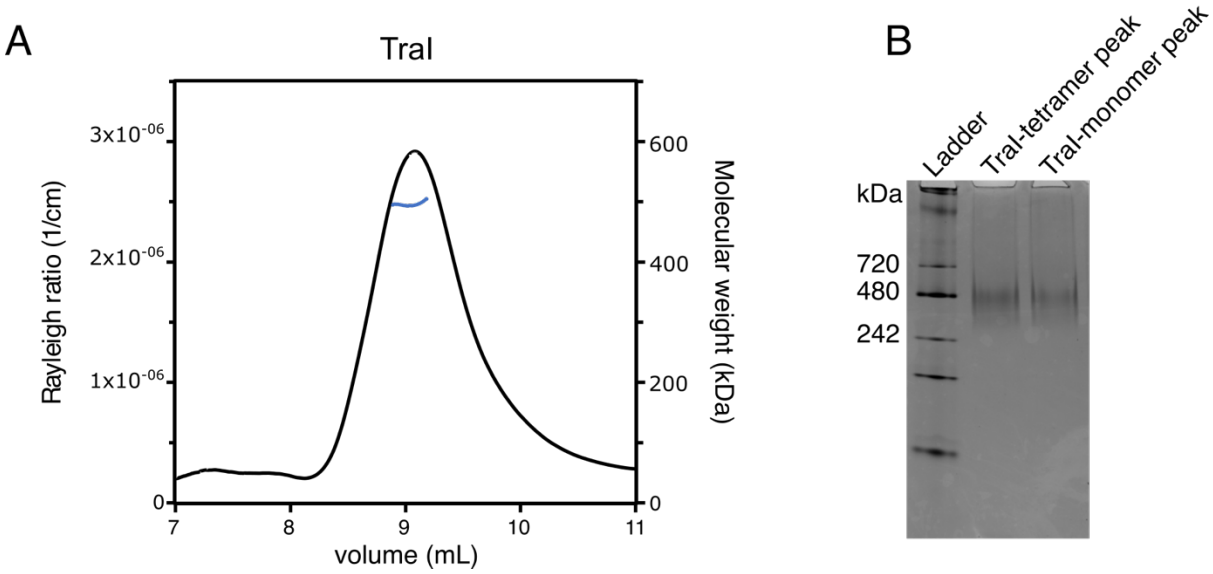

Figure 4. Oligomerization of TraI. (A) Size-exclusion chromatography coupled to multiangle light scattering (SEC-MALS) of TraI run over a Superdex 200 10/300 GL Increase column. The chromatogram shows the Rayleigh ratio which is directly proportional to the intensity of the scattered light in excess to the solvent (left axis). The blue line indicates the calculated molecular weight of the protein throughout the peak (right axis). The main elution peak corresponds to a tetramer (485 +/- 10 kDa measured vs the expected 492 kDa). (B) Native gel of samples from tetrameric and monomeric SEC fractions.

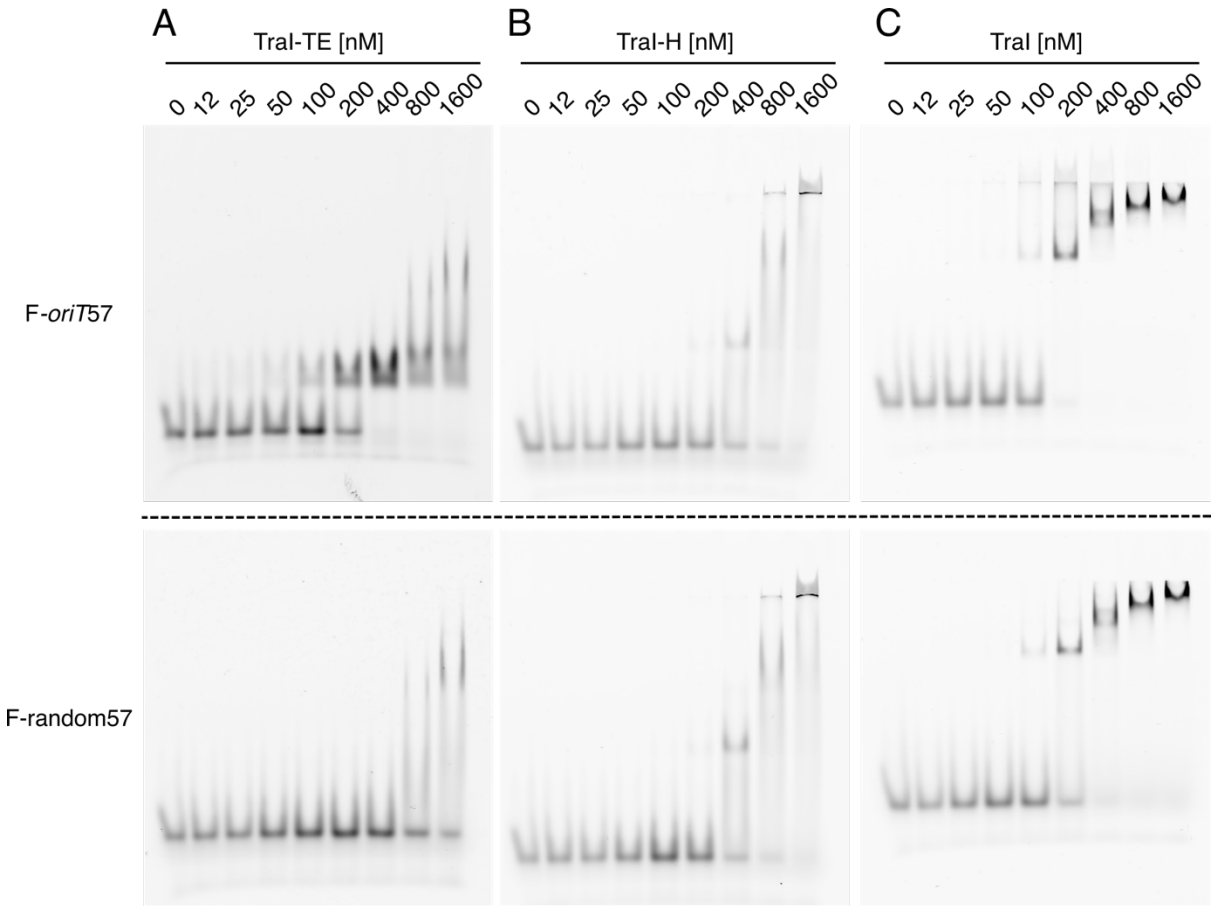

Figure 5. Electromobility shift assay using 50 nM fluorescent 57-mer *oriT* (F-*oriT*57) (upper panel) or random DNA (lower panel) of the trans-esterase domain (TraI-TE) (A), helicase domain (TraI-H) (B) or full-length TraI (C).

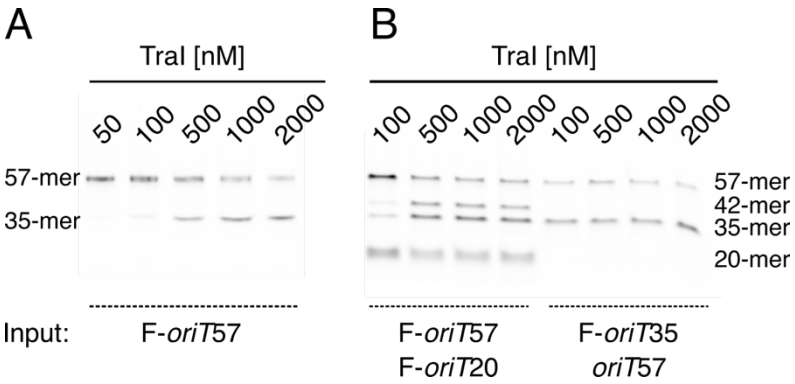

Figure 6. TraI activity assays. (A) Nicking assay. F-*oriT*57 was incubated with increasing protein concentrations and the nicked DNA was separated on a denaturing gel, showing the emergence of a 35-mer nicking product. (B) Religation assay. F-*oriT*57 and F-*oriT*20 were incubated with increasing protein concentrations (left). The appearance of a 35-mer product indicates nicking of F-*oriT*57 and the appearance of a 42-mer product indicates religation of the covalently bound post-nick DNA with the shorter F-*oriT*20 pre-nick DNA. F-*oriT*35 and *oriT*57 were incubated with increasing protein concentrations (right). The appearance of a 57-mer product shows nicking of the non-fluorescent *oriT*57 and religation with the fluorescent F-*oriT*35.

Supplementary figures

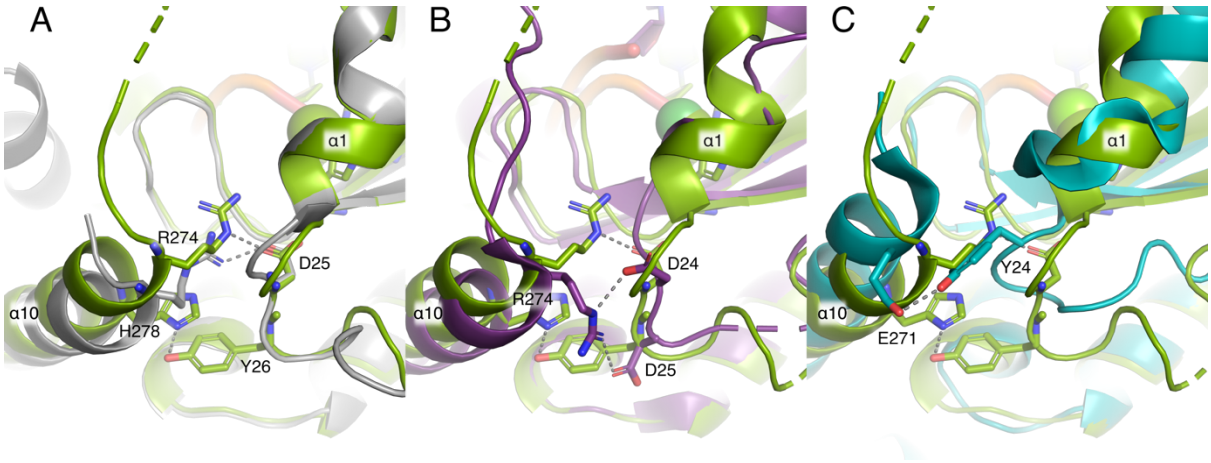

Figure S1. Interaction between the base of the thumb domain and loop after α-helix 1. (A) $TraI_{pKM101}$ DNA bound structure (green) superimposed with apo structure (grey). H-bonds are formed between R274 and D25 in both structures with an additional bond between H278 and Y26 in the DNA bound structure. (B) $TraI_{pKM101}$ DNA bound structure (green) superimposed with $TrwC_{R388}$ (PDB: 1S6M, purple). H-bonds between R274 and D24 or D25. (C) $TraI_{pKM101}$ DNA bound structure (green) superimposed with $TraI_{F}$ (PDB: 2Q7T, turquoise). H-bonds between E271 and Y24.

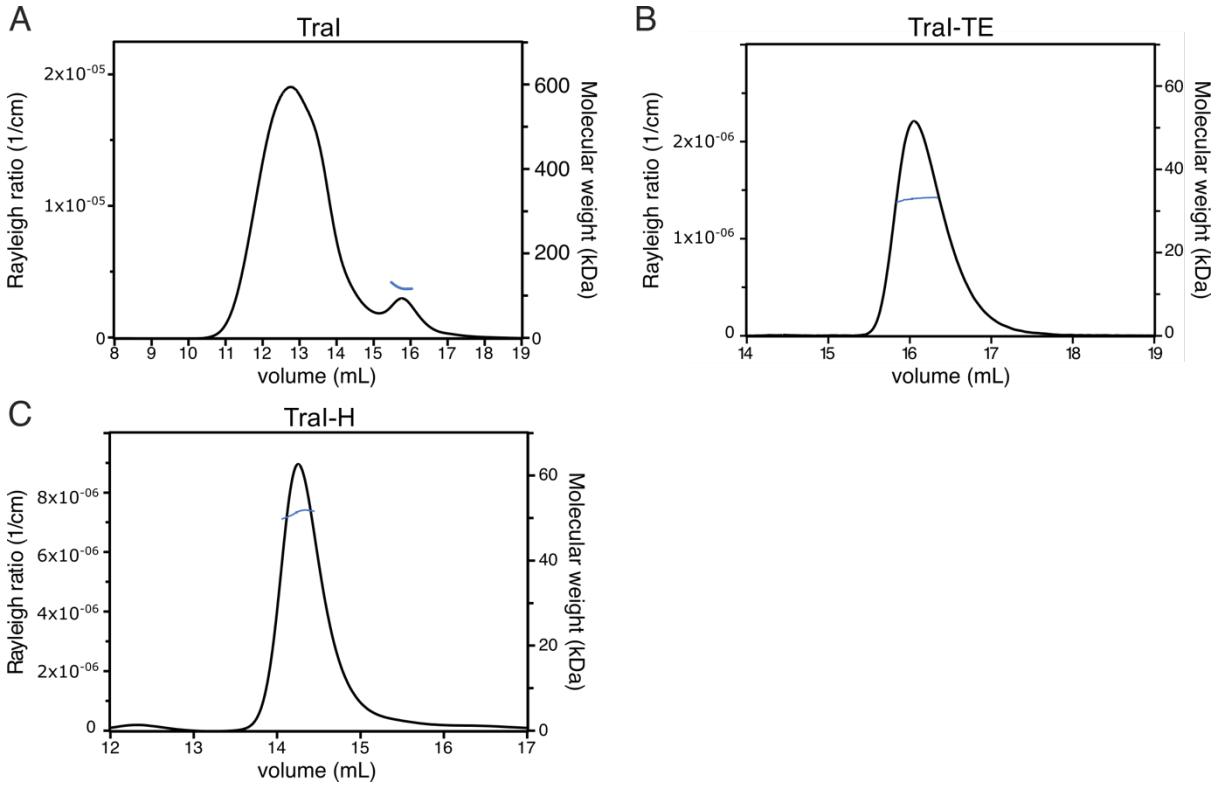

Figure S2. Size-exclusion chromatography coupled to multiangle light scattering (SEC-MALS) of TraI. The chromatograms show the Rayleigh ratio which is directly proportional to the intensity of the scattered light in excess to the solvent (left axis). The blue lines indicate the calculated molecular weight of the protein throughout the peak (right axis). (A) TraI elution profile using a Superose 6 column, showing both the tetrameric and monomeric peak with the latter corresponding to 123 +/- 5 kDa. (B) SEC-MALS using a Superdex200 column of the trans-esterase domain corresponding to a monomer with 33 +/- 1 kDa. (C) SEC-MALS, using a Superdex200 column, of the helicase domain corresponding to a monomer with 49 +/- 2 kDa.

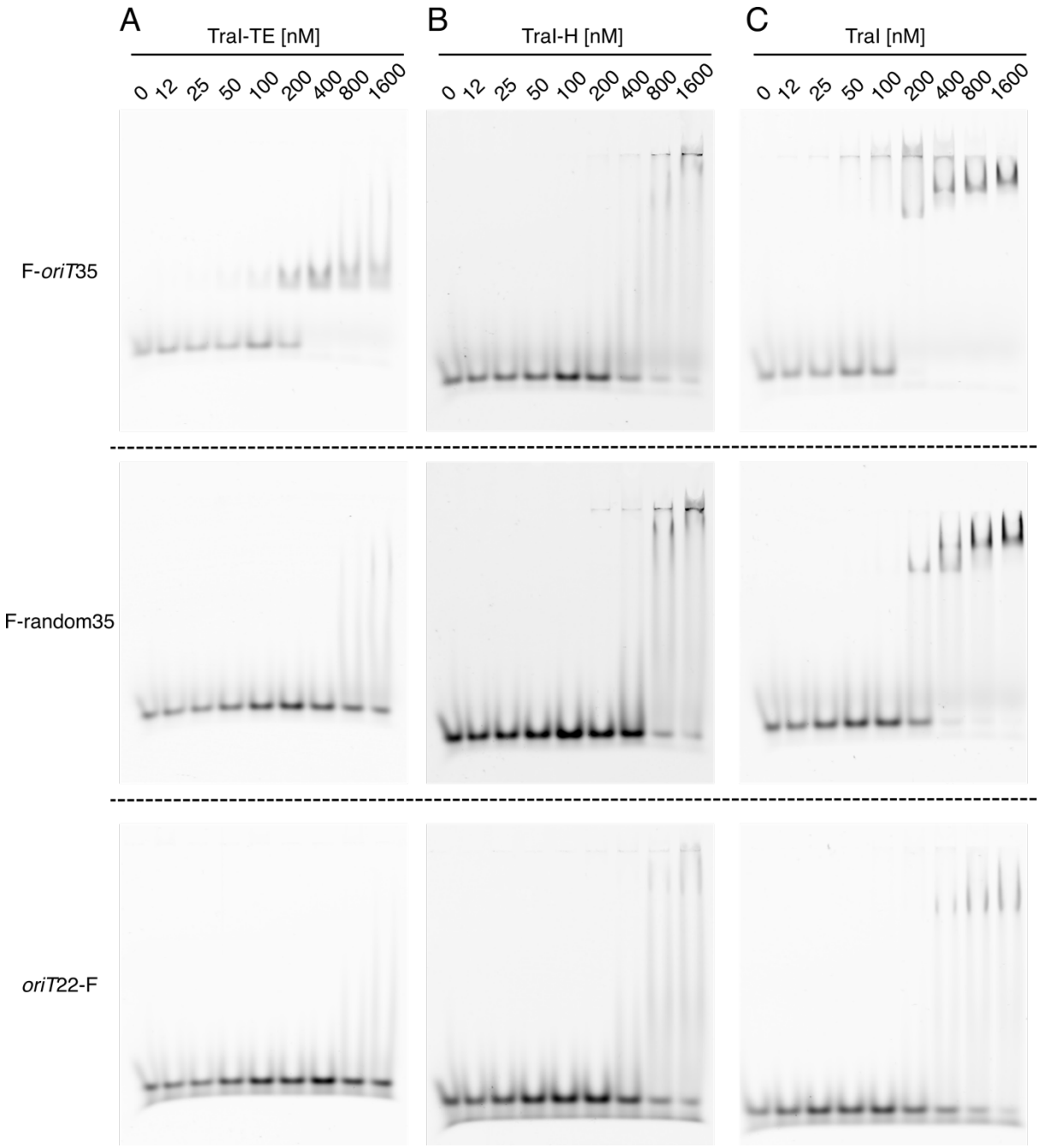

Figure S3. Electromobility shift assay using 50 nM F-*oriT*35 (upper panel), F-random35 (middle panel) or *oriT*22-F (lower panel) of the trans-esterase domain (TraI-TE) (A), helicase domain (TraI-H) (B) or full-length TraI (C).

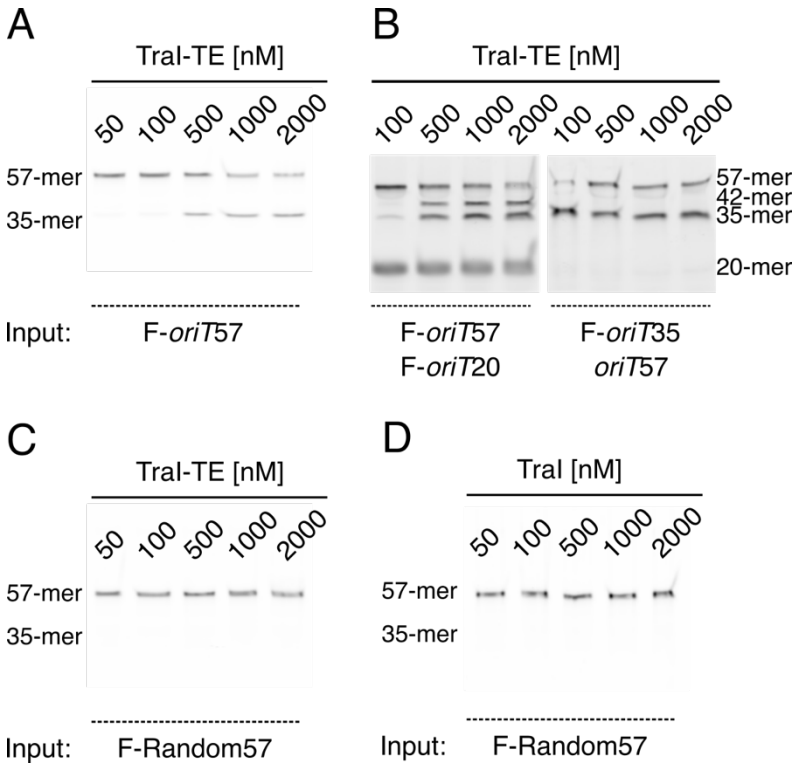

Figure S4. Activity assays of TraI-trans esterase domain (TraI-TE). (A) Nicking assay. F-*oriT*57 was incubated with increasing protein concentrations and the nicked DNA was separated on a denaturing gel, showing the emergence of a 35-mer nicking product. (B) Religation assay. F-*oriT*57 and F-*oriT*20 were incubated with increasing protein concentrations (left). The appearance of a 35-mer product indicates nicking of F-*oriT*57 and the appearance of a 42-mer product indicates religation of the covalently bound post-nick DNA with the shorter F-*oriT*20 pre-nick DNA. F-*oriT*35 and *oriT*57 were incubated with increasing protein concentrations (right). The appearance of a 57-mer product shows nicking of the non-fluorescent *oriT*57 and religation with the fluorescent F-*oriT*35. (C) Nicking assay with TraI-TE and F-random57 as input. (D) Nicking assay with TraI and F-random57 as input.

Supplementary tables

Table S1. List of strains and plasmids used for constructs design.

Strain, plasmid, or oligonucleotide	**Relevant features or sequences (5'-3')**	**References**
Strains (*E. coli*):		
TOP10 One Shot	Cloning host	Thermo-Fisher
BL21 (DE3)	Expression host	Thermo-Fisher
Plasmids:		
p7XC3H-TraI	P7XC3H expressing *traI-his*$_{10}$	This study
p7XNH3-TraI-H	P7XNH3 expressing *his*$_{10}$-*traI*$_{429-910}$	This study
p7XC3H-TraI-TE	P7XC3H expressing *traI*$_{1-299}$-*his*$_{10}$	This study
FX-cloning primers		
Fw-TraI	ATATATGCTCTTCTAGTCTTGATATAACCACGATTACCCGCCAG	Eurofins
Rv-TraI	TATATAGCTCTTCATGCGATTTCATGGCCCCCTTCTTCATGCTC	Eurofins
Rv-TraI-TE	TATATAGCTCTTCATGCTCCCTGCCATTCACGGTTATCAAAATC	Eurofins
Fw-TraI-H	ATATATGCTCTTCTAGTCTTAAAAAGACCAGCCACCGCGTCACT	Eurofins
Rv-TraI-H	TATATAGCTCTTCATGCGCTCACGCTCGATGCCAAAGATTTTTT	Eurofins
***ss*DNA oligomers**		
oriT57	TGTGATAGCGTGATTTATCGCGCTGCGTTAGGTGTATAGCAGGT TAAGGGATAAAAA	Eurofins
F-*oriT*57	**FITC**-TGTGATAGCGTGATTTATCGCGCTGCGTTAGGTGTATAGCAGGT TAAGGGATAAAAA	Eurofins
F-*oriT*35	**FITC**-TGTGATAGCGTGATTTATCGCGCTGCGTTAGGTGT	Eurofins
*F-oriT*20	**FITC**-TATCGCGCTGCGTTAGGTGT	Eurofins
*oriT*11	GCGTTAGGTGT	Eurofins
*oriT*22-F	ATAGCAGGTTAAGGGATAAAAA-**FITC**	Eurofins
F-Random57	**FITC**-TCCGCCATGCAGACGAGACCAGTCGGAGATTACCGAGCATTCTA TCAGGTCGGCGAC	Eurofins
F-Random35	**FITC**-CACTAGTGAGCTACTGGAGCCGAGGGGTAACCACG	Eurofins

Table S2. Data collection and refinement statistics.

Data collection summary	**TraI DNA bound**	**TraI apo**
Resolution range	39.67 - 2.1 (2.175 - 2.1)	45.29 - 1.7 (1.761 - 1.7)
Space group	P 21 21 21	P 21 21 21
Cell dimensions		
a, b, c (Å)	43.444 81.643 90.785	40.166 80.628 90.575
α, β, γ (°)	90 90 90	90 90 90
Total reflections	38982 (3816)	66280 (6506)
Unique reflections	19496 (1908)	33141 (3253)
Multiplicity	2.0 (2.0)	2.0 (2.0)
Completeness (%)	99.85 (99.79)	99.91 (99.79)
Mean I/sigma (I)	9.16 (1.15)	12.42 (2.18)
R-meas	0.07283 (0.6719)	0.03658 (0.4699)
CC(1/2)*	0.996 (0.751)	0.999 (0.782)
Refinement summary		
R-work	0.2255 (0.3359)	0.1709 (0.2550)
R-free	0.2695 (0.3514)	0.1913 (0.2818)
Number of non-hydrogen atoms	2648	2594
protein	2403	2376
DNA	208	
other ligands	1	42
solvent	36	198
RMS(bonds)	0.003	0.010
RMS(angles)	0.46	1.02
Ramachandran favored (%)	97.61	98.28
Ramachandran allowed (%)	2.39	1.37
Ramachandran outliers (%)	0.00	0.34
Average B-factor	60.80	31.54
protein	60.50	30.73
DNA	65.97	
other ligands	39.89	49.90
solvent	51.46	39.37

Statistics for the highest-resolution shell are shown in parentheses.