8136–8153 *Nucleic Acids Research, 2019, Vol. 47, No. 15* *Published online 5 July 2019*
doi: 10.1093/nar/gkz577

DNA processing by the MOB_H family relaxase TraI encoded within the gonococcal genetic island

Jan-Hendrik Heilers[1], Jens Reiners[2,3], Eva-Maria Heller, Annika Golzer[1], Sander H.J. Smits[2,3] and Chris van der Does[1,*]

[1]Institute for Biology II, Microbiology, Albert Ludwig University Freiburg, 79104 Freiburg, Germany, [2]Biochemie I, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany and [3]Center for Structural Studies, Heinrich Heine University, 40225 Düsseldorf, Germany

Received December 22, 2018; Revised June 18, 2019; Editorial Decision June 19, 2019; Accepted June 26, 2019

ABSTRACT

Relaxases of the MOB_H family are often found on large plasmids, genetic islands and integrative conjugative elements. Many members of this family contain an N-terminal relaxase domain (TraI_2) followed by a disordered middle part and a C-terminal domain of unknown function (TraI_2_C). The TraI_2 domain contains two putative metal-binding motifs, an HD domain motif and an alternative 3H motif. TraI, encoded within the gonococcal genetic island of *Neisseria gonorrhoeae*, is the prototype of the MOB_H family. SAXS experiments showed that TraI_2 and TraI_2_C form globular structures separated by an extended middle domain. The TraI_2 domain cleaves *oriT*-ssDNA in a site-specific Mn^{2+} or Co^{2+} dependent manner. The minimal *oriT* encompasses 50 nucleotides, requires an inverted repeat 3′ of the *nic*-site and several nucleotides around *nic* for efficient cleavage. Surprisingly, no stable covalent relaxase-DNA intermediate was observed. Mutagenesis of conserved tyrosines showed that cleavage was abolished in the Y212A mutant, whereas the Y212F and Y212H mutants retained residual activity. The HD and the alternative 3H motifs were essential for cleavage and the HD domain residues D162 and D267 for metal ion binding. We propose that the active site binds two metal ions, one in a high-affinity and one in a low-affinity site.

INTRODUCTION

Bacterial conjugative DNA transfer is a major driving force of horizontal gene transfer (HGT). Mobile Genetic Elements (MGE) like conjugative plasmids and Integrative Conjugative Elements (ICE) are transferred via conjugative Type-IV-Secretion Systems in a unidirectional fashion into recipient cells, thereby promoting genome plasticity and acquisition of novel traits (1–5). *Neisseria gonorrhoeae*, the causative agent of gonorrhea, encodes a unique T4SS that secrets chromosomal ssDNA into the extracellular milieu, where the DNA can be taken up by highly competent *Neisseria* species (6–8) (for review see: (9)). Moreover, the secreted ssDNA has been shown to facilitate the initial steps of biofilm formation (10). The T4SS is encoded on the Gonococcal Genetic Island (GGI), a 59 kb genetic island found in approximately 80% of *N. gonorrhoeae* and 17.5% of *N. meningitidis* clinical isolates (6,7,11,12). The GGI was most likely horizontally acquired, as suggested by differences in GC-content and frequency of *Neisseria*-specific DNA-uptake sequences (7,13). It is flanked by XerD-recombination sites (*dif*), which have been shown to be exploited by mobile genetic elements like the CTXϕ phage to integrate into the *Vibrio cholerae* chromosome (7,14). The excision rate of the GGI is very low, which is the result of mutations in one of the two flanking *dif* sites (15). Whether the GGI once transmitted in a classical cell-to-cell conjugative fashion remains unknown.

The GGI of *N. gonorrhoeae* strain MS11 encodes at least 63 putative proteins encoded within at least four operons (7,9). The first operon encodes the components involved in targeting the substrate to the transport complex, among which are the relaxase and the coupling protein. A second and third operon are located in the opposite direction and encode the components of the transport complex (16,17). The *origin of transfer* (*oriT*) is located between the operon encoding the targeting components and operons encoding the transport complex (18). The rest of the GGI encodes mostly for proteins with an unknown function, although several proteins show homology to DNA processing and modifying proteins (7). Additionally, this region encodes homologs of the ParA and ParB partitioning proteins that are the only other proteins within this region that were also required for ssDNA secretion (7,17). Previously, more T4SS have been discovered that encode proteins with homology to $T4SS_{GGI}$-encoded proteins, referred to as GGI-like T4SS (17). These GGI-like T4SS were mostly found on

*To whom correspondence should be addressed. Tel: +49 761 2032631; Fax: +49 761 2032773; Email: chris.van.der.does@biologie.uni-freiburg.de

chromosomes of α-, β- and γ-proteobacteria, often in proximity to genes coding for integrases and transposases, suggesting an integrative-conjugative lifestyle. The analysis further revealed that the majority of the GGI-like T4SS contains a MOB_H family relaxase (TraI) which often clusters together with a TraD-like coupling protein and the small inner-membrane protein Yaa, while the mating-pair formation complex belongs to the MPF_F-type (17). In the Neisserial T4SS, both the coupling protein TraD and the relaxase TraI are essential for DNA secretion (19).

Relaxases are essential for conjugative DNA transfer. They have been clustered into different families based on homology of their N-terminal relaxase domain and 9 different relaxase families have been identified (MOB_F, MOB_P, MOB_Q, MOB_V, MOB_L, MOB_B, MOB_C, MOB_T, MOB_H) (20–22). DNA transfer is initiated by binding of the relaxase and accessory proteins (also called the relaxosome) to the DNA within the *oriT* region (23,24). The relaxase site- and strand-specifically performs a transesterification reaction via a catalytic residue on a scissile phosphate of the DNA backbone (25). In the majority of relaxases the catalytic residue is a tyrosine, while it has recently been shown that the $MobM_{pMV158}$ relaxase of the MOB_V family employs a catalytic histidine instead (26). The nucleophilic attack is facilitated by a divalent cation that aids in positioning of the DNA substrate in the relaxase active site and polarization of the scissile phosphate (25,27,28). The transesterification reaction leads to a covalent relaxase-ssDNA adduct, which is piloted to the secretion complex by interaction with the coupling protein followed by transport through the secretion complex into the recipient cell, which requires unfolding of the relaxase (29–36). In the recipient cell, the relaxase rejoins the DNA molecule by a second transesterification reaction (25). Recently, a cryoEM structure of the MOB_F family relaxase $TraI_{R1}$ revealed how $TraI_{R1}$ binds as a dimer to *oriT*-DNA, where one molecule is in a closed 'helicase' state while the other one is in an open 'transesterase' state (37).

Of the 9 described relaxase families, 6 families (MOB_F, MOB_P, MOB_Q, MOB_V, MOB_L, MOB_B) share the conserved HUH (U denotes a hydrophobic residue) metal binding motif which is characteristic of the HUH-endonuclease superfamily (20–22,25). The remaining families (MOB_C, MOB_T, MOB_H) are only distantly related to HUH-relaxases. For instance, MOB_C family relaxases contain a DEE motif for metal ion coordination similar to type-I restriction endonucleases (38). MOB_T family relaxases, which are related to rolling-circle replication proteins of the Rep_trans family, have been shown to employ a DDE motif for metal ion coordination (39). The TraI relaxase encoded on the GGI of *N. gonorrhoeae* (from here on called TraI) is a representative of the MOB_H family of relaxases (22). The MOB_H family is further subdivided into two major clades, MOB_{H1} and MOB_{H2}, the latter of which includes TraI of *N. gonorrhoeae*, and into the less well-resolved clade MOB_{H3} (22). Members of the MOB_H family are usually found on large plasmids, genomic islands and ICEs (22). They are often found together with mating pair formation complexes of the MPF_G, MPF_I, MPF_F, MPF_T and MPF_{FATA} families (21). A domain architecture that is often found is (i) an N-terminal relaxase domain (TraI_2) followed by (ii) a putatively disordered middle part of variable length and (iii) a DUF1528 C-terminal domain of unknown function (recently renamed to TraI_2_C) (18). The presence of an amphipathic α-helix at the very N-terminus, which has been suggested to mediate membrane association, is unique to TraI and a subset of related relaxases (18). The N-terminal relaxase domain of the MOB_{H1} and MOB_{H2} clades contains the alternative 3H motif (HHH) which is reminiscent of the HUH motif of classical relaxases, but exhibits a deviant sequence environment (22). Moreover, these relaxases contain a conserved HD motif which is associated with the HD domain phosphohydrolase superfamily (22,40). These proteins are involved in various processes like DNA metabolism and signal transduction, as for instance the deoxyribonuclease YfbR of *Escherichia coli*, the cyclic-di-GMP phosphodiesterase HD-GYP of *Bdellovibrio bacteriovorus* or the OxsA protein involved in oxetanocin-A biosynthesis in *Bacillus megaterium* (41–44). Members of the MOB_{H3} clade do neither contain the alternative 3H motif nor the HD-phosphohydrolase motif and no other tractable motifs (22) suggesting that these proteins have an alternative function.

Little is known about the family of MOB_H relaxases. Relaxases of the MOB_H family have been shown to be essential for transfer of ICEs of *Salmonella enterica*, *Vibrio cholera*, *Neisseria gonorrhoeae*, *Proteus mirabilis* and *Pectobacterium atrosepticum*, but no member of this family has been characterized biochemically (18,45–48). A crystal structure of the N-terminal domain is available from an uncharacterized protein of *Coxiella burnetii* (PDB: 3KQ5) that harbors predicted TraI_2 and TraI_2_C domains. Another crystal structure is available from the TraI_2_C domain of the uncharacterized protein of *Salmonella enterica* (PDB: 2IPQ) that is connected to an N-terminal TraI_2 domain. Both structures belong to members of the MOB_{H3} clade and hence do not contain the conserved sequence motifs. Surprisingly, secretion assays with *N. gonorrhoeae* strains expressing different TraI mutants showed that secretion was not abolished when mutations were created in the alternative 3H motif or the HD motif (18). It was proposed that Y93 initiates DNA processing and Y201 is required for termination or acts in DNA binding (18).

Here, we present the first biochemical characterization of a MOB_H family relaxase. TraI is a monomer in solution with an extended shape revealed by Small-Angle X-Ray Scattering (SAXS). We determined the *nic*-site within *oriT* and show that cleavage is dependent on Mn^{2+} and Co^{2+}. We show that both the alternative 3H motif as well as the HD motif are essential for the cleavage reaction and propose an active site that coordinates two metal ions. Moreover, we demonstrate that TraI does not form a stable covalent intermediate with the DNA and suggest that the cleavage reaction involves a hydrolysis event.

MATERIALS AND METHODS

Construction of WebLogo and multiple sequence alignment for sequence comparison

Based on a previous phylogenetic analysis of MOB_H relaxases (22), 8 sequences covering different parts of the phylogenetic tree were selected and submitted for

BLAST-analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Sequences (maximum: 1000 per submitted sequence) were retrieved, merged and submitted to the CD-Hit Suite Server (http://weizhong-lab.ucsd.edu/cdhit_suite/cgi-bin/index.cgi) to cluster similar sequences into representative sequences with a threshold of <80%. The remaining 585 sequences were subsequently aligned in MEGA 7 (v7.0.25) using the MUSCLE-algorithm (standard parameters). Incomplete sequences and sequences not containing the (H)HH motif and HD motif were removed, which resulted in 461 sequences. Analysis of the alignment identified a conserved glycine (G422 in *N. gonorrhoeae* TraI) as last residue in the relaxase domain that was conserved in more than 80% of the sequences. Sequences C-terminal of this glycine were removed and remaining sequences were re-aligned in MEGA 7 using the MUSCLE-algorithm. The resulting multiple-sequence alignment was submitted to the WebLogo server (https://weblogo.berkeley.edu/logo.cgi). To simplify visualization, positions with an information entropy <3.5 bits and amino acids with <0.1 bits were removed. Furthermore, the sequences representative of the MOB_{H1} and MOB_{H2} clades and the two sequences of the MOB_{H3} clade, for which crystal structures were available (PDBs: 3KQ5; 2IPQ) (22), were included in the alignment above and re-aligned. The sequences were extracted and depicted using BioEdit (v. 7.0.5.3).

Structural modeling

Structural models of the TraI_2 domain of *N. gonorrhoeae* TraI were generated by submission of the first 422 amino acids of TraI to the Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) (49), the I-Tasser (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) (50) and the RaptorX (http://raptorx.uchicago.edu/) (51) servers. Structural models were aligned in PyMOL (v 1.3) with the closest homolog of TraI in the database (PDB: 3KQ5) and HD domain containing phosphohydrolase structures (PDBs: 3CCG, 2O08 and 2OGI).

Strains, plasmids and primers

Strains, plasmids and primers used in this study are shown in Supplementary Tables S1–S4. DNA encoding His_6-TraI(28−427) was obtained via PCR, using primers that contained the His_6-tag sequence, on genomic DNA of *N. gonorrhoeae* strain MS11. The DNA was subsequently cloned into the NdeI and XhoI restriction sites of vector pET20b resulting in vector pSVA3775. DNA encoding TraI(674−850) and TraI(42−850) was obtained via PCR on genomic DNA of *N. gonorrhoeae* strain MS11 using primers listed in Supplementary Table S3. The DNA was cloned by FX-Cloning into the pBXC3GH vector (a gift from Raimund Dutzler & Eric Geertsma (Addgene plasmid # 47070; http://n2t.net/addgene:47070; RRID:Addgene_47070)) (52) thereby adding a C-terminal GFP:His_{10}-fusion, which resulted in vectors pSVA4133 and pSVA4134. Site-directed mutagenesis of TraI(28−427) was performed using pSVA3775 as a template. Briefly, a circular PCR was performed with overlapping primers (plasmids listed in Supplementary Table S3) containing the desired mutations. The template vector was digested with DpnI at 4°C over-night followed by heat-inactivation at 65°C for 20 min. The DNA was transformed into *E. coli* NEB 10-beta cells (New England Biolabs) and transformants were analyzed by colony PCR. All plasmid sequences were confirmed by sequencing (Eurofins Genomics).

Protein expression

Expression plasmids were transformed into *E. coli* BL21*(DE3) cells. Precultures containing plasmids coding for TraI(28−427) and mutants were grown overnight at 37°C in LB + 1% glucose + 100 μg/ml Ampicillin. Cultures were inoculated to an OD_{600} = 0.05 in LB + 0.5% glucose + 100 μg/ml Ampicillin. At OD_{600} = 0.6, the cultures were cold-shocked for at least 3 min in an ice bath before induction with 0.5 mM IPTG followed by growth at 30°C for 3 h. Cultures were pelleted, frozen in liquid nitrogen and stored at −80°C. Precultures containing plasmids coding for TraI(674−850) and TraI(42–850) were grown over-night at 37°C in LB + 50 μg/ml Kanamycin. Main cultures were inoculated to an OD_{600} = 0.05 in LB + 50 μg/ml Kanamycin. At OD_{600} = 0.6, the cultures were cold-shocked for at least 3 min in an ice bath before induction with 0.1% arabinose followed by growth at 30°C for 3 h. The cultures were pelleted, frozen in liquid nitrogen and stored at −80°C.

Protein purification

Cell pellets were resuspended in Buffer A (Tris–HCl pH 8.0, 100 mM NaCl, 5% glycerol) and DNAse I (Roche) and Complete EDTA-free protease Inhibitor (Roche) was added. After 30 min incubation on ice, cells were disrupted by using the French Press (18 000 psi). The cell lysate was centrifuged at 40 000 rpm (TI-60 rotor, Beckman) for 1 h at 4°C. The supernatant was loaded to a HisTrap HP 5 ml column (GE Healthcare) that was pre-equilibrated with Buffer A. The column was washed step-wise with 30 ml buffer A, 10 ml buffer A + 10 mM imidazole, 10 ml buffer A + 20 mM imidazole and 10 ml buffer A + 30 mM imidazole. The protein was eluted with an imidazole gradient from 30 to 400 mM imidazole in buffer A over 10 ml followed by 10 ml buffer A + 400 mM imidazole. The purified proteins were concentrated in Amicon Ultra 4 ml Centrifugal Filters (30 kDa cut-off for TraI(28−427) and mutants, 10 kDa cut-off for TraI(674−850), 50 kDa cut-off for TraI(42−850)). The concentrated proteins were further purified by Size-Exclusion Chromatography (SEC) either over a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) or over a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated in buffer SEC (50 mM Tris pH 8.0, 100 mM NaCl). Eluted TraI(28−427) and mutants were concentrated to 2 ml as described above and aliquots were frozen in liquid nitrogen and stored at −80°C. Eluted TraI(42−850) and TraI(674−850) were incubated with His-tagged Human Rhinovirus 3C Protease overnight at 4°C to remove the C-terminal GFP-fusion protein. The cleaved-off GFP-protein as well as the His-tagged Human Rhinovirus 3C Protease were removed by loading the sample on a HisTrap HP 5 ml column (GE Healthcare). The flow-through containing the

protein was concentrated as described above. Aliquots were frozen in liquid nitrogen and stored at −80°C. The protein concentration was determined with the BCA Protein Assay Macro Kit (Serva).

Small-Angle X-ray Scattering (SAXS)

Before use, possible aggregates were removed using a Superdex 200 10/300 column (GE Healthcare) pre-equilibrated with buffer SEC (50 mM Tris–HCl pH 8.0, 100 mM NaCl) at a flow rate of 0.5 ml/min at the EMBL-Lab outstation (Grenoble). SAXS data was collected on beamline BM29 at the ESRF Grenoble (53,54). The beamline was equipped with a PILATUS 1M detector (Dectris) at a fixed distance of 2.867 m. The achievable s-range was 0.025–5 nm^{-1} and the maximum measurable radius of gyration (R_g) of the investigated particles was 20 nm. All measurements were performed at 4°C. For online-SEC one frame was collected each two seconds and for static measurements ten frames with an exposure time of one second. By comparison of these frames, the possibility of radiation damage during the measurement was excluded. All used programs for data processing were part of the ATSAS Software package (Version 2.8.4), available on the EMBL website (http://www.embl-hamburg.de/biosaxs/software.html) (55,56). The primary data reduction was performed with the programs CHROMIXS and PRIMUS (57,58). The radius of gyration (R_g) and the forward scattering $I(0)$ were determined with the Guinier approximation (59) (implemented in PRIMUS (57)). The pair-distribution function $p(r)$, computed with the program GNOM (60), was used to estimate the maximum particle dimension (D_{max}). Low resolution *ab initio* models were calculated with GASBOR (61) and superimposed with calculated Phyre2 models (49) of TraI(28−427) and TraI(674−850) (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index), modeling mode: intense) using the SUPCOMB program (62). Bovine serum albumin (66 kDa) was used as a reference solution to determine the molecular weight from the forward scattering. Further, molecular weight was calculated by using the MoW2 Server (63) and the volume of correlation (Vc) (64).

Size exclusion chromatography - multi angle light scattering (SEC-MALS)

A Superdex 200 Increase 10/300 GL (GE Healthcare) column was equilibrated with buffer SEC. 200 µl samples containing 1 mg TraI(28−427), 0.3 mg TraI(674−850) or 0.66 mg TraI(42−850) were injected. Data were collected for 50 min at a flow rate of 0.6 ml/min and evaluated using the software ASTRA 7.1.2 by Wyatt Technologies. The analysis was done with a combined system, an Agilent 1260 HPLC connected with a light scatterer (miniDAWN TREOS II) and a refractive index detector (Optilab T-rEX™) from Wyatt Technologies.

Cleavage assays using fluorescently labeled ssDNA

The standard cleavage assay reaction mixtures contained 500 nM [Cy3]-labeled ssDNA (Supplementary Table S4) with 5–10 µM of TraI variants in the presence or absence of 5 mM of the different metal ions in buffer SEC. For all metal ions the chloride salts were used. The mixtures were incubated at 25°C for 1 h, unless otherwise indicated. The reaction was stopped by addition of 2× gel loading mix (90% formamide, 0.5% EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue) (65). The samples were heat-denatured at 95°C for at least 5 min. The gel electrophoresis was carried out similarly as described in (65). Briefly, the samples were loaded to a 15%-polyacrylamide–urea gel containing 8 M urea. Electrophoresis was carried out in 1× TBE-buffer at 200 V for 45–55 min. Fluorescently labeled polyT oligonucleotides were used as size standards. The gels were scanned with the FAS Digi Gel Documentation System (Nippon Genetics).

For determination of the *nic*-site, the samples were loaded to a large format (20 cm) 15% polyacrylamide-urea gel and the electrophoresis was performed at 300 V for 9 h at 10°C with the Protean II xi cell (Biorad). Oligonucleotides representing 16, 17 and 18 nucleotides long 5′ cleavage products of the *oriT*-ssDNA were used as size standard in this experiment. The gel was documented with the FAS Digi Gel Documentation System (Nippon Genetics). To determine the turn-over rate of the cleavage reaction, cleavage assays were carried out with 2 and 5 µM TraI(28−427) or TraI(42−850), respectively, in the presence of 10 µM [Cy3]-*oriT*-ssDNA in buffer SEC containing 5 mM $MnCl_2$. Reactions were stopped and analyzed by gel electrophoresis as described above. Signal intensities of the bands were quantified with the ImageJ (NIH) Software and relative cleavage was calculated. To calculate the relative cleavage, loss of signal of the tested oligonucleotide was related to the sum of the uncleaved oligonucleotide and the cleavage product in each lane. The curves were fitted in GraphPad PRISM (v. 6.07) with the one-phase exponential decay function. To determine whether the 5′ end of the cleavage product retained a 5′ phosphate, 25 µM TraI(28−427) was incubated with 2 µM unlabeled *oriT*-ssDNA in buffer SEC in the absence and presence of 5 mM $MnCl_2$ for 2 h at 25°C. Samples were heat-inactivated at 95°C for at least 10 min. Subsequently, the samples were cooled on ice and 5′ labeled by T4-polynucleotide kinase (NEB). 20 U T4-PNK, 10× PNK buffer and 0.125 µCi γ-^{32}P-ATP (Hartmann Analytic) were added. The samples were incubated for 1 h at 37°C. The reactions were stopped and analyzed by gel electrophoresis as described above. The gels were exposed to an Imaging Plate BAS-MS 2040 (Fujifilm) over-night. The screen was developed with a Typhoon™ FLA 9500 Laser Scanner (GE Healthcare).

RecJ-Assay

500 nM of [Cy3]-labeled oligonucleotides representing the 3′ cleavage product (3′ of *nic*), *oriT*-ssDNA and a T_{57}-oligonucleotide were incubated in NEBuffer 2 (NEB) with (+) and without (–) 30 U recombinant $RecJ_F$ (NEB) for 20 min at 37°C. The reactions were stopped and analyzed by gel electrophoresis as described above.

Characterization of the *oriT*

Different oligonucleotides (Supplementary Table S3) were 5′-labeled with [γ-^{32}P]ATP (Hartmann Analytic). 2 µl

oligonucleotide (100 μM) was mixed with 0.4 μl T4 Polynucleotide Kinase (NEB), 0.4 μl 10 x T4 Polynucleotide Kinase Reaction Buffer (NEB), 0.45 μl nuclease-free water (Carl Roth GmbH) and 0.75 μl (0.1875 μCi) [γ-^{32}P]ATP. The mixture was incubated for 30 min at 37°C, followed by heat-inactivation at 65°C for 20 min. Subsequently, 36 μl buffer SEC was added to each reaction mixture to obtain a final concentration of 5 μM labeled oligonucleotide. For the cleavage assays, 200 nM [γ-^{32}P]-labeled oligonucleotides were incubated with 3 μM TraI(28−427) in buffer SEC in the presence and absence of 5 mM $MnCl_2$ for 1 h at 25°C. The reactions were stopped and analyzed by gel electrophoresis as described above. The gels were exposed to an Imaging Plate BAS-MS 2040 (Fujifilm) over-night and scanned with either the Pharos FX™ Plus Molecular Imager (Biorad) or the Typhoon™ FLA 9500 Laser Scanner (GE Healthcare). Signal intensities of the bands were quantified with the ImageJ (NIH) Software and relative cleavage activity was calculated. For calculation, loss of signal of the tested oligonucleotide was related to loss of signal the oligonucleotide control (*oriT*-ssDNA) present in each experiment.

The sequences of the different oligonucleotides were submitted to the Mfold web server (http://unafold.rna.albany.edu/?q=mfold/DNA-Folding-Form) for secondary structure prediction. The folding temperature was set to 25°C and ionic conditions were set to 50 mM Na^+ and 5 mM Mg^{2+} (Mn^{2+} could not be selected), respectively. The first hit of every secondary structure prediction was exported in the dot-bracket representation.

Differential scanning fluorimetry (DSF)

DSF experiments were performed with the CFX96 Touch™ Real-Time System (Biorad) in the FRET-channel. 5 μM of TraI(28−427) or TraI(28−427) mutants were mixed in buffer SEC with 800 μM $MnCl_2$ in presence of a final concentration of 10× SYPRO ORANGE protein gel stain (Thermo Fisher). Thermal denaturation was measured with a temperature gradient of 1°C/min from 20°C to 95°C. The melting temperatures (T_m) were derived from the first derivative of the thermal denaturation profiles using GraphPad PRISM (v. 6.07).

Comparison of metal binding and cleavage activity

DSF experiments were performed with 5 μM TraI(28−427) in the presence of increasing concentrations of $MnCl_2$ and $CoCl_2$ as described above. Derived melting temperatures were plotted against the metal ion concentration and fitted with the sigmoidal curve fit (4 Parameter Logistic; max: 5000 iterations) in GraphPad PRISM (v. 6.07). Cleavage assays were carried out in the presence of 10 μM TraI(28−427) with increasing concentrations of $MnCl_2$ and $CoCl_2$ as described above. Signal intensities of the bands were quantified with the ImageJ (NIH) software and specific cleavage was calculated. For calculation, loss of signal of the tested oligonucleotide was related to the signal of the uncleaved oligonucleotide control. Specific cleavage was plotted against the metal ion concentration and fitted with the sigmoidal curve fit (four parameter logistic; max: 5000 iterations) in GraphPad PRISM (v. 6.07).

RESULTS

TraI is a monomeric protein with an elongated structure

TraI of *N. gonorrhoeae* is a member of the uncharacterized family of MOB_H relaxases (22). TraI has a size of 850 amino acids and consists of at least two domains (Figure 1A). Members of the MOB_H family of relaxases contain a characteristic N-terminal (TraI_2) domain with a size of 300−400 amino acids which is predicted to cleave ssDNA (this domain will be referred to as 'relaxase domain'). This domain contains two conserved sequence motifs: the alternative 3H motif, which is reminiscent of the classical 3H motif described for other relaxase families (MOB_F, MOB_B, MOB_P, MOB_Q, MOB_V, MOB_L) and the HD domain phosphohydrolase motif, which is among relaxases only found in the MOB_H family (Figure 1A) (20,22,40). Many MOB_H relaxases contain a C-terminal (TraI_2_C) domain of unknown function (formerly known as DUF1528) of ∼140 amino acids. The relaxase domain and the TraI_2_C domain of MOB_H relaxases are interconnected by a region with little to no sequence conservation. This region is putatively disordered and varies in length between ∼50 and 200 amino acids. In TraI of *N. gonorrhoeae,* this region is comparably large with ∼300 amino acids. Specific for TraI of *N. gonorrhoeae* and a small subset of MOB_H relaxases is an amphipathic α-helix at the N-terminus, which has been shown to mediate membrane association (Figure 1A) (18).

In order to functionally characterize TraI, we set-out to overexpress and purify the full-length protein and truncated variants containing the relaxase and the TraI_2_C domains (Figure 1A). Optimization of the overexpression and purification procedures resulted in three different stable proteins: (i) TraI(42−850), which lacks the N-terminal amphipathic α-helix, (ii) TraI(28−427), which contains the relaxase domain and (iii) TraI(674−850) which contains the TraI_2_C domain (Supplementary Figures S1A and B). Previously, it has been shown that TraI migrates slower on SDS-PAGE as expected based on its molecular weight (18). Despite lack of the N-terminal amphipathic α-helix, TraI(42−850) (91.2 kDa) still migrated at a position corresponding to ∼120 kDa (Figure S1A, Table 1), even when higher concentrations of SDS were used. Elution profiles of Size-Exclusion Chromatography (SEC) revealed that TraI(42−850), TraI(28−427) and TraI(674−850) eluted at volumes corresponding to molecular weights of 295 kDa, 46 kDa and 22 kDa, which suggested formation of higher oligomers for TraI(42−850) (Supplementary Figure S1B, Table 1). However, SEC followed by multi-angle light-scattering (SEC-MALS) clearly demonstrated that all three TraI variants were monomeric (Table 1, Supplementary Figures S1C–E).

Since extensive attempts to crystallize the TraI variants were not successful, Small-Angle X-Ray Scattering (SAXS) experiments were performed to determine the overall shapes of the proteins (Figure 1B, Supplementary Figure S2). TraI(28−427) and TraI(674−850) form globular domains, which roughly correspond with structures predicted by the Phyre2, I-Tasser and RaptorX servers (Figure 1B, Supplementary Figure S3) (49). TraI(42−850) forms an extended structure (Figure 1A, B) that might explain the early elu-

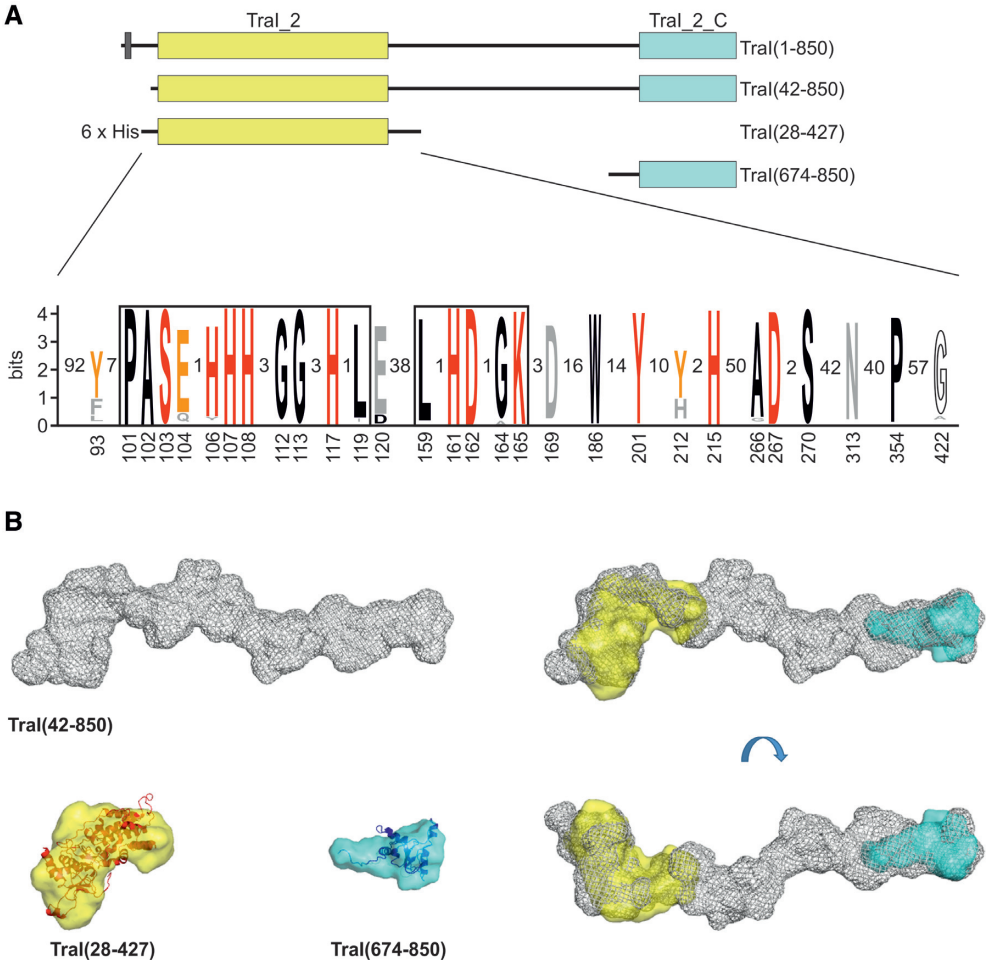

Figure 1. Purified TraI is a monomer with an extended structure. (**A**) Domain organization of *N. gonorrhoeae* TraI and variants used in this study. The dark grey box indicates the N-terminal amphipathic α-helix. The yellow and cyan boxes indicate the relaxase (TraI_2) and the TraI C-terminal (TraI_2_C) domains. 6× His indicates the N-terminal $(His)_6$-tag of TraI(28–427). A graphical representation (WebLogo) of the multiple sequence alignment of the TraI_2 domain is depicted with the overall height of the stack indicating the sequence conservation and the individual height of symbols the relative frequency of each amino acid at that position, respectively (99). Only positions with an information entropy above 3.5 bits are shown. Within these positions, amino acids with an individual height <0.1 bits were removed. Numbers below indicate the position in the *N. gonorrhoeae* TraI sequence. Red and orange symbols indicate amino acids analyzed in this study, with orange symbols indicating positions with an information entropy below 3.5 bits. Grey symbols indicate amino acids not conserved in TraI of *N. gonorrhoeae*. The alternative 3H motif and the HD domain phosphohydrolase motifs as defined by Garcillán-Barcía *et al.* (22) are indicated by black boxes. G422 (hollow letter) represents the last conserved amino acid before start of the putatively disordered middle part. (**B**) SAXS analysis of the TraI variants. The volumetric envelopes were calculated from scattering data using GASBOR (61) and are shown for TraI(42–850) (gray), TraI(28–427) (yellow) and TraI(674–850) (cyan). Structural models predicted by Phyre2 were docked into the volumetric envelopes using SUPCOMB (62).

tion during SEC (Supplementary Figure S1B). Molecular weight determinations from the SAXS data confirmed a monomeric state of all TraI variants in solution (Table 1, Supplementary Table S5).

TraI cleaves *oriT*-ssDNA in a site-specific and Mn^{2+}-dependent manner

Conjugative DNA transfer is initiated by recognition and processing of the *origin of transfer* (*oriT*) by the relaxosome. The relaxase cleaves ssDNA in a site- and strand-specific reaction within the *nic*-site (66). The cleavage reaction is specific to the relaxase domain and requires at least one divalent metal ion in the active site (38,67). Previously, the *oriT* of the GGI was identified in the intergenic region between the *yaf* and *ltgX* genes, the first genes of the *yaf-yaa* and *ltgX-traF* operons, respectively (Figure 2A) (9,18). This region contains an inverted repeat (IR) and disruption of this region by insertion of a *cat*-marker resulted in loss of DNA secretion (18).

To test *in vitro* cleavage of *oriT* containing ssDNA, a 75 nts long 5′-[Cy3]-labeled oligonucleotide (Cy3-*oriT*-ssDNA) spanning the IR was incubated with purified TraI(28−427) in the presence of different divalent metal ions (Figure 2A, Supplementary Figure S4). The samples were separated on denaturing polyacrylamide-urea gels. TraI(28−427) cleaved the Cy3-*oriT*-ssDNA in the presence of Mn^{2+} and Co^{2+} but not with Mg^{2+}, Zn^{2+}, Ca^{2+}, Cu^{2+} and Ni^{2+} (Supplementary Figure S4). No cleavage was observed with unspecific polyT DNA (Supplementary Figure S5). Cleavage in the presence of increasing concentrations

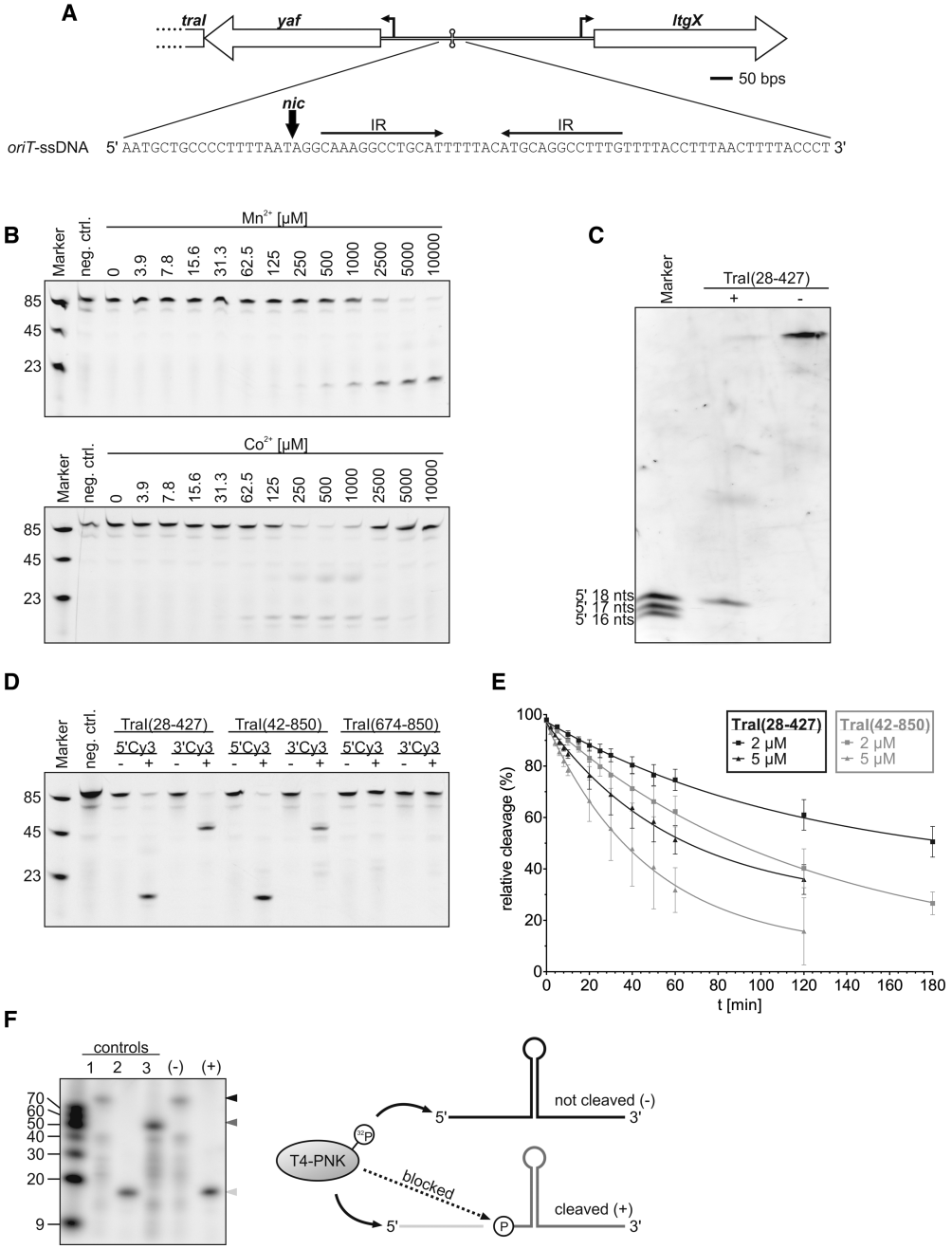

Figure 2. TraI cleaves ssDNA in a sequence specific manner. (**A**) Map of the GGI region containing the *oriT*. *yaf* encodes a protein of unknown function and *traI* the relaxase. *ltgX* encodes a putative lytic transglycosylase. Arrows mark the promoter regions for the *yaf-yaa* and *yaf-traF* transcripts. The 75 nucleotides long oligonucleotide (*oriT*-ssDNA) which was used for the cleavage assays is indicated below. The IR sequences and the *nic*-site are indicated. (**B**–**D**) Cy3-labeled oligonucleotides were incubated with TraI for 1 h at 25°C and separated on denaturing polyacrylamide-urea gels. Representative gels from at least 3 experiments are shown. (B) 500 nM [Cy3]-*oriT*-ssDNA was incubated with 10 μM TraI(28–427) in the presence of increasing concentrations of Mn^{2+} or Co^{2+}. Marker lanes show labeled polyT oligonucleotides with indicated lengths. *oriT* containing oligonucleotides migrated slower than expected based on their size, most likely because the hairpin structure was not completely denatured. (C) 500 nM [Cy3]-*oriT*-ssDNA was incubated with 10 μM TraI(28–427) in the absence (–) and presence (+) of 5 mM Mn^{2+} and separated on a high-resolution polyacrylamide-urea gel. Marker lane shows oligonucleotides representing 16, 17 and 18 nucleotides of the 5′ end of the *oriT*-ssDNA. (D) 500 nM [Cy3]-*oriT*-ssDNA (5′Cy3) and *oriT*-ssDNA-[Cy3] (3′Cy3) were incubated with 10 μM TraI(28–427), TraI(42–850) or TraI(674–850) in the absence (–) and presence (+) of 5 mM Mn^{2+}. Marker lanes show labeled polyT oligonucleotides with indicated lengths. (E) 2 μM (squares) and 5 μM (triangles) TraI(28–427) (black symbols and line) or TraI(42–850) (gray symbols and line) were incubated with 10 μM [Cy3]-*oriT*-ssDNA in the presence of 5 mM Mn^{2+} for up to 180 min. Samples were separated on denaturing polyacrylamide-urea gels and the relative cleavage of *oriT*-ssDNA was determined. Standard deviations derived from three to four independent experiments are shown. (**F**) 25 μM TraI(28–427) was incubated with 2 μM unlabeled *oriT*-ssDNA in the absence (–) and in the presence (+) of 5 mM Mn^{2+} for 2 h. TraI was heat-inactivated and subsequently incubated with T4-polynucleotide kinase (PNK) in the presence of γ-[^{32}P]-ATP for 1 h followed by separation on denaturing polyacrylamide-urea gels. As size controls, oligonucleotides representing the uncleaved *oriT*-ssDNA (1), the 5′ cleavage product (2) and the 3′ cleavage product (3) were subjected to 5′ labeling by PNK. A representative gel from 4 experiments is shown. In the schematic on the right, the different grey tones of the oligonucleotides correspond to the colors of the arrows.

Table 1. Comparison of molar masses (kDa) of TraI variants determined by different methods

	Theoretical mass	SDS-PAGE	SEC	SEC-MALS	SAXS (MoW2)	SAXS (Vc)
TraI(42–850)	91.2	120	294.6	94.6 ± 4.7	97.8	89.6
TraI(28–427)	46	50	45.7	46.9 ± 2.3	43.8	42.1
TraI(674–850)	20.8	22	30.5	21.8 ± 1.1	17.9	20.4

of Mn^{2+} and Co^{2+} revealed optimal activity at 5 mM Mn^{2+} and 0.5 mM Co^{2+}, respectively (Figure 2B). At higher Co^{2+} concentrations, the cleavage activity decreased due to protein instability (Figures 2B and 6). Mn^{2+}-supported cleavage resulted in one specific cleavage product (<23 nts) with a faintly visible unspecific product (between 23 nts and 45 nts), while Co^{2+}-supported cleavage resulted in the same main cleavage product but with more additional side products (Figure 2B). Therefore, experiments were continued in the presence of Mn^{2+}. To determine the *nic*-site, which is the exact position where TraI cleaves within the *oriT*, the cleavage product was analyzed on a high-resolution denaturing polyacrylamide-urea gel (Figure 2C). The size of the cleavage product was compared to Cy3-labeled oligonucleotides representing the first 16, 17 and 18 nucleotides of the *oriT*-ssDNA oligonucleotide, which revealed cleavage at a position 18 nucleotides from the 5′ end of the *oriT*-ssDNA (Figure 2C).

TraI cleaves *oriT*-ssDNA without formation of a stable TraI-DNA intermediate

Relaxases form a stable covalent intermediate with the DNA (25). This enables the transport of the relaxase with the DNA in an unfolded state into the recipient cell (36). The majority of known relaxases forms a stable phosphotyrosyl-DNA intermediate. An exception is the MOB_V family relaxase $MobM_{pMV153}$ which has been suggested to form a phosphoamidate-DNA intermediate which involves a histidine residue (26). Remarkably, no covalent relaxase-DNA intermediate could be detected in the MOB_C family relaxase $MobC_{pCloDF13}$ (68). To examine whether TraI forms a stable covalent bond with the cleaved DNA, TraI(28−427) was incubated with 5′ and 3′ [Cy3]-labeled *oriT* containing oligonucleotides (5′Cy3 and 3′Cy3) in the presence and absence of 5 mM Mn^{2+} (Figure 2D). The covalent attachment of the relaxase to the 3′-labeled DNA leads to an upshift of the protein-DNA complex in polyacrylamide-urea gels (69). Surprisingly, no shift was observed indicating that TraI(28−427) cleaves the *oriT*-ssDNA without formation of a stable covalent intermediate (Figure 2D). A similar result was obtained for TraI(42−850) which contains both the relaxase domain and the C-terminal TraI_2_C domain. As expected, TraI(674−850) containing only the TraI_2_C domain did not cleave the *oriT*-ssDNA, which demonstrates that the cleavage activity resides solely in the relaxase domain (Figure 2D).

Next, we compared the rate of cleavage of TraI(28−427) and TraI(42−850) to examine whether the presence of the putatively disordered middle part and the C-terminal TraI_2_C domain affect the cleavage rate. The initial cleavage rate was slightly higher for TraI(42−850) compared to TraI(28−427) at the same concentration, demonstrating an effect of the C-terminal domains on cleavage activity (Figure 2E). Moreover, the results for TraI(42−850) show that it performs multiple cleavage reactions further demonstrating the absence of a stable covalent TraI-ssDNA intermediate (Figure 2E).

The 3′ cleavage product most likely contains a 5′ phosphate group

The nucleophilic attack carried out by relaxases results in attachment of the relaxase to the 5′ phosphate of the cleaved ssDNA. This generates a free 3′-OH which can serve as primer for DNA replication and as nucleophile to rejoin the transferred DNA after conjugation terminates (25). Since no stable covalent TraI-ssDNA intermediate could be detected, we wondered whether the 5′ end of the 3′ cleavage product retained its phosphate, as usually observed for DNA hydrolysis reactions (70). TraI(28−427) was incubated with unlabeled *oriT*-ssDNA in the presence and absence of 5 mM Mn^{2+} until cleavage was complete. After heat-inactivation of TraI, the samples were subjected to [^{32}P]-labeling by T4-polynucleotide kinase (PNK) which adds a phosphate group to the free 5′-OH ends of DNA (Figure 2F). Indeed, only the 5′ cleavage product was labeled, while the 5′ end of the 3′ cleavage product was blocked. As controls, the *oriT*-ssDNA oligonucleotide and oligonucleotides representing the 5′ and the 3′ cleavage product harboring phosphate-free 5′ ends could successfully be labeled by PNK (Figure 2F). These results suggest the presence of a 5′ phosphate on the 3′ cleavage product supporting the idea that a hydrolysis reaction takes place.

We showed that no stable covalent intermediate was formed between TraI and the DNA. This seemed to contradict the previous observation that secreted ssDNA was protected from degradation by RecJ, which is a 5′ specific exonuclease that exclusively targets ssDNA (18). It was previously shown that RecJ requires a ≥7 nts ssDNA tail to properly initiate its exonuclease activity (71). Our results showed that the *nic*-site is located only 3 nucleotides 5′ of the double-stranded hairpin structure (Figure 2C). Indeed, while incubation of RecJ with the *oriT*-ssDNA oligonucleotide harboring a 18 nucleotides long 5′ ssDNA tail resulted in degradation of the oligonucleotide, RecJ was unable to degrade an oligonucleotide representing the 3′ cleavage product (Supplementary Figure S6). Thus the previously observed 5′ protection of the secreted ssDNA from degradation by RecJ was most likely caused by the presence of the double-stranded hairpin structure of the IR.

Determination of the minimal requirements of the *oriT*

The *oriT* is an essential element in the process of conjugative DNA transfer to discriminate between transfer DNA and host DNA. Crystal structures of relaxases bound to their cognate *oriT*s revealed specific interactions with single nucleotides and the importance of DNA hairpin structures (26,27,72). To obtain more insight into the TraI-*oriT* interaction, cleavage studies with oligonucleotides containing various mutations and truncations were performed (Figure 3). All tested oligonucleotides are derivatives of the 75 nucleotides long *oriT*-ssDNA oligonucleotide used in the cleavage assays above (Figure 2). A secondary structure prediction of the oligonucleotides is shown in Supplementary Figure S7. Cleavage of each oligonucleotide was normalized to the *oriT*-ssDNA oligonucleotide. Mutations to the complementary bases at positions –3 to +5 relative to the *nic*-site (Figure 3, lanes 1–8) showed that mutation of nucleotides at positions –3, –2 and +1 (Figure 3, lane 1, 2, 4) resulted in reduced cleavage, whereas mutation of the nucleotide at the –1 position (Figure 3, lane 3) abolished cleavage. Mutation of nucleotides at positions +2, +3 and +5 (Figure 3, lanes 5, 6, 8) had almost no effect, while mutation of the nucleotide at position +4 (Figure 3, lane 7) strongly affected cleavage. The importance of this nucleotide was further confirmed by the observation that the triple-mutation of nucleotides +2, +3 and +4 (Figure 3, lane 9) also strongly affected cleavage. Thus, nucleotides around the *nic*-site, especially at position -1, are important for cleavage.

The role of the IR was analyzed using oligonucleotides with mutations that impaired base-pairing in the hairpin structure (Figure 3, lanes 10–14, Supplementary Figure S7). An AAA to TTT triple-mutation (Figure 3, lane 10) in the base of the IR proximal arm resulted in reduced cleavage. Introducing 4-nucleotide or 6-nucleotide substitutions (Figure 3, lanes 11, 12) in the middle of the IR proximal arm abolished cleavage, suggesting that hairpin formation at this position is important for cleavage. This was further confirmed by additional complementary mutations in the distal arm (Figure 3, lanes 13, 14) which restored the hairpin structure. These mutations resulted in cleavage activities comparable to the control. Therefore, rather the structure than the sequence of the IR is important for cleavage. Surprisingly, introducing the 4-nucleotide or 6-nucleotide complementary substitutions (Figure 3, lanes 15, 16) only within the IR distal arm resulted in reduced, but not abolished cleavage. Possibly, alternative secondary structures are formed that mimic a hairpin structure. Removal of 4 nucleotides of the hairpin loop (Figure 3, lane 17) did not affect cleavage, showing that the sequence of the hairpin loop is not important for cleavage.

Finally, the minimal *oriT* sequence that is required for efficient cleavage was determined (Figure 3, lanes 18–25). To narrow down the minimal required length at the 3′ end, the *oriT*-ssDNA was truncated for 10 and 14 nucleotides (Figure 3, lanes 18, 19), respectively. While truncation of 10 nucleotides resulted in cleavage similar to the control, truncation of 14 nucleotides abolished cleavage. This demonstrates that a length of 43 and 47 nucleotides is required 3′ of the *nic*-site for efficient cleavage. Consecutive removal of nucleotides 5′ of the *nic*-site revealed full cleavage when 3–4 nucleotides were still present (Figure 3, lanes 20, 21), but resulted in reduced cleavage with only two nucleotides present (Figure 3, lane 22). This demonstrates that 3 nucleotides are required 5′ of the *nic*-site for efficient cleavage. Cleavage was abolished when only one nucleotide was present 5′ of the nic (Figure 3, lane 23). Further 3′ truncation (Figure 3, lanes 24, 25) revealed that cleavage of the oligonucleotides with a length of 48 and 49 was further impaired, whereas the 50 nucleotide long oligonucleotide was still cleaved. Thus, the minimal required sequence for cleavage by TraI is 50 nucleotides long, requires an inverted repeat 3′ of the *nic*-site and requires specific nucleotides around the *nic*-site and at the +4 position.

Y212 is important for cleavage of ssDNA

The majority of characterized relaxases employs a catalytic tyrosine that acts as a nucleophile to form a covalent phospotyrosyl intermediate with the DNA. TraI contains several tyrosines in the relaxase domain. Y201 is highly conserved, while Y93 and Y212 show only limited conservation (Figure 1A, Supplementary Figure S8). Y93 is frequently replaced by the structurally similar phenylalanine in other MOB_H-relaxases. Y212 is mostly present in the MOB_{H2} clade, whereas it is often replaced by a histidine in the MOB_{H1} clade (Figure 1A, Supplementary Figure S8). Mutation of these tyrosines to phenylalanines resulted for the Y93F mutant in abolishment of DNA secretion, whereas the Y201F and Y212F mutants still secreted DNA (18). In order to examine the function of these tyrosines *in vitro*, TraI(28–427) mutants containing either a single tyrosine-to-alanine substitution or a tyrosine-to-phenylalanine substitution were purified and analyzed in cleavage assays (Figure 4, Supplementary Figure S9). Surprisingly, both the Y93 mutants and the Y201 mutants still showed cleavage activity, demonstrating that these tyrosines do not function as catalytic tyrosines. Cleavage activity was abolished for Y212A, whereas residual activity was observed for Y212F (Figure 4). Since Y212 is often replaced by a histidine in the MOB_{H1} clade, Y212 was also replaced by a histidine. Similar to the Y212F mutant, the Y212H mutant showed residual activity. To prove correct folding of the Y212A, Y212F and Y212H mutants, differential scanning fluorimetry (DSF) experiments were conducted in absence and presence of Mn^{2+} (Figure 5B). TraI(28–427) wild-type (WT) and Y212A, Y212F and Y212H mutants revealed melting temperatures between 44 and 46°C in absence of Mn^{2+}, while presence of Mn^{2+} increased melting temperatures up to 49–51°C, demonstrating that the proteins were similarly stable and capable of Mn^{2+} ion binding. Thus, our *in vitro* results show that Y212 is important, but not essential for cleavage.

TraI might contain a two metal ion binding site

Classical HUH-relaxases coordinate a single divalent metal ion in their active site (26,72,73). TraX of the MOB_C family, which bears similarity with PD-(D/E)XK restriction enzymes like BamHI, has been proposed to coordinate two metal ions similarly to these restriction enzymes (38). TraI contains two motifs that are suggested to be involved in coordination of divalent metal ions, the alternative 3H motif

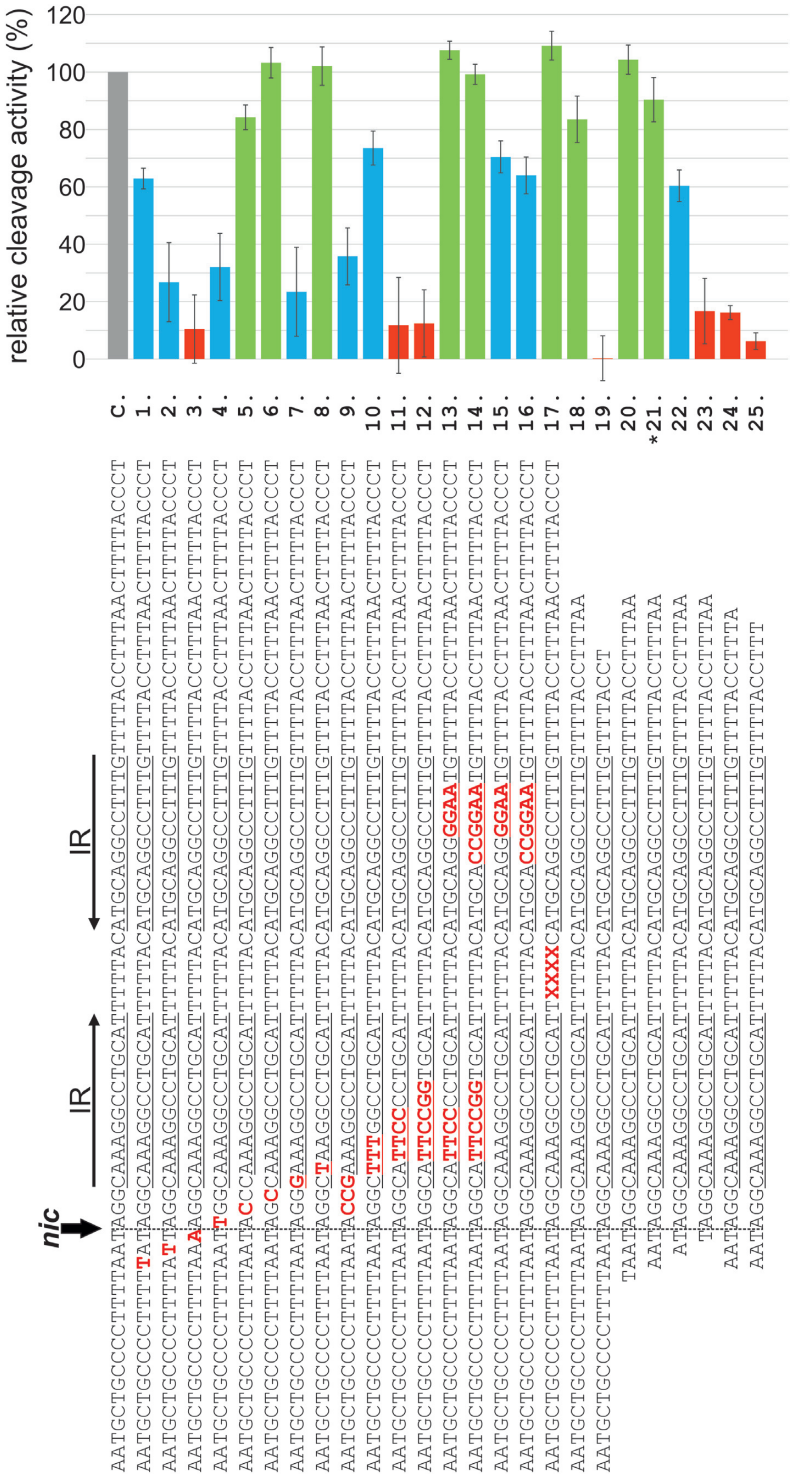

Figure 3. Identification of the minimal functional *oriT*. 3 μM TraI(28–427) was incubated with 200 nM of different [^{32}P]-labeled *oriT*-ssDNA oligonucleotides in the absence and presence of 5 mM Mn^{2+} for 1 h at 25°C. Samples were analyzed on denaturing polyacrylamide-urea gels and signal intensities were quantified with the ImageJ software. Signal loss of the uncleaved oligonucleotide was determined and normalized to a [^{32}P]-*oriT*-ssDNA control. Relative cleavage activity was determined from three to four replicates for each oligonucleotide. Relative cleavage activities <20% are shown in red, 20–80% in blue and >80% in green. Nucleotides that deviate from the original sequence are colored red. The oligonucleotide representing the minimal *oriT*-ssDNA (50 nts) required for efficient cleavage is indicated by a star. The *nic* site is indicated by a dashed line. The sequences forming the IR are underlined.

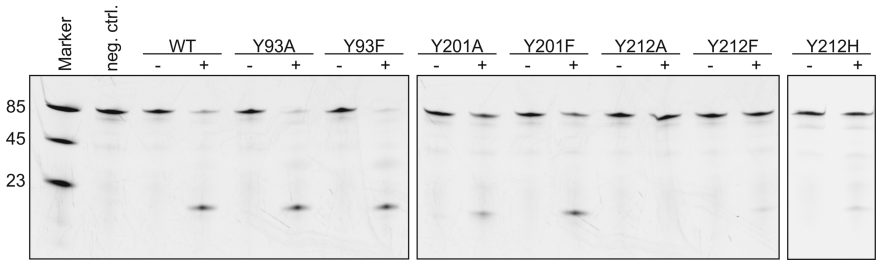

Figure 4. Y212 is required for cleavage. 5 μM of TraI(28–427) or TraI(28–427) variants containing different tyrosine mutations was incubated with 500 nM [Cy3]-*oriT*-ssDNA in the absence (–) and presence (+) of 5 mM Mn^{2+} for 1 h at 25°C. The samples were analyzed on denaturing polyacrylamide-urea gels. Marker lane shows [Cy3]-polyT oligonucleotides with indicated lengths. A representative gel from 3 experiments is shown.

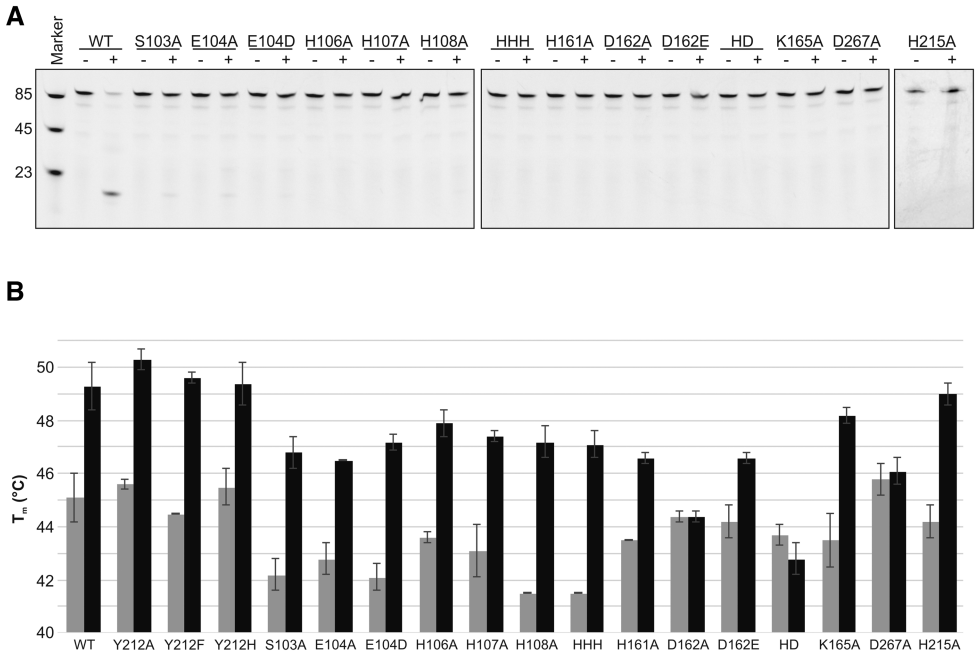

Figure 5. Identification of amino acids involved in cleavage activity and Mn^{2+}-binding. (**A**) 10 μM of TraI(28–427) or TraI(28–427) variants containing different mutations in the putative metal binding residues was incubated with 500 nM [Cy3]-*oriT*-ssDNA in the absence (–) and presence (+) of 5 mM Mn^{2+} for 1 h at 25°C. The samples were analyzed on denaturing polyacrylamide-urea gels. Marker lane shows [Cy3]-polyT oligonucleotides with indicated lengths. WT = wildtype; HHH = H106A/H107/H108A; HD = H161A/D162A. (**B**) TraI(28–427) and TraI(28–427) containing different mutations in conserved tyrosine residues or in putative metal binding residues were analyzed for Mn^{2+} binding by differential scanning fluorimetry (DSF). The thermal denaturation profile of each protein was determined at a concentration of 5 μM in absence and presence of 800 μM Mn^{2+} with a stepwise temperature increase of 1°C/min from 20°C to 98°C. Melting temperatures (T_m) were derived from the first derivatives of the thermal denaturation profiles. Each mutant was analyzed in 3–4 independent experiments with each 2–3 technical replicates. WT = wildtype; HHH = H106A/H107/H108A; HD = H161A/D162A.

and the HD domain phosphohydrolase motif. The latter belongs to the HD superfamily of metal-dependent phosphohydrolases which have been identified with mononuclear-, dinuclear- or trinuclear metal-centers (40–42,44,74). To address which residues are involved in metal binding, 11 conserved residues in the relaxase domain that might contribute to the metal binding site (Figure 1A), were selected, mutated and the respective mutants were overexpressed. The mutants, except for H117A which was highly instable, were purified and analyzed in cleavage assays in the presence and the absence of Mn^{2+} (Figure 5A). Cleavage activity was strongly reduced or abolished for all of the tested mutants. This demonstrated that these residues are essential for ssDNA cleavage *in vitro* and suggests that they are either involved in metal ion binding, reaction catalysis and/or DNA binding. To further examine the role of these residues in metal binding, stabilization of TraI(28−427) and the TraI(28−427) mutants was assayed in the presence of Mn^{2+} by DSF (Figure 5B). All mutants exhibited melting temperatures of 41.5 – 46°C compared to 44.5°C of the TraI(28−427) WT in absence of Mn^{2+}, showing that the mutants were overall correctly folded. The TraI proteins containing the D162A and D267A mutations were not stabilized in the presence of Mn^{2+} suggesting loss of their ability to bind Mn^{2+}. However, TraI proteins containing mutations of the putative metal coordinating residues S103, E104, H106, H107 and H108 of the alternative 3H motif, H161 and K165 of the HD domain motif and the conserved H215, were still stabilized by the addition of Mn^{2+}. Even a TraI protein containing a H106A/H107A/H108A triple-mutation was still stabilized in the presence of Mn^{2+} (Figure 5B). Although it has been observed that mutation of single

residues in a metal binding site does not necessarily abolish metal ion coordination, it is surprising that the HHH-triple mutant still binds Mn^{2+}, particularly with regard to examples of relaxases where metal binding is strongly affected upon mutagenesis of a single amino acid in the HUH motif (28,75).

To confirm the presence of two metal ions in the active site, the change of melting temperature and cleavage activity in the presence of an increasing concentration of Mn^{2+} ions was analyzed (Figure 6A, left panel). The half-maximal melting temperature increase was reached at ∼50 μM Mn^{2+} whereas the half-maximal cleavage activity was reached at ∼1300 μM Mn^{2+}. This strongly suggests that TraI indeed harbors two metal binding sites, one with high-affinity and one with low-affinity for Mn^{2+}. To examine whether this observation is specific for Mn^{2+}, the experiment was repeated in the presence of increasing concentrations of Co^{2+} (Figure 6A, right panel). Similarly, we observed stabilization of the protein at low Co^{2+} concentrations and stimulation of the cleavage activity only at higher Co^{2+} concentrations. Thus, we propose that TraI harbors two metal ions in its active site, where one metal ions binds to a higher-affinity binding site which results in stabilization of TraI, while binding of a second metal ion to a low-affinity binding site completes the active site and is therefore required for activity.

To model the structure of the putative two metal binding site, the RCSB Protein Data Bank was searched for crystal structures of (putative) HD domain phosphohydrolases which contained a metal binding site and showed structural similarities with the closest homolog of TraI in the database (PDB: 3KQ5). Indeed, several structures (PDBs: 3CCG, 2O08, 2OGI) were identified that showed structural similarities to the 3KQ5 structure (Supplementary Figure S10). These structures all contained two Fe^{2+} ions in a dinuclear metal center with a $HX_nHDX_nHX_nHX_nD$ consensus motif (42). This motif mostly overlaps with the $HX_nHDX_nHX_nD$ motif found in MOB_H relaxases and suggests that H117 is not part of the alternative 3H motif, but of the HD domain. Comparison of the positions of the metal binding residues in the 3CCG, 2O08 and 2OGI structures with the proposed metal binding residues in the modeled TraI structure shows a similar arrangement of the coordinating residues (for comparison of the modeled TraI structure with the 3CCG structure, see Supplementary Figure S11). In this two metal ion binding site, one site is coordinated by residues H117, H161, D162 and D267 and the second site is coordinated by D162 and H215. In comparison to the 3CCG, 2O08 and 2OGI structures, the second site lacks a coordinating histidine. In the modeled structures, the alternative 3H motif locates on a random coil close the active site, leaving it uncertain whether this motif is involved in metal ion coordination or possibly in DNA binding.

DISCUSSION

TraI(42–850) has an extended structure

Relaxases have previously been clustered into different families based on their N-terminal relaxase domain which harbors the cleavage and DNA rejoining activities. While considerable progress has been made in studying several relaxase families, little is known about the MOB_H family of relaxases. Several relaxases contain additional domains with helicase and primase functions, while many members of the MOB_H family contain the TraI_2_C domain of unknown function, which is separated of the relaxase domain by a putatively disordered middle domain (18,22,24). Previously, relaxases have been identified as monomers or dimers (27,37,76–78). To determine whether the additional

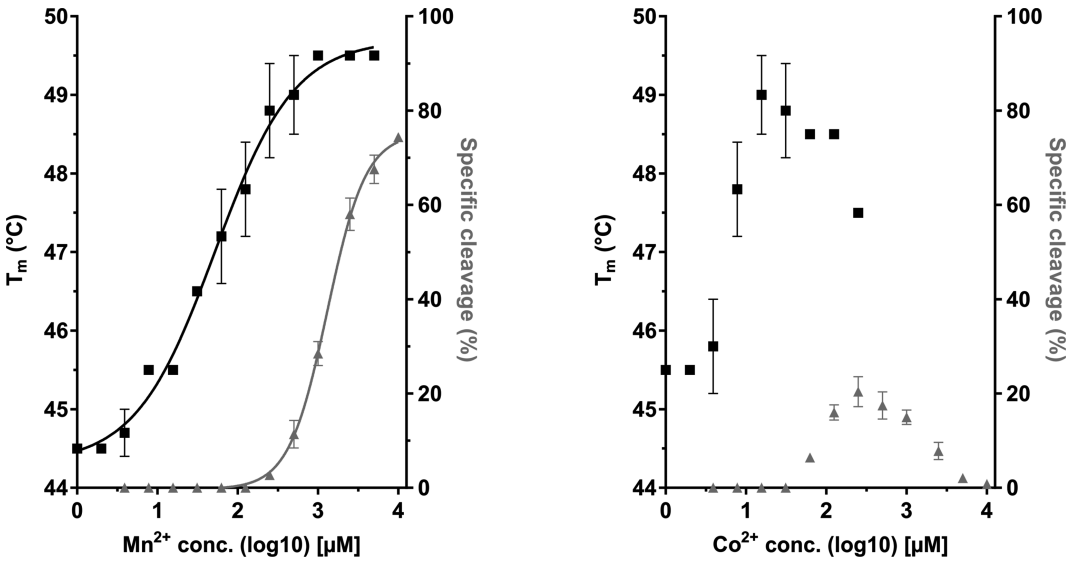

Figure 6. Comparison of metal binding and cleavage activity strongly suggests two metal binding sites. (**A**) Melting temperatures (black squares) and specific cleavage activities (gray triangles) of TraI(28–427) were determined in the presence of increasing Mn^{2+} (left) and Co^{2+} (right) concentrations. Melting temperatures were determined by incubation of 5 μM of TraI(28–427) with a stepwise temperature increase of 1°C/min from 20 to 98°C. Each data point is an average of three independent experiments with 2 technical replicates. Cleavage activity was determined by incubation of 10 μM TraI(28–427) with 500 nM [Cy3]-*oriT*-ssDNA for 1 h at 25°C. Signal increase of the 18 nucleotides cleavage product was normalized to the negative control. Data points are averages of 3–4 independent experiments.

domains influence the structure or the activity of TraI, three TraI variants were purified. SEC-MALS analysis demonstrated that all three TraI variants formed monomers in solution (Supplementary Figures S1C-E). SAXS experiments revealed that the relaxase and the TraI_C_2 domains formed globular structures, while TraI(42−850) formed a long extended structure (Figure 1B). The purified proteins were used in cleavage assays with *oriT*-ssDNA and, as expected, cleavage of *oriT*-ssDNA was observed in the presence of the TraI(28−427) relaxase domain, while the TraI_2_C domain showed no cleavage activity. Cleavage activity was slightly increased for TraI(42−850) compared to TraI(28−427) indicating that the presence of the domains C-terminal of the relaxase domain might enhance the cleavage activity of the relaxase domain. While several structures of isolated relaxase domains or C-terminal domains have been solved, the only available structure of a relaxase with multiple domains is the structure of MOB_F family TraI of the R1-plasmid ($TraI_{R1}$) (26,27,37,67,72,73,78–82). $TraI_{R1}$ contains four domains, a relaxase domain, a vestigial and an active helicase domain and a C-terminal domain that functions as a recruitment platform for relaxosome components (37). The structure of the highly homologous MOB_F family $TraI_F$ obtained from SAXS experiments also predicted a monomeric elongated structure, whereas the high-resolution cryo-EM structure of the DNA bound $TraI_{R1}$ showed a more globular structure with large contacts between the different domains (83). TraI of *N. gonorrhoeae* might also undergo such a conformational change upon DNA binding. However, more research will be required to elucidate the role of the TraI_2_C domain. In this study we focused on the characterization of the N-terminal relaxase domain.

Analysis of the TraI-*oriT* interaction

The interaction between the relaxase and its cognate *oriT* is essential to identify the target DNA. *oriT* sequences usually contain at least one IR sequence with the *nic* site adjacent to it ((84) and references therein). Cleavage assays were performed with different variants of the *oriT*-ssDNA oligonucleotide to determine the *nic*-site and the requirements for efficient cleavage. The minimal *oriT* for cleavage contains 50 nucleotides, of which 32 form the IR, 6 nucleotides that are required 5′ of the IR and 12 nucleotides 3′ of the IR. Introduction of complementary mutations in the IR sequence revealed that the hairpin structure rather than the sequence of the IR is required for cleavage. The position of the *nic*-site was determined 3 nucleotides 5′ of the IR, which differs from relaxases of the well characterized MOB_F, MOB_P, MOB_Q and MOB_V families. These relaxases cleave 3′ of the IR and often contain 8–12 nucleotides between the *nic*-site and the IR ((84,85) and references therein). Crystal structures of MOB_F-relaxases revealed that the nucleotides between the IR and the *nic*-site surround the relaxase in a U-turn shape which properly positions the *nic*-site in the relaxase active site (82,86,87). The crystal structures of the MOB_Q relaxases NES and MobM revealed similar roles for the nucleotides between the IR and the *nic*-site forming a U-turn like shape (26,72). Several nucleotides around the *nic*-site have been found to be required for efficient cleavage by TraI, indicating that these nucleotides form contacts with the protein. Moreover, 12 nucleotides 3′ of the IR are required for cleavage possibly contributing to the TraI-DNA interaction. In summary, these observations suggest that the relaxase-*oriT* interaction in MOB_H relaxases differs from other relaxases. Recently, DNA substrates that contained a *nic*-site moved to the 5′ end of the IR were efficiently cleaved by HUH-relaxases (88). However, conjugation efficiencies were strongly reduced for these substrates, most likely because the re-ligation of the substrate was shifted towards the covalently-attached state (88). In some relaxases only the IR proximal arm sequence was required for cleavage and it has been suggested that the distal arm sequence serves to form a dsDNA substrate which is required for the re-ligation reaction (82,89). Since in TraI the IR is located 3′ of the *nic*-site, such a function in re-ligation appears unlikely. Indeed, neither does TraI form a stable covalent intermediate with the DNA nor seems a re-ligation reaction be required for the secreted ssDNA. Therefore, the location of the *nic*-site relative to the IR might have a different function in TraI of *N. gonorrhoeae* compared to other relaxases. A possible function might be the protection of the secreted ssDNA from nuclease degradation, which is in agreement with the findings that the 3′ cleavage product could not be degraded by the ssDNA nuclease RecJ (Supplementary Figure S6). Protection of the secreted ssDNA might be important, since the secreted ssDNA of *N. gonorrhoeae* promotes the initial stages of biofilm formation and serves to transform highly naturally competent *Neisseria* species (6,10).

TraI requires Mn^{2+} or Co^{2+} for cleavage

Both relaxases and HD domain phosphodiesterases require divalent metal ions for their activity. Cleavage of classical HUH relaxases is supported by various divalent metal ions, the most prevalent being Mn^{2+} and Mg^{2+}, Zn^{2+} and Ni^{2+} (24). HD domain proteins have also been found with a variety of divalent metal ions (41,44,74,75,90,91). Cleavage activity of TraI was supported by Mn^{2+} and Co^{2+}, whereas no activity was observed with Mg^{2+}, Zn^{2+}, Ca^{2+}, Cu^{2+} and Ni^{2+}. Similar results were found for the HUH-replication protein RepB which cleaved ssDNA in the presence of Mn^{2+} and Co^{2+}. RepB was highly active at concentrations of 20 mM Mn^{2+} and Co^{2+}, respectively (92). However, the cleavage activity of TraI with increasing concentrations of Mn^{2+} and Co^{2+}, revealed that cleavage activity is highest at concentrations of 2.5–10 mM Mn^{2+} while Co^{2+}-dependent cleavage activity was highest at concentrations of around 500 nM Co^{2+}. Higher Co^{2+} concentrations resulted in aggregation of TraI. Such a negative effect on protein stability has also been observed for other relaxases. For instance the MOB_F relaxases $TraI_{pCU1}$ loses cleavage activity at higher concentrations of Cu^{2+} and Zn^{2+}, probably due to the preferred tetrahedral arrangement of these ions (78). Whether Mn^{2+} or Co^{2+} is the physiological relevant metal for TraI is unclear. However, Mn^{2+} is an important ion for *N. gonorrhoeae* survival. *N. gonorrhoeae* accumulates Mn^{2+} via the MntABC transporter to protect itself from oxidative stress

(93,94). Moreover, Mn^{2+} was also found to be linked with regulation of certain virulence factors (95).

TraI binds at least two metal ions in its active site

To identify residues involved in metal coordination, mutations were introduced in the majority of the conserved amino acids of the 3H alternative motif and the HD domain motif. The respective mutants were analyzed for their ability to bind Mn^{2+} and to cleave *oriT*-ssDNA. In DSF assays only D162 and D267, which belong to the HD domain, were affected in Mn^{2+} binding. However, also mutants that were still stabilized in the presence of Mn^{2+} revealed strongly reduced or abolished cleavage activity. This showed that both, the 3H-alternative motif and the HD domain motif, are essential for cleavage activity. Previously, residues belonging to these motifs (H106, H108 as well as H161 and D162) were found to be not essential for DNA-secretion *in vivo* (18). These differences might be caused by the different conditions under which the activity assays were performed. *In vivo*, TraI is present in much lower concentrations than in the *in vitro* assays and possible accessory proteins are present that might influence the cleavage reaction. Furthermore, the *in vitro* assays were performed with TraI variants lacking the N-terminal hydrophobic domain. However, considering the drastic effects on cleavage activity and especially the essential role of D162 in metal binding, it is difficult to envision how these residues would not be required for DNA-secretion.

Based on the structural similarity with related HD-phosphohydrolases and apparent conservation of the active site, an active site containing more than one metal ion seemed likely. Indeed, the half-maximal concentration of Mn^{2+} and Co^{2+} observed in DSF assays was 10 – 20-fold lower than the half-maximal metal ion concentration required for the cleavage reaction. In the HD domain Cas-3 protein from *Thermus thermophilus*, a two metal ion mechanism has been proposed based on structural comparison with other available HD domain crystal structures (75). Similar to TraI, all HD domain residues were required for activity but only mutation of D70 resulted in strongly reduced metal binding as determined by DSF assays (75). For the human DNA polymerase, it was shown that two aspartates coordinate two metal ions, but only at higher metal ion concentrations and in the presence of a substrate, a third metal ion appears that stabilizes the substrate without being directly coordinated by the enzyme's active site residues (96). Based on our data, we propose that TraI contains two metal ions in its active site. The first metal ion binds at a high-affinity binding site, which leads to a structural rearrangement within the protein as observed by DSF assays. For cleavage activity, a second metal ion binds at a lower-affinity binding site.

TraI does not form a stable covalent intermediate with the DNA

HUH-Endonucleases perform a transesterification reaction which includes formation of a covalent bond between the catalytic residue and the cognate DNA (25,26). The covalent bond conserves the energy for the reverse reaction that serves to rejoin the cleaved DNA. Moreover, the covalent bond enables relaxases to be transferred together with the DNA into the recipient cell in a process which involves unfolding of the relaxase (36). An exception is the relaxase CloDF13 of the MOB_C relaxase family, where a free 5′ end of the cleavage product has been reported (68). Relaxases of the MOB_C family contain a DEE motif for metal ion coordination. Moreover, they bear structural similarities with the BamHI-restriction enzyme as shown for the MOB_C-relaxase TraX (38). How conjugation would occur without a covalent relaxase-DNA intermediate is difficult to envision. In TraX, Y181 was found to be essential for pAD1 conjugal transfer and was therefore suggested to form the covalent intermediate with the DNA (38). Previously, the secreted ssDNA of the GGI-encoded T4SS was found protected from degradation by the ssDNA specific 5′ to 3′ exonuclease RecJ in nuclease-protection assays, which led to the assumption that TraI formed a covalent intermediate with the secreted DNA (18). However, we did not observe a stable covalent TraI–DNA intermediate. Furthermore, determination of cleavage rates showed a multiple turn-over reaction. Here, we found a possible explanation for this discrepancy. RecJ requires a 7 nucleotides long ssDNA 5′ tail to properly initiate its exonuclease activity (71). Cleavage of the *oriT*-ssDNA leads to a 3′ cleavage product that contains the sequence that forms the IR-hairpin structure only 3 nucleotides close to the 5′ end. Indeed, the 3′ *oriT-ssDNA* cleavage product containing the 5′ hairpin could not be degraded by rRecJ.

TraI cleavage involves a hydrolysis reaction

The absence of a stable covalent intermediate is unexpected for a relaxase but common for a HD domain phosphohydrolase of which the principal reaction mechanism is hydrolysis (41,44,90). The majority of nucleases cleave DNA leaving a phosphate on the 5′ end and a free 3′ OH (70). Indeed, the 5′ end of the cleavage product of TraI was protected from 5′ PNK-labeling. An interesting mechanistic question is whether the hydrolysis occurs or in a two-step reaction, where first a TraI-DNA covalent intermediate is formed by a transesterification reaction followed by a hydrolysis reaction in the second step or in a one-step reaction by direct hydrolysis. The former has previously been observed, e.g. for the restriction enzyme BfiI. BfiI first forms a transient covalent intermediate with DNA via a histidine-nucleophile, which is then released by a hydrolysis reaction within the same active site (97). While the Y212A mutant of TraI was found inactive in cleavage, the Y212F and Y212H mutants showed residual activity making Y212 as a nucleophile very unlikely. Due to the observed similarities of the conserved metal binding residues of TraI and HD phosphohydrolases and the absence of an apparent nucleophilic amino acid, we propose that TraI cleaves the *oriT*-ssDNA by hydrolysis. Similar to HD phosphohydrolases containing a dinuclear metal center, we propose that the bridging water molecule of the dinuclear center acts as nucleophile (Figure 7). The proposed mechanism does not involve the residues of the alternative 3H motif or of Y212, as, despite their essential role for the cleavage reaction, a precise role of these residues could not be determined. Whether they play a

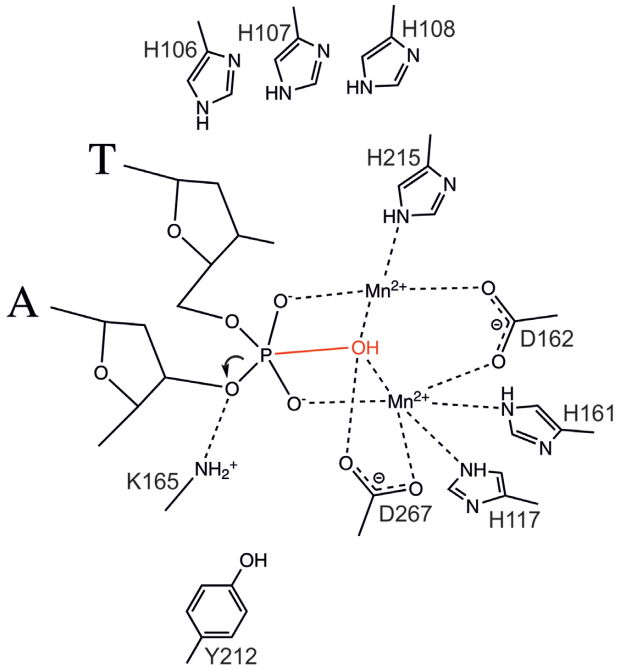

Figure 7. Proposed mechanism for a one-step hydrolysis reaction. The model shows a dinuclear metal center formed by residues of the HD domain and how the nucleophilic water molecule in the dinuclear metal center (shown in red) attacks the phosphate backbone of the ssDNA. The histidines of the alternative 3H motif and Y212 are important for cleavage; however, the precise role of these residues remains unknown.

role in the cleavage reaction or whether they fulfil a different function, like interaction with the DNA substrate, remains a question of future research. Finally, to conclusively show whether the cleavage reaction occurs with a one-step or in a two-step mechanism, the stereochemistry of the cleavage reaction might be studied. A chiral phosphorothioate bond is introduced at the *nic*-site, which enables to distinguish whether one reaction (hydrolysis) or two reactions (transesterification followed by hydrolysis) occur at the phosphorothioate bond (98). Unfortunately, such experiments were hampered by the difficulty to synthesize an oligonucleotide with a chiral phosphorothioate bond, because of both, the length of the minimal *oriT* (50 nucleotides) and its highly stable hairpin structure close to the *nic*-site.

The process of conjugation is difficult to envision without a covalent intermediate between the relaxase and the DNA. However, TraI is involved in secretion of ssDNA which is taken up by recipient cells via type IV pili and subsequently integrated in the genome by homologous recombination (9). For this mechanism, secretion without formation of a covalent bond might even be advantageous. However, most other relaxases of the MOB_H family are most likely transferred via conjugation (22). Whether these relaxases form a covalent intermediate with the DNA should be a topic of further study.

DATA AVAILABILITY

The structural data of the SAXS experiments is deposited on https:\\www.sasbdb.org.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Lotte Søgaard-Andersen, Department of Ecophysiology, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany, and Sonja-Verena Albers, Molecular Biology of Archaea, Institute for Biology II, Microbiology, Albert Ludwig University Freiburg, Freiburg where these experiments were initiated and have been finished, respectively, for their support. We thank Samta Jain and Jürgen Koch for their support in the initial stages of this project and Olivia Spitz for performing the SEC-MALS experiment. We thank Joseph P. Dillard and Alexander Wagner for helpful discussions. We also thank the ESRF beamline BM29 assistant for their help during SAXS data collection and the people of the European Molecular Biology Laboratory (ESRF Grenoble) for their support.

FUNDING

The Center of Structural Studies is funded by the Deutsche Forschungsgemeinschaft [417919780]. The article processing charge was funded by the German Research Foundation (DFG) and the University of Freiburg in the funding programme Open Access Publishing.
Conflict of interest statement. None declared.

REFERENCES

1. Soucy,S.M., Huang,J. and Gogarten,J.P. (2015) Horizontal gene transfer: Building the web of life. *Nat. Rev. Genet.*, **16**, 472–482.
2. Johnson,C.M. and Grossman,A.D. (2015) Integrative and conjugative elements (ICEs): What they do and how they work. *Annu. Rev. Genet.*,**49**, 13.1–13.25.
3. Christie,P.J., Valero,L.G. and Buchrieser,C. (2017) Biological diversity and evolution of type IV Secretion systems. *Curr. Top. Microbiol. Immunol.*, **413**, 1–30.
4. Waksman,G. and Terradot,L. (2017) Structural and molecular biology of type IV secretion systems. *Curr. Top. Microbiol. Immunol.*, **413**, 31–60.
5. Grohmann,E., Keller,W. and Muth,G. (2017) Mechanisms of conjugative transfer and type IV Secretion-Mediated effector transport in Gram-Positive bacteria. *Curr. Top. Microbiol. Immunol.*, **413**, 115–141.
6. Dillard,J.P. and Seifert,H.S. (2001) A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. *Mol. Microbiol.*, **41**, 263–277.
7. Hamilton,H.L., Domínguez,N.M., Schwartz,K.J., Hackett,K.T. and Dillard,J.P. (2005) Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system. *Mol. Microbiol.*, **55**, 1704–1721.
8. Hamilton,H.L. and Dillard,J.P. (2006) Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Mol. Microbiol.*, **59**, 376–385.
9. Callaghan,M.M., Heilers,J.-H., van der Does,C. and Dillard,J.P. (2017) Secretion of chromosomal DNA by the neisseria gonorrhoeae Type IV secretion system. *Curr. Top. Microbiol. Immunol.*, **413**, 323–345.
10. Zweig,M., Schork,S., Koerdt,A., Siewering,K., Sternberg,C., Thormann,K., Albers,S.-V., Molin,S. and van der Does,C. (2014) Secreted single-stranded DNA is involved in the initial phase of biofilm formation by neisseria gonorrhoeae. *Environ. Microbiol.*, **16**, 1040–1052.
11. Snyder,L.A.S., Jarvis,S.A. and Saunders,N.J. (2005) Complete and variant forms of the 'gonococcal genetic island' in Neisseria meningitidis. *Microbiology*, **151**, 4005–4013.

12. Woodhams,K.L., Benet,Z.L., Blonsky,S.E., Hackett,K.T. and Dillard,J.P. (2012) Prevalence and detailed mapping of the gonococcal genetic island in Neisseria meningitidis. *J. Bacteriol.*, **194**, 2275–2285.
13. Spencer-Smith,R., Roberts,S., Gurung,N. and Snyder,L.A.S. (2016) DNA uptake sequences in Neisseria gonorrhoeae as intrinsic transcriptional terminators and markers of horizontal gene transfer. *Microb. Genomics*, **2**, 1–11.
14. Castillo,F., Benmohamed,A. and Szatmari,G. (2017) Xer site specific recombination: double and single recombinase systems. *Front. Microbiol.*, **8**, 1–18.
15. Domínguez,N.M., Hackett,K.T. and Dillard,J.P. (2011) XerCD-mediated site-specific recombination leads to loss of the 57-kilobase gonococcal genetic Island. *J. Bacteriol.*, **193**, 377–388.
16. Remmele,C.W., Xian,Y., Albrecht,M., Faulstich,M., Fraunholz,M., Heinrichs,E., Dittrich,M.T., Müller,T., Reinhardt,R. and Rudel,T. (2014) Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic Acids Res.*, **42**, 10579–10595.
17. Pachulec,E., Siewering,K, Bender,T., Heller,E.-M., Salgado-Pabon,W., Schmoller,S.K., Woodhams,K.L., Dillard,J.P. and van der Does,C. (2014) Functional analysis of the gonococcal genetic island of neisseria gonorrhoeae. *PLoS One*, **9**, e109613.
18. Salgado-Pabón,W., Jain,S., Turner,N., van der Does,C. and Dillard,J.P. (2007) A novel relaxase homologue is involved in chromosomal DNA processing for type IV secretion in Neisseria gonorrhoeae. *Mol. Microbiol.*, **66**, 930–947.
19. Salgado-Pabón,W., Du,Y., Hackett,K.T., Lyons,K.M., Arvidson,C.G. and Dillard,J.P. (2010) Increased expression of the type IV secretion system in piliated Neisseria gonorrhoeae variants. *J. Bacteriol.*, **192**, 1912–1920.
20. Ramachandran,G., Miguel-Arribas,A., Abia,D., Singh,P.K., Crespo,I., Gago-Córdoba,C., Hao,J.A., Luque-Ortega,J.R., Alfonso,C., Wu,L.J. *et al.* (2017) Discovery of a new family of relaxases in Firmicutes bacteria. *PLOS Genet.*, **13**, e1006586.
21. Guglielmini,J., Quintais,L., Garcillán-Barcia,M.P., de la Cruz,F. and Rocha,E.P.C. (2011) The repertoire of ice in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS Genet.*, **7**, e1002222.
22. Garcillán-Barcia,M.P., Francia,M.V. and De La Cruz,F. (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.*, **33**, 657–687.
23. Ilangovan,A., Connery,S. and Waksman,G. (2015) Structural biology of the Gram-negative bacterial conjugation systems. *Trends Microbiol.*, **23**, 301–310.
24. Zechner,E.L., Moncalián,G. and de la Cruz,F. (2017) Relaxases and plasmid transfer in Gram-Negative bacteria. *Curr. Top. Microbiol. Immunol.*, **413**, 93–113.
25. Chandler,M., de la Cruz,F, Dyda,F. and Hickman,A.B., Moncalian,G. and Ton-Hoang,B. (2013). Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat. Rev. Microbiol.*, **11**, 525–538.
26. Pluta,R., Boer,D.R., Lorenzo-Díaz,F., Russi,S., Gómez,H., Fernández-López,C., Pérez-Luque,R., Orozco,M., Espinosa,M. and Coll,M. (2017) Structural basis of a histidine-DNA nicking/joining mechanism for gene transfer and promiscuous spread of antibiotic resistance. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E6526–E6535.
27. Boer,R., Russi,S., Guasch,A, Lucas,M., Blanco,A.G., Pérez-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.
28. Larkin,C., Haft,R.J.F.F., Harley,M.J., Traxler,B. and Schildbach,J.F. (2007) Roles of active site residues and the HUH motif of the F plasmid tral relaxase. *J. Biol. Chem.*, **282**, 33707–33713.
29. Llosa,M., Gomis-Rüth,F.X., Coll,M. and De la Cruz,F. (2002) Bacterial conjugation: a two-step mechanism. *Mol. Microbiol.*, **45**, 1–8.
30. Cascales,E. and Christie,P.J. (2004) Definition of a bacterial Type IV secretion pathway for a DNA substrate. *Science*, **304**, 1–9.
31. Dostál,L., Shao,S. and Schildbach,J.F. (2011) Tracking F plasmid TraI relaxase processing reactions provides insight into F plasmid transfer. *Nucleic Acids Res.*, **39**, 2658–2670.
32. Christie,P.J., Whitaker,N. and González-Rivera,C. (2014) Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta*, **1843**, 1578–1591.
33. Cabezón,E., Ripoll-Rozada,J., Peña,A., de la Cruz,F., Arechaga,I., Cruz,F., De Cabezón,E., Ripoll-Rozada,J., Peña,A., de la Cruz,F. *et al.* (2014) Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.*, **39**, 81–95.
34. Low,H.H., Gubellini,F., Rivera-Calzada,A., Braun,N., Connery,S., Dujeancourt,A., Lu,F., Redzej,A., Fronzes,R., Orlova,E. V *et al.* (2014) Structure of a type IV secretion system. *Nature*, **508**, 550–553.
35. Redzej,A., Ukleja,M., Connery,S., Trokter,M., Felisberto-Rodrigues,C., Cryar,A., Thalassinos,K., Hayward,R.D., Orlova,E.V. and Waksman,G. (2017) Structure of a VirD4 coupling protein bound to a VirB type IV secretion machinery. *EMBO J.*, **36**, 3080–3095.
36. Trokter,M. and Waksman,G. (2018) Translocation through the conjugative Type 4 secretion system requires unfolding of its protein substrate. *J. Bacteriol.*, **200**, 23.
37. Ilangovan,A., Kay,C.W.M., Roier,S., El Mkami,H., Salvadori,E., Zechner,E.L., Zanetti,G. and Waksman,G. (2017) Cryo-EM structure of a relaxase reveals the molecular basis of DNA unwinding during bacterial conjugation. *Cell*, **169**, 708–721.
38. Francia,M.V., Clewell,D.B., de la Cruz,F. and Moncalián,G. (2013) Catalytic domain of plasmid pAD1 relaxase TraX defines a group of relaxases related to restriction endonucleases. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 13606–13611.
39. Soler,N., Robert,E., Beauchêne,I.C. De, Monteiro,P., Libante,V., Maigret,B., Staub,J., Ritchie,D.W., Guédon,G., Payot,S. *et al.* (2019) Characterization of a relaxase belonging to the MOB T family, a widespread family in Firmicutes mediating the transfer of ICEs. *Mob DNA*, **10**, 18.
40. Aravind,L. and Koonin,E. V. (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.*, **23**, 469–472.
41. Zimmerman,M.D., Proudfoot,M., Yakunin,A. and Minor,W. (2008) Structural insight into the mechanism of substrate specificity and catalytic activity of an HD domain phosphohydrolase: the 5′-deoxyribonucleotidase YfbR from Escherichia coli. *J. Mol. Biol.*, **378**, 215–226.
42. Lovering,A.L., Capeness,M.J., Lambert,C., Hobley,L. and Sockett,R.E. (2011) The structure of an unconventional HD-GYP protein from bdellovibrio reveals the roles of conserved residues in this class of cyclic-di-GMP phosphodiesterases. *MBio.*, **2**, 1–8.
43. Galperin,M.Y. and Koonin,E. V. (2012) Divergence and convergence in enzyme evolution. *J. Biol. Chem.*, **287**, 21–28.
44. Bridwell-Rabb,J., Kang,G., Zhong,A., Liu,H. and Drennan,C.L. (2016) An HD domain phosphohydrolase active site tailored for oxetanocin-A biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 201613610.
45. Sherburne,C.K., Lawley,T.D., Gilmour,M.W., Blattner,F.R., Burland,V., Grotbeck,E., Rose,D.J. and Taylor,D.E. (2000) The complete DNA sequence and analysis of R27, a large IncHI plasmid from Salmonella typhi that is temperature sensitive for transfer. *Nucleic Acids Res.*, **28**, 2177–2186.
46. Beaber,J.W., Hochhut,B. and Waldor,M.K. (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from vibrio cholerae. *J. Bacteriol.*, **184**, 4259–4269.
47. Flannery,E.L., Antczak,S.M. and Mobley,H.L.T. (2011) Self-Transmissibility of the integrative and conjugative element ICE Pm1 between clinical isolates requires a functional integrase, relaxase, and Type IV secretion system. *J. Bacteriol.*, **193**, 4104–4112.
48. Vanga,B.R., Ramakrishnan,P., Butler,R.C., Toth,I.K., Ronson,C.W., Jacobs,J.M.E. and Pitman,A.R. (2015) Mobilization of horizontally acquired island 2 is induced in planta in the phytopathogen Pectobacterium atrosepticum SCRI1043 and involves the putative relaxase ECA0613 and quorum sensing. *Environ. Microbiol.*, **17**, 4730–4744.
49. Kelley,L.A., Mezulis,S., Yates,C.M., Wass,M.N. and Sternberg,M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
50. Yang,J., Yan,R., Roy,A., Xu,D., Poisson,J. and Zhang,Y. (2014) The I-TASSER suite: protein structure and function prediction. *Nat. Methods*, **12**, 7–8.
51. Wang,H., Wang,S., Peng,J., Wang,Z., Lu,H. and Xu,J. (2016) Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.*, **7**, 1511–1522.

52. Geertsma,E.R. and Dutzler,R. (2011) A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry*, **50**, 3272–3278.
53. Pernot,P, (2010) New beamline dedicated to solution scattering from biological macromolecules at the ESRF. *J. Phys. Conf. Ser.*, **247**, 012009.
54. Pernot,P., Round,A., Barrett,R., Antolinos,D.M., Gobbo,A., Gordon,E., Huet,J., Mattenet,M., Morawe,C., Mueller-dieckmann,C. *et al.* (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.*, **20**, 660–664.
55. Petoukhov,M., Franke,D., Shkumatov,A. V., Tria,G., Kikhney,A.G., Gajda,M., Gorba,C., Mertens,H.D.T., Petr,V. and Svergun,D.I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.*, **45**, 342–350.
56. Franke,D., Petoukhov,M.V., Konarev,P.V., Panjkovich,A., Tuukkanen,A., Mertens,H.D.T., Kikhney,A.G., Hajizadeh,N.R., Franklin,J.M., Jeffries,C.M. *et al.* (2017) computer programs ATSAS 2. 8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.*, **50**, 1212–1225.
57. Konarev,P.V, Volkov,V.V, Sokolova,A.V, Koch,H.J. and Svergun,D.I. (2003) PRIMUS: a Windows PC-based system for small- angle scattering data analysis. *J. Appl. Crystallogr.*, **36**, 1277–1282.
58. Panjkovich,A. and Svergun,D.I. (2018) Structural bioinformatics CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics*, **34**, 1944–1946.
59. Guinier,A. (1939) Diffraction of x-rays of very small angles-application to the study of ultramicroscopic phenomenon. *Ann. Phys. (Paris).*, **12**, 161–237.
60. Svergun,D.I. (1992) Determination of the regnlarization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
61. Svergun,D.I., Petoukhov,M.V. and Koch,M.H.J. (2001) Determination of domain structure of proteins from X-Ray solution scattering, *Biophys J.* **80**, 2946–2953.
62. Kozin,M.B. and Svergun,D.I. (2001) Automated matching of high- and low-resolution structural models. *Appl. Crystallogr.*, **34**, 33–41.
63. Fischer,H., Neto,M.D.O., Napolitano,H.B., Polikarpov,I. and Craievich,A.F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale research papers. *J. Appl. Crystallogr.*, **43**, 101–109.
64. Rambo,R.P. and Tainer,J.A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature*, **496**, 477–481.
65. Summer,H., Grämer,R. and Dröge,P. (2009) Denaturing urea polyacrylamide gel electrophoresis (Urea PAGE). *J. Vis. Exp.*, **32**, 1485.
66. De La Cruz,F., Frost,L.S., Meyer,R.J. and Zechner,E.L. (2010) Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol. Rev.*, **34**, 18–40.
67. Larkin,C., Datta,S., Harley,M.J., Anderson,B.J., Ebie,A., Hargreaves,V. and Schildbach,J.F. (2005) Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure*, **13**, 1533–1544.
68. Núñez,B. and De la Cruz,F. (2001) Two atypical mobilization proteins are involved in plasmid CloDF13 relaxation. *Mol. Microbiol.*, **39**, 1088–1099.
69. Pansegrau,W., Schoumachert,F., Hohnt,B. and Lankat,E. (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of Agrobacterium tumefaciens Ti plasmids: analogy to bacterial conjugation. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11538–11542.
70. Yang,W. (2011) Nucleases: diversity of structure, function and mechanism. *Q Rev Biophys.*, **44**, 1–93.
71. Han,E.S., Cooper,D.L., Persky,N.S., Sutera,V.A., Whitaker,R.D., Montello,M.L. and Lovett,S.T. (2006) RecJ exonuclease: substrates, products and interaction with SSB. *Nucleic Acids Res.*, **34**, 1084–1091.
72. Edwards,J.S., Betts,L., Frazier,M.L., Pollet,R.M., Kwong,S.M., Walton,W.G., Ballentine,W.K., Huang,J.J., Habibi,S., Del Campo,M. *et al.* (2013) Molecular basis of antibiotic multiresistance transfer in Staphylococcus aureus. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 2804–2809.
73. Datta,S., Larkin,C. and Schildbach,J.F. (2003) Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure*, **11**, 1369–1379.
74. Bellini,D., Caly,D.L., Mccarthy,Y., Bumann,M., An,S.Q., Dow,J.M., Ryan,R.P. and Walsh,M.A. (2014) Crystal structure of an HD-GYP domain cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron centre. *Mol. Microbiol.*, **91**, 26–38.
75. Mulepati,S. and Bailey,S. (2011) Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J. Biol. Chem.*, **286**, 31896–31903.
76. de Anotnio,C., Farías,M.E., de Lacoba,M.G. and Espinosa,M. (2004) Features of the Plasmid pMV158-encoded MobM, a Protein Involved in its mobilization. *J. Mol. Biol.*, **335**, 733–743.
77. Kopec,J., Bergmann,A., Fritz,G., Grohmann,E. and Keller,W. (2005) TraA and its N-terminal relaxase domain of the Gram-positive plasmid pIP501 show specific oriT binding and behave as dimers in solution. *Biochem. J.*, **387**, 401–409.
78. Nash,R.P., Habibi,S., Cheng,Y., Lujan,S.A. and Redinbo,M.R. (2010) The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res.*, **38**, 5929–5943.
79. Guasch,A., Lucas,M., Moncalián,G., Cabezas,M., Pérez-Luque,R., Gomis-Rüth,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.
80. 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 A resolution. *J. Mol. Biol.*, **366**, 165–178.
81. Guogas,L.M., Kennedy,S.A., Lee,J.H. and Redinbo,M.R. (2009) A novel fold in the TraI Relaxase-Helicase C-terminal domain is essential for conjugative DNA transfer. *J. Mol. Biol.*, **386**, 554–568.
82. Carballeira,J.D., González-Pérez,B., Moncalián,G. and de la Cruz,F. (2014) A high security double lock and key mechanism in HUH relaxases controls oriT-processing for plasmid conjugation. *Nucleic Acids Res.*, **42**, 10632–10643.
83. Cheng,Y., McNamara,D.E., Miley,M.J., Nash,R.P. and Redinbo,M.R. (2011) Functional characterization of the multidomain F plasmid TraI relaxase-helicase. *J. Biol. Chem.*, **286**, 12670–12682.
84. Zrimec,J. and Lapanje,A. (2018) DNA structure at the plasmid origin-of-transfer indicates its potential transfer range. *Sci. Rep.*, **8**, 1–10.
85. Kishida,K., Inoue,K., Ohtsubo,Y., Nagata,Y. and Tsuda,M. (2016) Host range of the conjugative transfer system of IncP-9 naphthalene-catabolic plasmid NAH7 and characterization of its orit region and relaxase. *Appl. Environ. Microbiol.*, **83**, e02359-16.
86. Hekman,K., Guja,K., Larkin,C. and Schildbach,J.F. (2008) An intrastrand three-DNA-base interaction is a key specificity determinant of F transfer initiation and of F TraI relaxase DNA recognition and cleavage. *Nucleic Acids Res.*, **36**, 4565–4572.
87. Lucas,M., González-Pérez,B., Cabezas,M., Moncalian,G., Rivas,G. and De La Cruz,F. (2010) Relaxase DNA binding and cleavage are two distinguishable steps in conjugative DNA processing that involve different sequence elements of the nic site. *J. Biol. Chem.*, **285**, 8918–8926.
88. Sagredo,S., de la Cruz,F. and Moncalián,G. (2016) Design of novel relaxase substrates based on rolling circle replicases for bioconjugation to DNA nanostructures. *PLoS One*, **11**, e0152666.
89. Parker,C., Becker,E., Zhang,X., Jandle,S. and Meyer,R. (2005) Elements in the co-evolution of relaxases and their origins of transfer. *Plasmid*, **53**, 113–118.
90. Jeon,Y.J., Park,S.C., Song,W.S., Kim,O.H., Oh,B.C. and Yoon,S.I.L. (2016) Structural and biochemical characterization of bacterial YpgQ protein reveals a metal-dependent nucleotide pyrophosphohydrolase. *J. Struct. Biol.*, **195**, 113–122.
91. Munir,A. and Shuman,S. (2017) Characterization of Runella slithyformis polynucleotide kinase domain. *J. Bacteriol.*, **199**, e00739-1.
92. Boer,D.R., Ruíz-Masó,J.A., López-Blanco,J.R., Blanco,A.G., Vives-llcer,M., Chacón,P., Usón,I., Gomis-Rüth,F.X., Espinosa,M., Llorca,O. *et al.* (2009) Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains. *EMBO J.*, **28**, 1666–1678.
93. Tseng,H.J., Srikhanta,Y., McEwan,A.G. and Jennings,M.P. (2001) Accumulation of manganese in Neisseria gonorrhoeae correlates with

resistance to oxidative killing by superoxide anion and is independent of superoxide dismutase activity. *Mol. Microbiol.*, **40**, 1175–1186.
94. Seib,K.L., Tseng,H., McEwan,A.G., Apicella,M.A. and Jennings,M.P. (2004) Defenses against oxidative stress in neisseria gonorrhoeae and Neisseria meningitidis: Distinctive systems for different lifestyles. *J. Infect. Dis.*, **190**, 136–147.
95. Wu,H.J., Seib,K.L., Srikhanta,Y.N., Edwards,J., Kidd,S.P., Maguire,T.L., Hamilton,A., Pan,K.-T.T., Hsiao,H.-H.H., Yao,C.-W.W. *et al.* (2010) Manganese regulation of virulence factors and oxidative stress resistance in Neisseria gonorrhoeae. *J. Proteomics*, **73**, 899–916.
96. Gao,Y. and Yang,W. (2016) Capture of a third Mg 2+ is essential for catalyzing DNA synthesis. *Science*, **352**, 1334–1338.
97. Sasnauskas,G., Connolly,B.A., Halford,S.E. and Siksnys,V. (2007) Site-specific DNA transesterification catalyzed by a restriction enzyme. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 2115–2120.
98. Mizuuchi,K., Nobbs,T.J., Halford,S.E., Adzuma,K. and Jun,Q. (1999) A new method for determining the stereochemistry of DNA cleavage reactions: Application to the SfiI and HpaII restriction endonucleases and to the MuA transposase. *Biochemistry*, **38**, 4640–4648.
99. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.