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

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

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

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

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

Abstract

Type 4 Secretion Systems (T4SSs) are large and versatile protein machineries that facilitate the spread of antibiotic resistance and other virulence factors via horizontal gene transfer. Conjugative T4SSs depend on relaxases to process the DNA in preparation for transport. TraI from the well-studied conjugative plasmid pKM101 is one such relaxase. Here, we report the crystal structure of the trans-esterase domain of TraI in complex with its substrate *oriT* DNA, highlighting the conserved DNA binding mechanism of conjugative relaxases. Additionally, we present an apo structure of the trans-esterase domain of TraI that includes most of the flexible thumb region. This allows us for the first time to visualize the large conformational change of the thumb domain upon DNA binding. We also characterize the DNA binding, nicking and religation activity of the trans-esterase domain, helicase domain and full-length TraI. Unlike previous indications in the literature, our results reveal that the TraI trans-esterase domain from pKM101 behaves in a conserved manner with its homologs from the R388 and F plasmids.

Introduction

Antibiotic resistance is one of the most pressing health challenges in today's world. One of the main drivers of this world-wide problem is the ability of pathogens to spread resistance genes using horizontal gene transfer. This is a process that is often facilitated by conjugative Type 4 Secretion Systems (T4SSs). These are highly versatile systems that transport DNA and proteins from a bacterial donor cell into a recipient cell. In recent years, researchers have generated structural and functional insight into these systems, especially from Gram-negative (G-) bacteria. A few recent examples are the T4SSs from the R388 plasmid and the pKM101 system[1,2]. For an overview of the current structural knowledge of T4SSs, please see the recent review of Costa *et al*[3]. The focus on this work is the T4SS of pKM101, which belongs to the class of minimal T4SS that consist of twelve proteins homologous to the paradigmatic VirB/VirD4 T4SS from *Agrobacterium tumefaciens*.

For all known T4SSs that transfer conjugative plasmids, the plasmid DNA must be processed before it can be transported. This is done via the relaxosome complex, which consists of a relaxase, accessory factor protein(s) and the DNA[4]. To form this complex, one or several accessory proteins bind to the origin of transfer (*oriT*) on the DNA and locally melt the double stranded DNA to promote relaxase binding to a defined sequence, often forming a hairpin, close to the nicking site[5]. The relaxase binds this single stranded *oriT* DNA via its N-terminal trans-esterase domain[6]. This domain reacts with the DNA via a transesterification reaction at the specific *nic*-site, which generates the transfer intermediate consisting of the relaxase covalently bound to the 5' end of the cleaved transfer strand (T-strand)[7]. Many relaxases have a second functional domain in the more variable C-terminal part of the enzyme. This is often a helicase domain, which unwinds the DNA to allow for transport of the single stranded transfer-strand DNA[6]. The relaxase-transfer-strand complex (T-complex) is recruited to the T4SS by the type IV coupling protein (T4CP) and is transported through the T4SS channel into the recipient cell[8]. Once present in the recipient cell, the trans-esterase domain religates the DNA to regenerate the circularized plasmid[9].

Relaxases have been phylogenetically classified into eight MOB families: MOB_F, MOB_H, MOB_Q, MOB_C, MOB_P, MOB_V, MOB_T and MOB_B. Of these, pKM101-encoded TraI ($TraI_{pKM101}$) belongs to the MOB_{F11} subclade together with its closest relative TrwC from plasmid R388 ($TrwC_{R388}$) and the identical TraI from the sister-plasmid pCU1 ($TraI_{pCU1}$)[6,10,11].

While structural information is available for the trans-esterase domains of several MOB_F relaxases[5,12] structural data for the substrate DNA-bound state is only available for $TrwC_{R388}$ and TraI from the F-plasmid ($TraI_F$), which belongs to subclade MOB_{F12}[13,14]. Furthermore, there is only very limited data on full-length relaxases, again mostly from $TraI_F$. $TraI_F$ consists of one trans-esterase domain, a vestigial helicase domain and an active helicase domain (Fig. 1A) and binds *oriT* DNA as a heterogenous dimer. One $TraI_F$ monomer adapts an open conformation and binds *oriT* at a hairpin positioned 5' of the *nic*-site with the trans-esterase domain, while a 2nd $TraI_F$ binds ssDNA with the helicase domains in a closed conformation. These two states are incompatible with each other in a single monomer[15]. A cryo-EM structure of the helicase-bound state is available and shows how the single stranded DNA is almost entirely surrounded by the helicase domains. This structure is of full-length $TraI_F$ and visualizes all of the protein except the very flexible C-terminal domain[15]. It is not clear to what extent this structure and mode of action apply to $TraI_{pKM101}$, as it is much shorter and completely lacks the vestigial helicase domain (Fig. 1A).

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

Materials and Methods

Plasmids

The full-length *traI* gene, and shorter constructs corresponding to the trans-esterase domain (TraI-TE) (residues 1-299) and the helicase domain (TraI-H) (residues 429-910) were PCR amplified and cloned into p7XC3H and p7XNH3 via the FX cloning system[16] under the T7 promotor, with an N- or C-terminal 10-His-tag and a 3C protease sequence (see Table S1).

Protein expression and purification

All TraI variants were produced in *E. coli* BL21 (DE3). The cells were grown in 1.5 L TB medium at 37°C using a LEX (Large-scale Expression) bioreactor (Epiphyte3). Once the

96 cultures reached an OD_{600} between 1.0 and 1.5, the temperature was reduced to 18°C and
97 expression was induced by adding 0.4 mM IPTG for ~16 hours. Cells were centrifugated at
98 6000 x g and pellets were resuspended in resuspension buffer (50 mM HEPES pH 7.0, 300 mM
99 NaCl, 15 mM Imidazole, 0.2 mM AEBSF, and ca 0.02 mg/mL DNase I). Resuspended cells
100 were lysed in a Cell Disruptor (Constant Systems) at 25 kPsi at 4°C, followed by centrifugation
101 at 30,000 g for 30 min at 4°C. The cell lysate was then incubated for 1h at 4°C with Ni-NTA
102 resin (~2 mL/L culture for TraI and TraI-H or ~4 mL/L culture for TraI-TE) before transfer to
103 a gravity flow column. The resin was subsequently washed with 10 column volumes (CV) of
104 wash buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 50 mM Imidazole, 0.2 mM AEBSF),
105 followed by 10 CV wash buffer with 2 M LiCl and another 10 CV wash buffer. For
106 experiments without tag-cleavage, $TraI_{His}$ was eluted with 5 mL/CV elution buffer (50 mM
107 HEPES pH 7.0, 300 mM NaCl, 500 mM Imidazole).
108 For experiments in which the tag was cleaved off, 3C PreScission protease was added, while
109 the protein was still bound to the IMAC resin, in an estimated ratio of 1:50 to 1:100 and
110 incubated ~16 hours at 4°C. The cleaved protein was then recovered by collecting the flow
111 through.
112 All samples were diluted 3x with dilution buffer (25 mM HEPES pH 7.0, 50 mM NaCl) after
113 Ni-NTA purification and loaded on a HiTraP Heparin HP (5 mL) column, equilibrated with
114 Buffer A (25 mM HEPES pH 7.0, 150 mM NaCl) with a peristaltic pump. After loading the
115 protein, the column was connected to an ÄKTA Pure (Cytiva) and washed with Buffer A for
116 ~2 CV until absorbance at 280 nm was stable. All TraI variants were subsequently eluted via a
117 salt gradient to 100% Buffer B (25 mM HEPES pH 7.0, 1.0 M NaCl) over 80 mL. Peak
118 fractions with pure protein, as judged by a 260 nm / 280 nm ratio of <0.8, were concentrated
119 using Amicon Ultra centrifugal filters with molecular weight cutoffs at 50 kDa (TraI or $TraI_{His}$),
120 30 kDa (TraI-H) or 10 kDa (TraI-TE). Size exclusion chromatography was performed in
121 25 mM HEPES pH 7.0, 300 mM NaCl on a Superdex 200 10/300 GL Increase column and
122 fractions were analyzed on 12% SDS-PAGE stained with Coomassie blue.
123
124 **Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS)**
125 SEC-MALS was performed with 250 µl of protein sample at a minimum concentration of
126 1 mg/mL in 25 mM HEPES pH 7.0, 300 mM NaCl. Experiments were done on a Superdex 200
127 10/300 GL Increase column, or a Superose 6 10/300 GL Increase column for $TraI_{His}$, on an
128 ÄKTA Pure (Cytiva) coupled to a light scattering (Wyatt Treas II) and refractive index (Wyatt
129 Optilab T-Rex) detector. Data was collected and analyzed using Astra software (Wyatt

130 Technology; version 7.2.2) as described by Some *et al.* 2019[17]. The molecular weight of the
131 protein samples was calculated as an average from a minimum of 3 measurements and is
132 reported with the standard deviation.
133
134 **DNA oligomers used for binding assays and crystallization**
135 Single stranded DNA oligomers were purchased from Eurofins Genomics, with and without
136 the fluorescent label fluorescein isothiocyanate (FITC). The position of the label at the 5' or 3'
137 end is indicated in Table S1 and a schematic of the different *oriT* DNA used in this study can
138 be found in Fig. 1B. Stock solutions of 100 µM were dissolved in MilliQ water, but all further
139 dilutions were done in 25 mM HEPES pH 7.0, 300 mM NaCl. Oligomers were heated at 95°C
140 for 5 min and cooled down to room temperature over at least 20 minutes prior to use.
141
142 **Electrophoretic mobility shift assay**
143 Fluorescently labeled DNA (50 nM, Table S1) was incubated for 15 min with 0 to 1600 nM of
144 protein (see results) in 25 mM HEPES pH 7.0, 300 mM NaCl at room temperature before
145 adding 6x native loading dye (3x TBE, 30% glycerol, 0.125% bromophenol blue). The samples
146 were resolved on a native gel (5% polyacrylamide, 0.75x TBE) in 0.75x TBE running buffer
147 at 50 V for 90 min at 4°C. Gels were imaged on an Amersham Typhoon 5 scanner with
148 excitation at 488 nm, using the Cy2 emission filter (515-535 nm).
149
150 **Nicking and religation assays**
151 Nicking and religation assays were performed in 25 mM HEPES pH 7.0, 300 mM NaCl, 5 mM
152 $MgCl_2$. Nicking assays contained 100 nM F-*oriT*57, religation assays (version 1) contained
153 100 nM F-*oriT*57 and 100 nM F-*oriT*20, religation assays (version 2) contained 100 nM F-
154 *oriT*35 and 100 nM non-fluorescent *oriT*57 (Fig. 1B & Table S1). DNA oligomers were
155 incubated with increasing concentrations of TraI in a shaking heatblock at 37°C and 300 rpm,
156 for 1h. The reactions were stopped by adding 2x stop solution (96% formamide, 20 mM EDTA,
157 0.1% bromophenol blue). The samples were boiled for 5 min at 95°C before loading on a
158 denaturing gel (16% polyacrylamide, 7 M Urea, 1x TBE). The gel was run at 50 V for 15 min
159 followed by 100 V for 1h at room temperature and then imaged on an Amersham Typhoon 5
160 scanner with excitation at 488 nm, using the Cy2 emission filter (515-535 nm).
161
162 **Crystallization and structure determination**

163 Protein crystals were obtained with the sitting drop vapor diffusion method at 20°C. Substrate
164 DNA, *oriT*11 (Fig. 1B & Table S1) was added to TraI-TE, to yield a final concentration of
165 9 mg/mL TraI-TE and an equimolar amount of the DNA. This mixture was used for co-
166 crystallization experiments, which yielded both the apo structure of TraI-TE and the DNA-
167 bound structure. The apo structure crystallized in 10% w/v PEG 4000, 20% v/v glycerol,
168 0.03 M $MgCl_2$, 0.03 M $CaCl_2$ and 0.1 M MOPS/HEPES-Na pH 7.5. The DNA-bound structure
169 crystallized in 10% w/v PEG 20000, 20% v/v PEG MME 550, 0.03 M sodium nitrate, 0.03 M
170 disodium hydrogen phosphate, 0.03 M ammonium sulfate and 0.1 M MES/imidazole pH 6.5.
171 Crystals were flash-frozen in liquid nitrogen, without the addition of extra cryo-protectant. X-
172 ray diffraction data was collected on ID30B, ESRF, France. The data was processed using
173 XDS[18,19]. Both crystals had the space group P212121 and contained a single copy of the protein
174 in the asymmetric unit. The phase-problem was solved using molecular replacement with the
175 trans-esterase domain of $TraI_{pCU1}$ (PDB: 3L6T) in PHENIX phaser[20]. The structures were built
176 in Coot[21] and refined at 2.1 Å (DNA-bound structure) and 1.7 Å (apo structure) in PHENIX
177 refine[22], to R_{work}/R_{free} values of 22.5/26.9% and 17.1/19.1% respectively. Further refinement
178 statistics can be found in Table S2. Atomic coordinates and structure factors of both the apo
179 and DNA-bound structure of TraI have been deposited with the Protein Data Bank (PDB: 8A1B
180 and 8A1C).
181

182 ## Results

183 **Purification of TraI and its functional domains**

184 TraI as well as its functional domains, the trans-esterase domain (TraI-TE) and the helicase
185 domain (TraI-H), were expressed in *E. coli* and purified to homogeneity in a three-step process.
186 TraI, TraI-TE and TraI-H, all with their His-tags cleaved off, were shown using SEC-MALS
187 to be monomeric in solution (Fig. S1). It is interesting to note that full-length TraI was initially
188 purified without removing the His-tag ($TraI_{His}$) and was then shown to be in an oligomer -
189 monomer equilibrium, with a molecular weight of the oligomer determined to be 485 +/- 10
190 kDa, which corresponds well to the expected molecular weight of a tetramer (492 kDa) (Fig.
191 S2). Since we only observed oligomerization of $TraI_{His}$, but not for TraI without the His-tag
192 nor TraI-TE and TraI-H, we conclude that TraI is a monomer and that the initially observed
193 oligomerization of the full-length protein was an artifact induced by the His-tag.
194

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

196 To biochemically characterize TraI and its functional domains we performed several activity
197 assays. To estimate the differences in binding affinity and specificity of the different TraI
198 domains we used electrophoretic mobility shift assays (EMSAs). We investigated DNA
199 binding to a fluorescent 22-mer with the post-*nic* DNA sequence from the *oriT* of pKM101
200 (*oriT*22-F), a fluorescent 35-mer of the pre-*nic* DNA containing the hairpin structure (F-
201 *oriT*35) and a fluorescent full-length 57-mer with both the pre- and post-*nic* sequence (F-
202 *oriT*57), as well as randomly sequenced 57- and 35-mers (Fig. 1B & Table S1).
203
204 We found that the TraI trans-esterase domain binds both F-*oriT*57 and F-*oriT*35 DNA with
205 high affinity (apparent K_D is below 200 nM), while this domain only interacts with post-*nic* or
206 random DNA at protein concentrations above 800 or 1600 nM (Fig. 2A & S3A). This indicates
207 that the trans-esterase domain binds sequence specific to the pre-*nic* region. In contrast, the
208 TraI helicase domain shows only low affinity binding to both *oriT* and random DNA and seems
209 to have a stronger interaction with the longer 57-mer than with the 35- and 22-mer DNA
210 constructs (Fig. 2B & S3B). This domain thus appears to have no sequence specificity and to
211 prefer DNA fragments longer than 35 nucleotides.
212
213 As full-length TraI includes both the helicase and the trans-esterase domains, we expected to
214 find a combination of their DNA binding capabilities. TraI was indeed found to bind both *oriT*
215 and random DNA. Interestingly, the random DNA shows a significant mobility shift with TraI
216 protein concentrations of 200 nM and above (Fig. 2C), while 400 nM is needed when the TraI
217 helicase domain is used in isolation. Both F-*oriT*57 and F-*oriT*35, show a complete mobility
218 shift with TraI concentrations of 200 nM and above, compared to 400 nM for the trans-esterase
219 domain. Both observations therefore indicate a significant increase in the apparent affinity of
220 the full-length TraI as compared to its individual domains. At higher TraI concentrations
221 additional supershifts are observed. An additional observation was an initial very high shift that
222 was observed for F-oriT57 and F-oriT35 already with 50 nM TraI (similar to the high shift seen
223 at the highest concentration of TraI-H). However, this high-molecular weight band was only
224 observed up to 400 nM TraI and disappears as the protein concentration is increased further.
225 This result has been consistently observed in mobility shift assays with *oriT* DNA, but not with
226 random DNA (Fig. 2C & S3C).
227
228 We also examined the capacity of TraI to nick and religate its substrate *oriT* DNA. TraI was
229 incubated with the fluorescent full-length F-*oriT*57, containing the recognition hairpin, the *nic*-

230 site and a post-*nic* stretch of 22 bases, or a random DNA sequence of the same length. DNA
231 products were consequently separated on a denaturing polyacrylamide gel. Our results (Fig.
232 4A & C) demonstrate that both TraI and TraI-TE can nick about 50% of the tested *oriT* DNA
233 substrate, while neither nicked the random DNA control under the same conditions. We were
234 unable to detect a higher degree of cleaved F-*oriT*57 DNA and predicted that this was due to
235 the expected religation ability of TraI. To test this, we added a shorter version of the fluorescent
236 pre-*nic* DNA (F-*oriT*20) to the incubation mixture of TraI. In this experiment F-*oriT*57 is
237 expected to be (partly) cleaved by TraI into a fluorescent 35-mer and a non-fluorescent 22-mer
238 that stays covalently bound to TraI. If religation occurs, the 22-mer could either be religated to
239 the fluorescent 35-mer to re-form the original F-*oriT*57 or to the F-*oriT*20 to form a fluorescent
240 42-mer. This 42-mer was indeed seen on the denaturing gel (Fig. 4B & D, left). This nicking
241 and religation activity of TraI was further investigated by incubating the protein with a
242 fluorescent F-*oriT*35 and a non-fluorescent *oriT*57. The appearance of F-*oriT*57 indicates that
243 non-fluorescent *oriT*57 was nicked and that the resulting 22-mer was ligated to F-*oriT*35 (Fig.
244 4B & D, right).
245
246 **Crystal structure of TraI trans-esterase domain in complex with *oriT* DNA**
247 To obtain a mechanistic understanding of the sequence specific DNA binding ability of TraI,
248 purified TraI-TE was co-crystallized with single stranded 11-mer *oriT* DNA (*oriT*11) (Fig. 1B).
249 The crystals grew in space group P212121 and contained one molecule in the asymmetric unit.
250 The crystallographic phase problem was solved with molecular replacement using the
251 previously solved structure of the trans-esterase domain of pCU1 (PDB: 3L6T) as a search
252 model, and the crystal structure was refined at a resolution of 2.1 Å. Virtually all residues were
253 visible in the electron density, apart from five residues in flexible loop regions and at the C-
254 terminus. We were also able to model ten of the eleven bases into the electron density of the
255 bound DNA (Fig. 4).
256
257 The overall architecture of TraI-TE can be described as a hand, with the bound DNA located
258 between the palm and the C-terminal thumb domain (Fig. 4A). The determined structure is
259 similar to the previously described homologs $TrwC_{R388}$, and $TraI_F$ (Fig. 4B)[13,14,23]. A conserved
260 histidine triad coordinates a divalent metal that is required for the trans-esterase reaction in the
261 homologs[7,14,24]. In TraI, this triad is formed by H149, H160 and H162 (Fig. 4C). Electron
262 density was observed in the middle of this triad and refinement against the possible metal ions

263 yielded the best fit for manganese. The conserved aspartate, D84 in $TraI_{pKM101}$, is situated
264 within hydrogen-bonding distance of both H162 and Y18 (2.8 Å and 2.7 Å respectively, Fig.
265 4C) as earlier described for $TrwC_{R388}$ and $TraI_F$[13,14]. It is therefore in position to activate the
266 catalytic tyrosine Y18, the first tyrosine of the conserved YY-$X_{(5-6)}$-YY motif[5,6], which is
267 predicted to occur during the trans-esterification reaction with the scissile phosphate at the *oriT*
268 *nic*-site.
269
270 The DNA substrate used for co-crystallization was *oriT*11, representing the 11 bases directly
271 5' of the *nic*-site. Of these, the first base at the 5' end, T1, did not yield any defined electron
272 density, likely because it was not properly stabilized (Fig. 4D). While the first visible base, C2,
273 does not show any specific contacts with TraI-TE, the following guanine (G3) forms four
274 hydrogen bonds with R80 and D182 using both the Hoogsteen and Watson-Crick edge, and the
275 following thymine (T4), forms two hydrogen bonds with N181. The following two bases, T5
276 and A6, are stacked on top of each other perpendicular to the orientation of the previous three
277 bases and form hydrogen bonds with K188 (T5) and Q256 (A6). These initial bases are oriented
278 in a position that is pointing away from the active site. The following bases place the DNA in
279 a conformation resembling a U-turn that locates the scissile phosphate of T11 in close
280 proximity to the catalytic tyrosine. This bent DNA conformation is stabilized with multiple
281 hydrogen bonds, combined with hydrophobic and pi-pi interactions between the stacked bases
282 G8 and G10. G7 forms two hydrogen bonds with L2 at its Watson-Crick edge and an additional
283 one with K196 from its Hoogsteen edge. G8 forms two Watson-Crick hydrogen bonds with D3
284 and G10 forms two Hoogsten hydrogen bonds with R250. The phosphates of the three most 3'
285 bases are also stabilized with several hydrogen bonds (S236 for T9, S236 and K91 for G10,
286 and R235 and R153 for T11) (Fig. 4D). Taken together, these DNA-DNA and DNA-protein
287 interactions illustrate a binding mechanism that relies both on the secondary structure and the
288 specific sequence of the DNA.
289
290 **Apo structure reveals rearrangement of the thumb domain upon DNA binding**
291 An important feature of DNA binding is the thumb domain, consisting of two α-helices
292 connected by a loop region. This domain lies on top of the 3' part of the DNA and forms
293 hydrogen bonds with the *oriT* DNA via residues R235, R250 and Q256 (Fig. 4A & D). The
294 thumb domain is highly flexible and has previously only been observed in relaxase structures
295 with DNA bound. Unexpectedly, our co-crystallization experiment resulted in an additional

296 crystal with the space group $P2_12_12_1$ that resolved to 1.7 Å. This structure had a different unit
297 cell, which had the fortuitous outcome of supporting an apo state of TraI-TE, in which we were
298 able to build most of the thumb domain in an open conformation (Fig. 5A). The electron density
299 corresponding to residues 261-271 was of insufficient quality to build these residues in the
300 model, but inspection of the density suggests that the α-helix continues towards the base of the
301 thumb. Between the apo- and the DNA-bound state, the thumb moves up to 40 Å and undergoes
302 important structural changes. The observed conformational differences include a
303 reorganization of α-helix 8 and the extension of α-helix 9 (numbered based on the DNA-bound
304 structure as one helix disappears in the apo structure) to transition from the apo to the DNA-
305 bound structure, as illustrated in Fig. 5A and Movie S1. These rearrangements reveal the
306 structural changes that are necessary to first allow DNA to enter the binding site and to
307 subsequently keep the DNA in place, highlighting the importance of the thumb domain in this
308 process. The C-terminal helix at the base of the thumb domain (α-helix 10) interacts with the
309 loop at the end of α-helix 1 via a hydrogen bond between R274 and D25 in both the DNA-
310 bound and apo-structures (Fig. S4A). An additional interaction is made between H278 and
311 Y26, the third tyrosine of the YY-$X_{(5-6)}$-YY motif, but only in the DNA-bound structure.
312
313 Both the apo- and the DNA-bound structure of TraI-TE presented here show differences when
314 compared to the available structure from $TraI_{pCU1}$, even though the domains have an identical
315 primary sequence (Fig. 5B, PDB: 3L6T)[25]. In $TraI_{pCU1}$, the C-terminal part of the protein
316 containing the thumb domain and the base of the thumb are missing and the loop at the end of
317 α-helix 1 is pointing away from the active site, possibly because it's not stabilized by the
318 missing α-helix 10 (Fig. 5B).

319 **Discussion**

320 In this study, we structurally and biochemically characterized the TraI relaxase from the MOB_F
321 family plasmid pKM101 to better understand the diversity of the relaxases in this protein family
322 as they are of key importance to type IV secretion.
323
324 We performed mobility shift assays (Fig. 2 & S3) to estimate the DNA binding affinity of full-
325 length TraI and to compare this to the affinity of the trans-esterase (TraI-TE) and helicase (TraI-
326 H) domains individually. Both full-length TraI and the helicase domain bound DNA in a
327 sequence unspecific manner, as both bound to a similar degree to *oriT* and random DNA (Fig.
328 2 & S3). This is as expected since these helicases are predicted to unwind the entire conjugative
329 plasmid and therefore should not have any sequence specificity. In contrast, the trans-esterase
330 domain of TraI (TraI-TE) showed a higher affinity for *oriT* DNA as compared to random DNA
331 (Fig. 2 & S3). While the mobility shift assays presented here are not suitable for precise K_D
332 determinations, the apparent K_D of the trans-esterase domain of $TraI_{pKM101}$ for *oriT* DNA is <
333 200 nM, while interactions with random DNA were only observed at protein concentrations
334 above 800 nM (Fig. 2A). Nicking and religation assays show that full-length TraI and trans-
335 esterase domain (TraI-TE) are both equally capable of nicking and religating the *oriT* DNA,
336 but not random DNA. This shows that the trans-esterase domain is active on its own and
337 confirms its *oriT* specificity (Fig. 3).
338
339 The observed affinity and specificity of the trans-esterase domain for its cognate *oriT* DNA is
340 a feature that has been described in various $TraI_{pKM101}$ homologs, such as $TrwC_{R388}$ and
341 $TraI_F$[23,26,27]. However, it is contradictory to a previous report on the trans-esterase domain of
342 $TraI_{pCU1}$[25]. pCU1 is an antibiotic resistance plasmid that shares a common ancestor with
343 pKM101 and R46[11]. The pKM101 and pCU1 conjugative plasmids have an identical *oriT*
344 region and encode identical relaxases: $TraI_{pKM101}$ / $TraI_{pCU1}$. In spite of this, the trans-esterase
345 domain of $TraI_{pCU1}$ was previously described to bind DNA in a weak and sequence independent
346 manner, with K_D values for binding to *oriT* and non-*oriT* oligomers ranging from 0.7 to
347 1.2 μM[25]. The DNA binding of $TraI_{pKM101}$ and the relaxase domain of $TraI_{pCU1}$ were measured
348 using different methods and using slightly different protein constructs. There are slight
349 differences in the protein constructs, mainly in the tags used, which could explain the
350 differences observed for the N-terminally tagged TraIpCU1 sine adding additional residues at
351 the N-terminus would likely disturb the active site. However, the differences are consistent also

352 with the C-terminally tagged $TraI_{pCU1}$, so we have no obvious explanation for the observed
353 discrepancies between our experimental results and those of $TraI_{pCU1}$. We do note that what we
354 observe in our study aligns well with what has been shown in homologous proteins and
355 conclude that $TraI_{pKM101}$ binds and processes DNA in a similar fashion as its homologs
356 $TrwC_{R388}$ and $TraI_F$.
357
358 While His-tags sometimes cause dimerization[28,29], it is very uncommon that they promote any
359 higher defined oligomerization. This led us to initially think that TraI formed tetramers in
360 solution (Fig. S2), which turned out to be an artefact of the His-tag (Fig. S1). However, an
361 interesting feature in the EMSAs with full-length TraI, and to some extent its helicase domain,
362 is the appearance of supershifts at higher protein concentrations (Fig. 2 & S3). The exact same
363 data was observed both with His-tagged TraI and TraI with the His-tagged cleaved. This
364 suggests that multiple copies of the protein can bind to the same DNA molecule. $TraI_{pKM101}$
365 was previously reported to self-interact *in vivo*[30] and $TraI_F$ was shown to bind *oriT* as a dimer[15].
366 Thus, we cannot exclude oligomerization of $TraI_{pkM101}$ when binding to DNA.
367
368 To understand the interaction between TraI and its substrate DNA on a mechanistic level, we
369 solved the crystal structure of the trans-esterase domain in complex with its cognate *oriT* DNA.
370 The DNA-bound structure of TraI-TE_{pKM101} (Fig. 4A) shows a protein fold that can be
371 described as a palm with the active site at the center and a thumb domain that is involved in
372 holding the substrate DNA in place. This structure is similar to the previously solved trans-
373 esterase domains of its homologs[13,14,23] (Fig. 4B). The highest structural homology is found
374 with another member of the MOB_{F11} subclade: $TrwC_{R388}$ (PDB: 2CDM, 50% identity, root
375 mean square deviation (RMSD) = 1.7 Å over 276 residues). The structure is also similar to that
376 of $TraI_F$, a MOB_{F12} relaxase (PDB: 2A0I, 37% identity, RMSD = 2.3 Å over 269 residues)[31].
377 The DNA binding site is highly conserved between these proteins with the bound ssDNA
378 making a characteristic U-turn[26]. The electron density for the ion that is coordinated by the
379 histidine triad in the active site could best be modeled by a manganese ion. The ion was likely
380 retained during purification since no metals were added during purification or crystallization
381 of the DNA-bound structure (the apo structure was crystallized in the presence of both Mg^{2+}
382 and Ca^{2+}). Our activity assays were performed in the presence of magnesium. No other divalent
383 metals were tested as TraI homologs were shown to be able to function with various divalent
384 metals[14,25].

385

386 While the DNA-bound structures of the MOB_F family relaxases are very similar, there is more
387 variability between the structures in the absence of DNA. This probably reflects the greater
388 flexibility of the structures in the absence of the ligand. The apo-structure of TraI-TE_{pKM101}
389 presented here (Fig. 5A) shows the thumb domain in an open conformation. Previously, this
390 thumb domain could only be modeled in crystal structures with bound DNA[13,14,23]. The RMSD
391 between our ligand-free and DNA-bound structures is 3.3 Å over 277 residues, which is
392 indicative of the large molecular movement of up to 40 Å (Fig. 5A and Movie S1). It is likely
393 that the open conformation is important for *oriT* DNA to access and bind to the active site. The
394 exact degree of opening observed in our structure is dependent on crystal contacts that
395 stabilizes the thumb domain. In solution, the protein is very likely in an open-closed
396 equilibrium, which is shifted strongly towards many different open states in the absence of
397 cognate DNA.

398

399 There are important differences between the structures of TraI-TE_{pKM101} and the apo structure
400 of TraI-TE_{pCU1} (PDB: 3L6T) despite their 100% sequence identity (Fig. 5B). The overall fold
401 is similar, with a RMSD of 1.4 Å over 208 residues, but the entire C-terminal region of $TraI_{pCU1}$
402 is missing. This includes the thumb domain, and the helix at the base of the thumb (α-helices
403 7-10 of the $TraI_{pKM101}$ DNA-bound structure). More importantly, α-helix 1, which contains the
404 beginning of the conserved YY-$X_{(5-6)}$-YY motive that provides the catalytic tyrosine, mostly
405 appears as a loop in the previously published $TraI_{pCU1}$ structure. As a consequence, it points far
406 away from the active site, resulting in a >10 Å displacement compared to the positions in the
407 homologs[25]. Possibly, the displacement of α-helix 1 is related to the destabilization of the
408 thumb domain of the protein. In contrast, both structures presented here, as well as other
409 relaxase structures ($TrwC_{R388}$ PDB: 1S6M, $TraI_F$ PDB: 2Q7T), show interactions between the
410 base of the thumb domain (α-helix 10) and the loop region following α-helix 1 (R274 to D25
411 in $TraI_{pKM101}$) (Fig. S4). In the DNA-bound structure of $TraI_{pKM101}$ an additional interaction
412 was found between H278 (α-helix 10) and Y26 (the third Y of the YY-$X_{(5-6)}$-YY motif), which
413 might further contribute to this stabilization.

414

415 To conclude, we have shown that $TraI_{pKM101}$ has a high affinity and specificity to its cognate
416 *oriT* and can both nick and religate *oriT* in a similar manner as other MOB_F family relaxases.
417 The crystal structures of its trans-esterase domain (TraI-TE_{pKM101}) visualize the large

418 conformational change the protein undergoes upon DNA binding. Both the biochemical
419 function and structures of $TraI_{pKM101}$ reported here resemble those of the homologs $TrwC_{R388}$
420 and $TraI_{F}$ more closely than the previously published work on the identical $TraI\text{-}TE_{pCU1}$[25]. Our
421 findings thus highlight the conserved mechanism of relaxases and the role they play in
422 conjugation.
423

Acknowledgments

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

CRediT statement

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

442 **References**

443 1. Macé, K. *et al.* Cryo-EM structure of a type IV secretion system. *Nature* (2022)
444 doi:10.1038/s41586-022-04859-y.
445 2. Khara, P., Song, L., Christie, P. J. & Hu, B. *In Situ* Visualization of the pKM101-Encoded Type
446 IV Secretion System Reveals a Highly Symmetric ATPase Energy Center. *mBio* e02465-21
447 (2021) doi:10.1128/mBio.02465-21.
448 3. Costa, T. R. D. *et al.* Type IV secretion systems: Advances in structure, function, and
449 activation. *Mol. Microbiol.* **115**, 436–452 (2021).
450 4. Grohmann, E., Christie, P. J., Waksman, G. & Backert, S. Type IV secretion in Gram-
451 negative and Gram-positive bacteria: Type IV secretion. *Mol. Microbiol.* **107**, 455–471
452 (2018).
453 5. Zechner, E. L., Moncalián, G. & de la Cruz, F. Relaxases and Plasmid Transfer in Gram-
454 Negative Bacteria. in *Type IV Secretion in Gram-Negative and Gram-Positive Bacteria* (eds.
455 Backert, S. & Grohmann, E.) vol. 413 93–113 (Springer International Publishing, 2017).
456 6. Garcillán-Barcia, M. P., Francia, M. V. & de La Cruz, F. The diversity of conjugative
457 relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* **33**, 657–687
458 (2009).
459 7. Byrd, D. R. & Matson, S. W. Nicking by transesterification: the reaction catalysed by a
460 relaxase. *Mol. Microbiol.* **25**, 1011–1022 (1997).
461 8. Alvarez-Martinez, C. E. & Christie, P. J. Biological Diversity of Prokaryotic Type IV Secretion
462 Systems. *Microbiol. Mol. Biol. Rev.* **73**, 775–808 (2009).
463 9. Waksman, G. From conjugation to T4S systems in Gram-negative bacteria: a mechanistic
464 biology perspective. *EMBO Rep.* **20**, (2019).

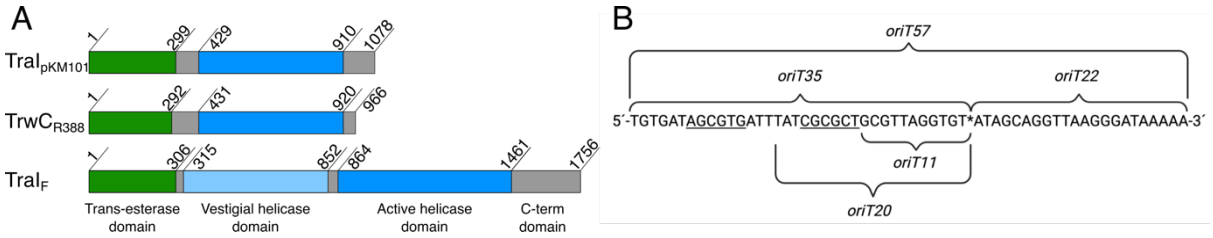
465 10. Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., de la Cruz, F. & Rocha, E. P. C. The
466 Repertoire of ICE in Prokaryotes Underscores the Unity, Diversity, and Ubiquity of
467 Conjugation. *PLoS Genet.* **7**, e1002222 (2011).
468 11. Paterson, E. S. *et al.* Genetic Analysis of the Mobilization and Leading Regions of the IncN
469 plasmids pKM101 and pCU1. *J. Bacteriol.* **181**, 2572 (1999).
470 12. De La Cruz, F., Frost, L. S., Meyer, R. J. & Zechner, E. L. Conjugative DNA metabolism in
471 Gram-negative bacteria. *FEMS Microbiol. Rev.* **34**, 18–40 (2010).
472 13. Larkin, C. *et al.* Inter- and Intramolecular Determinants of the Specificity of Single-
473 Stranded DNA Binding and Cleavage by the F Factor Relaxase. *Structure* **13**, 1533–1544
474 (2005).
475 14. Boer, R. *et al.* Unveiling the Molecular Mechanism of a Conjugative Relaxase: The
476 Structure of TrwC Complexed with a 27-mer DNA Comprising the Recognition Hairpin and
477 the Cleavage Site. *J. Mol. Biol.* **358**, 857–869 (2006).
478 15. Ilangovan, A. *et al.* Cryo-EM Structure of a Relaxase Reveals the Molecular Basis of DNA
479 Unwinding during Bacterial Conjugation. *Cell* **169**, 708-721.e12 (2017).
480 16. Geertsma, E. R. & Dutzler, R. A Versatile and Efficient High-Throughput Cloning Tool for
481 Structural Biology. *Biochemistry* **50**, 3272–3278 (2011).
482 17. Some, D., Amartely, H., Tsadok, A. & Lebendiker, M. Characterization of Proteins by Size-
483 Exclusion Chromatography Coupled to Multi-Angle Light Scattering (SEC-MALS). *J. Vis.*
484 *Exp.* 59615 (2019) doi:10.3791/59615.
485 18. Kabsch, W. *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
486 19. Monaco, S. *et al.* Automatic processing of macromolecular crystallography X-ray
487 diffraction data at the ESRF. *J. Appl. Crystallogr.* **46**, 804–810 (2013).

488 20. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
489 (2007).
490 21. Emsley, P. & Cowtan, K. *Coot* : model-building tools for molecular graphics. *Acta*
491 *Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
492 22. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with
493 *phenix.refine*. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
494 23. Guasch, A. *et al.* Recognition and processing of the origin of transfer DNA by conjugative
495 relaxase TrwC. *Nat. Struct. Mol. Biol.* **10**, 1002–1010 (2003).
496 24. Datta, S., Larkin, C. & Schildbach, J. F. Structural Insights into Single-Stranded DNA Binding
497 and Cleavage by F Factor TraI. *Structure* **11**, 1369–1379 (2003).
498 25. Nash, R. P., Habibi, S., Cheng, Y., Lujan, S. A. & Redinbo, M. R. The mechanism and control
499 of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res.*
500 **38**, 5929–5943 (2010).
501 26. Carballeira, J. D., González-Pérez, B., Moncalián, G. & la Cruz, F. de. A high security double
502 lock and key mechanism in HUH relaxases controls oriT-processing for plasmid
503 conjugation. *Nucleic Acids Res.* **42**, 10632–10643 (2014).
504 27. Stern, J. C. & Schildbach, J. F. DNA Recognition by F Factor TraI36: Highly Sequence-
505 Specific Binding of Single-Stranded DNA. *Biochemistry* **40**, 11586–11595 (2001).
506 28. Amor-Mahjoub, M., Suppini, J.-P., Gomez-Vrielyunck, N. & Ladjimi, M. The effect of the
507 hexahistidine-tag in the oligomerization of HSC70 constructs. *J. Chromatogr. B* **844**, 328–
508 334 (2006).
509 29. Singh, M. *et al.* Effect of N-terminal poly histidine-tag on immunogenicity of Streptococcus
510 pneumoniae surface protein SP0845. *Int. J. Biol. Macromol.* **163**, 1240–1248 (2020).

511 30. Li, Y. G. & Christie, P. J. The TraK Accessory Factor Activates Substrate Transfer through
512 the pKM101 Type IV Secretion System Independently of its Role in Relaxosome Assembly.
513 *Mol. Microbiol.* (2020) doi:10.1111/mmi.14507.
514 31. Holm, L. Using Dali for Protein Structure Comparison. in *Structural Bioinformatics* (ed.
515 Gáspári, Z.) vol. 2112 29–42 (Springer US, 2020).
516

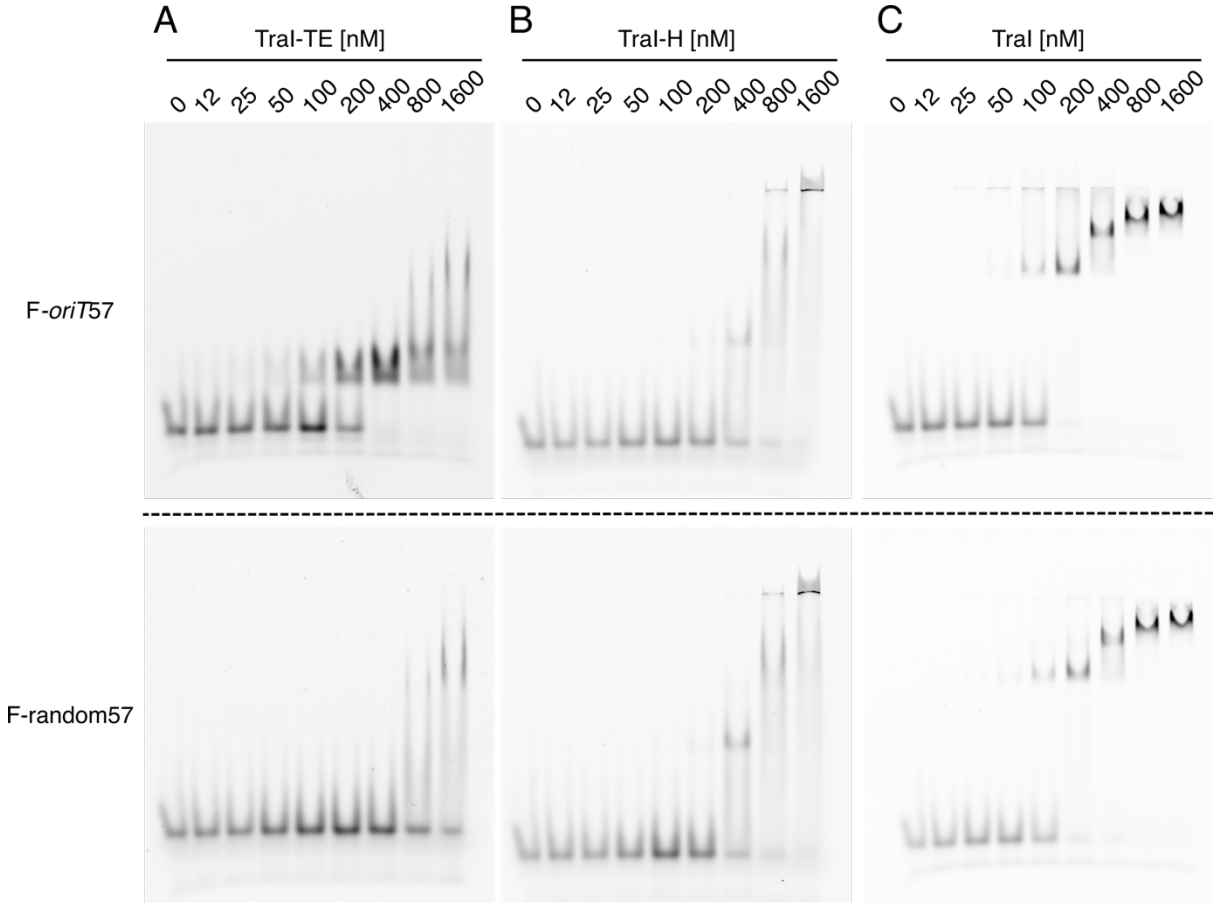
517 **Figures**

518



519

520 **Figure 1**. Schematic overviews of domain organization of relaxases and the cognate *oriT* of
521 pKM101. (A) Domain structure of three MOB_F family relaxases $TraI_{pKM101}$ (identical to
522 $TraI_{pCU1}$), $TrwC_{R388}$ and $TraI_F$[5,15]. Trans-esterase domains are shown in green, vestigial helicase
523 domain in light blue, active helicase domains in dark blue and the C-terminal domains and
524 linker regions are shown in grey. (B) *oriT* sequence of pKM101 and an overview of the
525 different *oriT* nucleotides used in this study. * indicates the *nic*-site, underlined regions of the
526 sequence indicate the bases participating in hairpin formation.



527

528 **Figure 2**. Electromobility shift assay using 50 nM fluorescent 57-mer *oriT* (F-*oriT*57) (upper
529 panel) or random DNA (lower panel) of the trans-esterase domain (TraI-TE) (A), helicase
530 domain (TraI-H) (B) or full-length TraI (C).

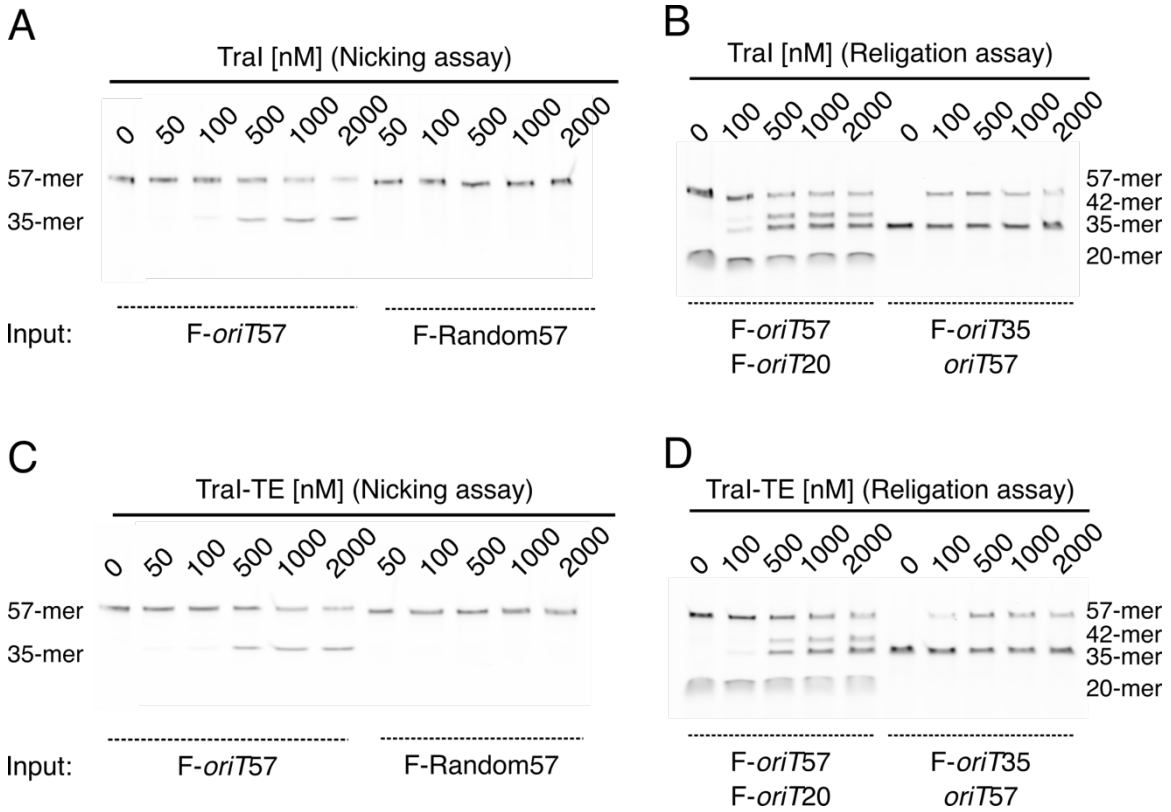


Figure 3. Activity assays of TraI (A, B) and TraI-TE (C, D). (A, C) Nicking assays. F-*oriT*57 (left) and F-Random57 (right) were incubated with increasing protein concentrations and the nicked DNA was separated on a denaturing gel, showing the emergence of a 35-mer nicking product for *oriT* DNA only. (B, D) Religation assays. F-*oriT*57 and F-*oriT*20 were incubated with increasing protein concentrations (left). The appearance of a 35-mer product indicates nicking of F-*oriT*57 and the appearance of a 42-mer product indicates religation of the covalently bound post-nick DNA with the shorter F-*oriT*20 pre-nick DNA. F-*oriT*35 and *oriT*57 were incubated with increasing protein concentrations (right). The appearance of a 57-mer product shows nicking of the non-fluorescent *oriT*57 and religation with the fluorescent F-*oriT*35.

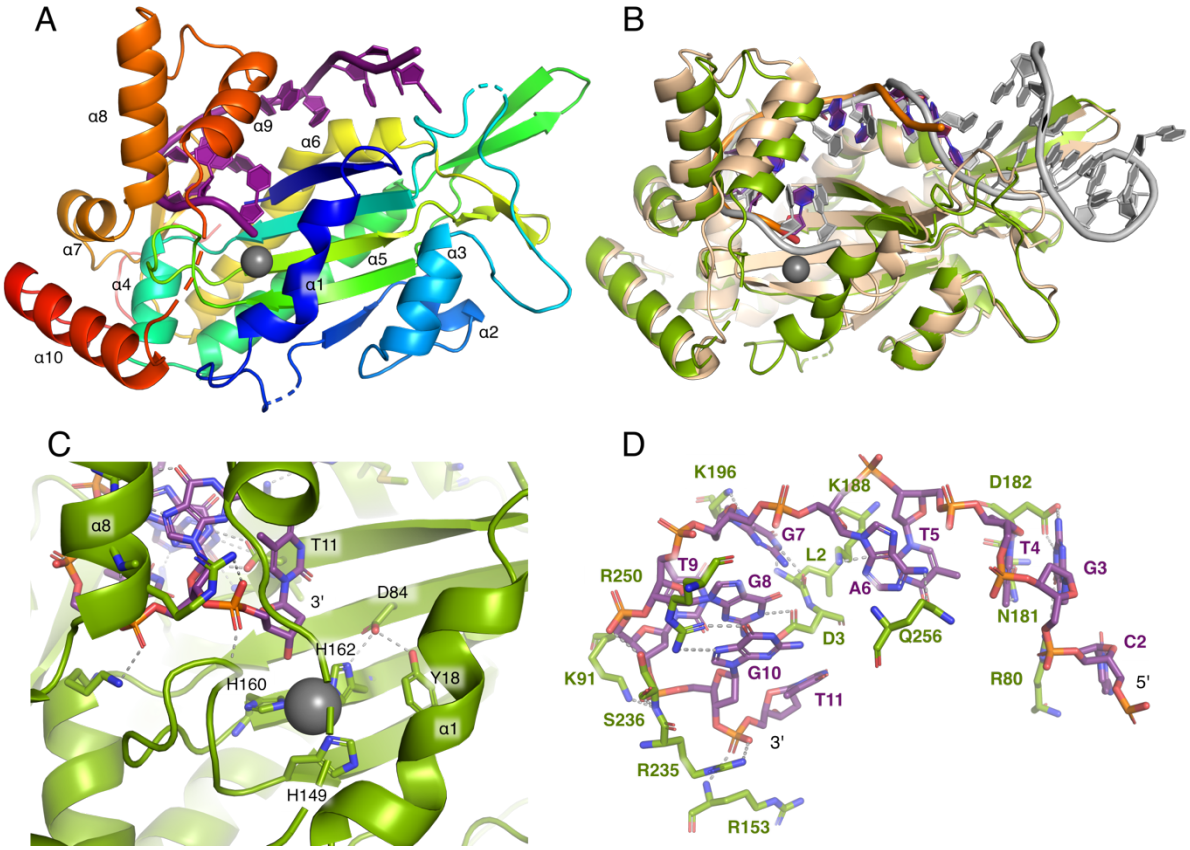
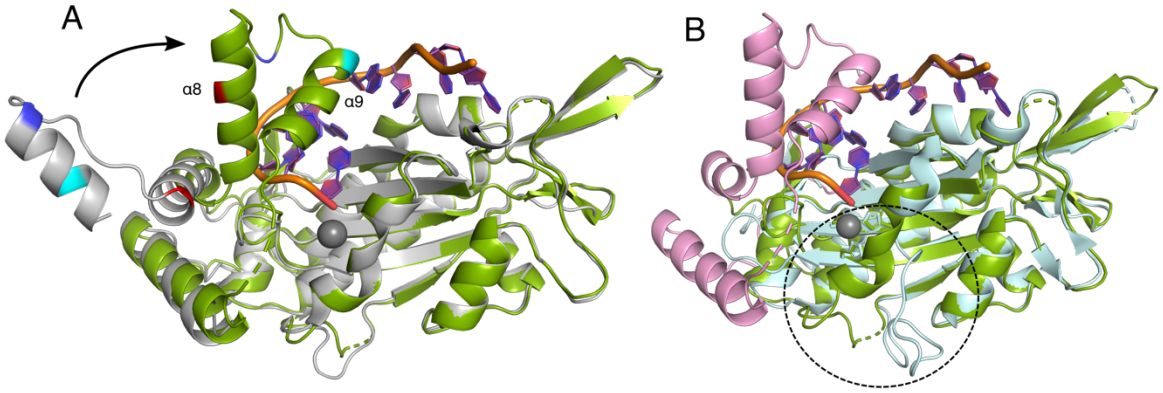


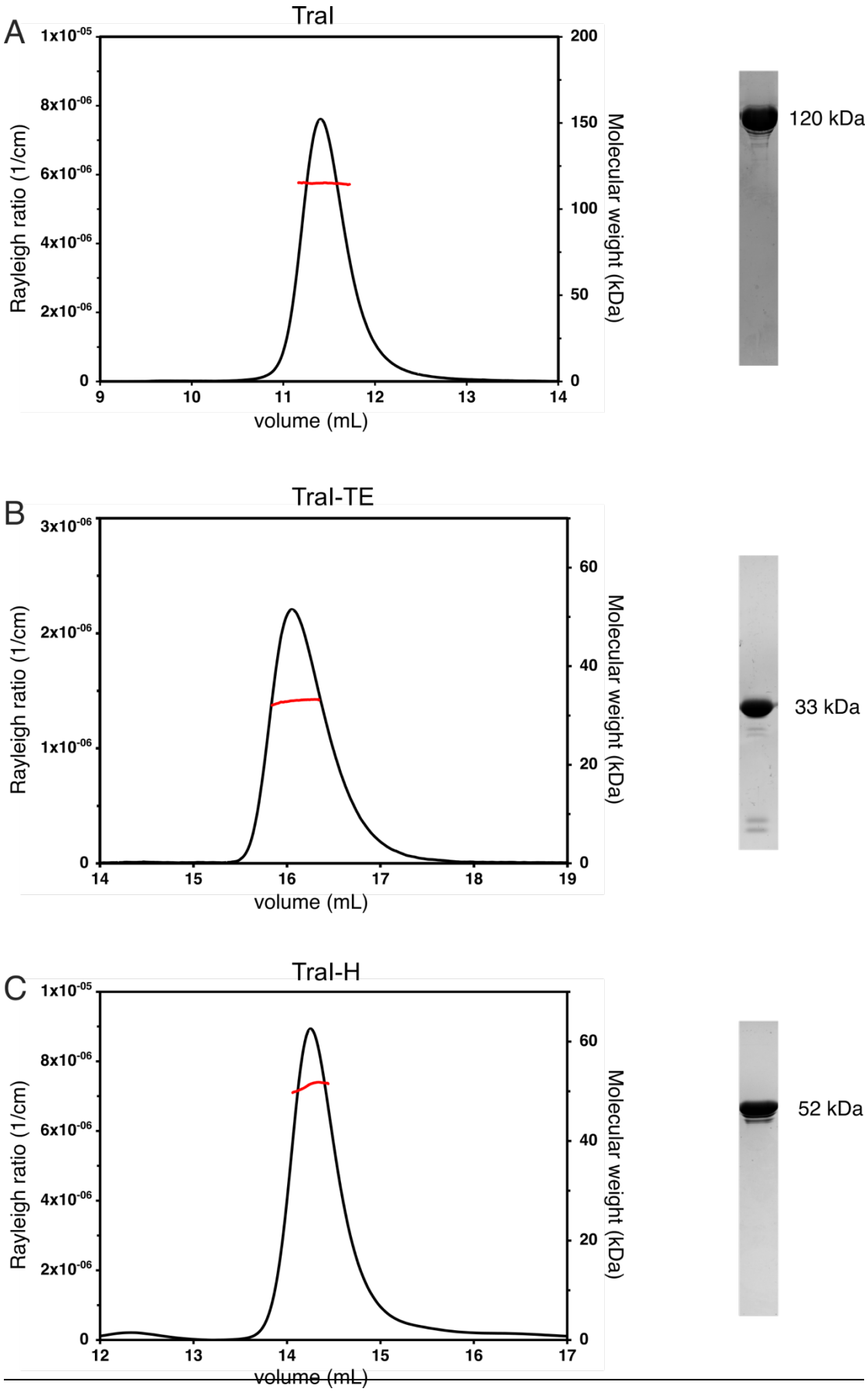
Figure 4. Crystal structure of DNA-bound trans-esterase domain of $TraI_{pKM101}$. In all panels TraI is colored green, the DNA bound by $TraI_{pKM101}$ is shown in purple and the bound Mg^{2+} ion in gray, unless otherwise indicated. (A) Cartoon representation of the trans-esterase domains of $TraI_{pKM101}$ colored from blue at the N-terminal to red at the C-terminal, with alpha helices labeled from 1 to 10. (B) Superimposition of the trans-esterase domains of $TraI_{pKM101}$ and $TrwC_{R388}$ (protein in beige, DNA in grey, PDB: 2CDM). (C) Active site of TraI-TE_{pKM101} consisting of the histidine triad coordinating a Mn^{2+} (H149, H160 and H162), conserved aspartate (D84) and catalytic tyrosine (Y18). (D) The bound *oriT* DNA and its interactions with TraI-TE. Residues that are forming hydrogen bonds with DNA and the DNA itself are represented as sticks.



553

554 **Figure 5**. Overview of the conformational changes between the apo- and the DNA-bound
555 structures and the differences with the previously published structure of the identical TraI-
556 TE_{pCU1}. (A) DNA-bound structure of $TraI_{pKM101}$ (protein in green, DNA in purple, Mn^{2+} as grey
557 sphere), superimposed with the apo structure (light grey). To provide a visual aid in
558 understanding the conformational changes that occur in the thumb domain upon DNA binding,
559 residue K241 (red), R251 (blue) and Q256 (cyan) are highlighted in both structures and alpha-
560 helix 8 and 9 are indicated in the DNA-bound structure. (B) DNA-bound structure of $TraI_{pKM101}$
561 (colored as in panel A) with thumb domain and C-terminus highlighted in pink, superimposed
562 with the previously published structure of TraI-TE_{pCU1} (cyan, PDB: 3L6T). TraI-TE_{pCU1} lacks
563 the thumb domain, and differences with TraI-TE_{pKM101} are also observed close to the active
564 site, indicated by the dashed circle.

565 **Supplementary figures**



566

567 **Figure S1**. Results from size-exclusion chromatography coupled to multi-angle light
568 scattering (SEC-MALS) experiments of TraI constructs without His-tag (left panels) and
569 purified fractions of these protein constructs on SDS-PAGE (right panels). In the left panels,
570 the black chromatograms show the Rayleigh ratio which is directly proportional to the
571 intensity of the scattered light in excess to the solvent (left axis). The red lines indicate the
572 calculated molecular weight of the protein throughout the peak, plotted on the right axis. The
573 average molecular weight throughout the peaks were calculated to be: 116 +/- 1 kDa for TraI
574 (A), 33 +/- 1 kDa for TraI-TE (B) and 49 +/- 2 kDa for TraI-H. These corresponded to the
575 theoretical molecular weight of the monomers, which is indicated next to the SDS-PAGE
576 fractions in each of the panels.

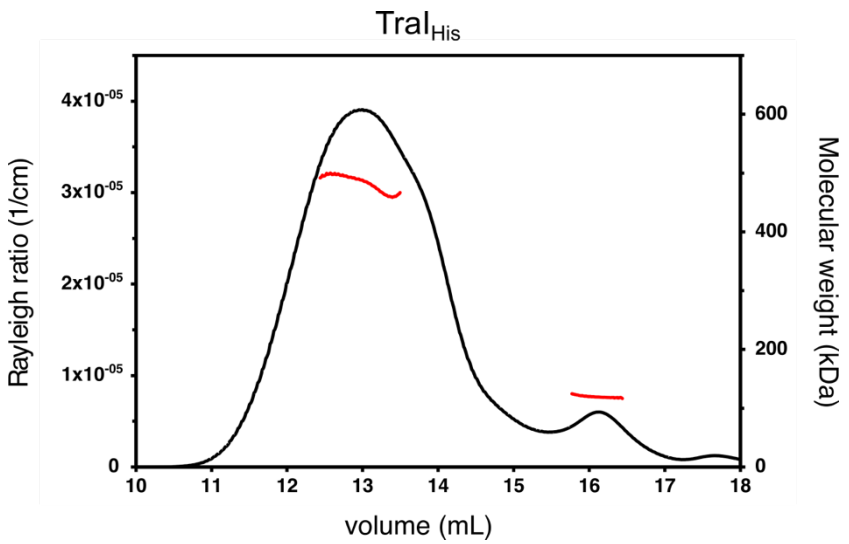
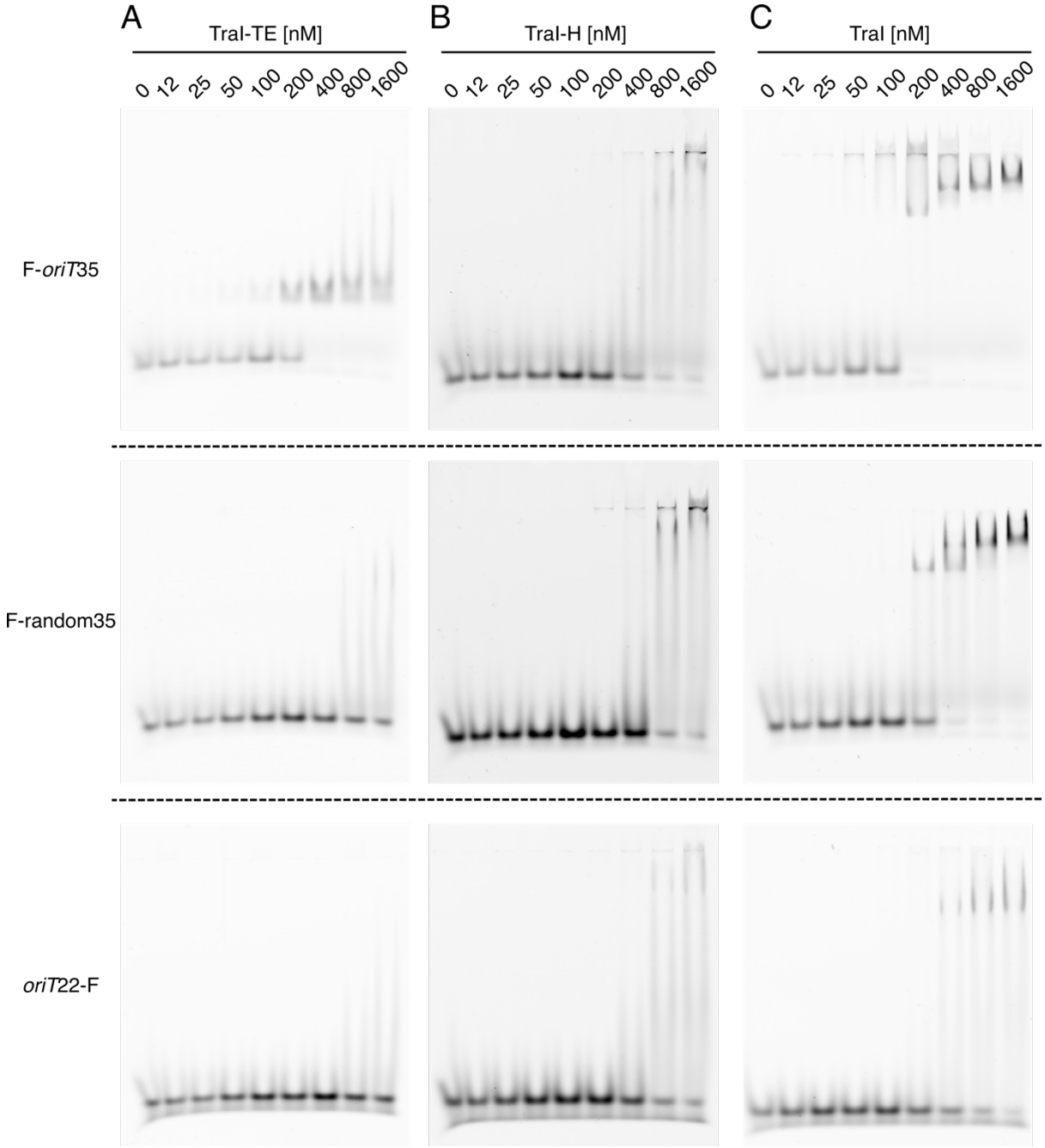


Figure S2. His-tag induced tetramerization of TraI$_{His}$. The chromatogram shows the elution profile and calculated molecular weight from size-exclusion chromatography coupled to multiangle light scattering (SEC-MALS) of TraI without removal of the His-tag (TraI$_{His}$). The Rayleigh ratio which is directly proportional to the intensity of the scattered light in excess to the solvent is indicated with a black line plotted on the left axis and the calculated molecular weight of the protein throughout the peak is indicated with a red line and plotted on the right axis. The average molecular weight throughout the first elution peak is 485 +/- 10 kDa, corresponding to a tetramer, and throughout the second elution peak this was calculated to be 123 +/- 5 kDa corresponding to a monomer.



587

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

591

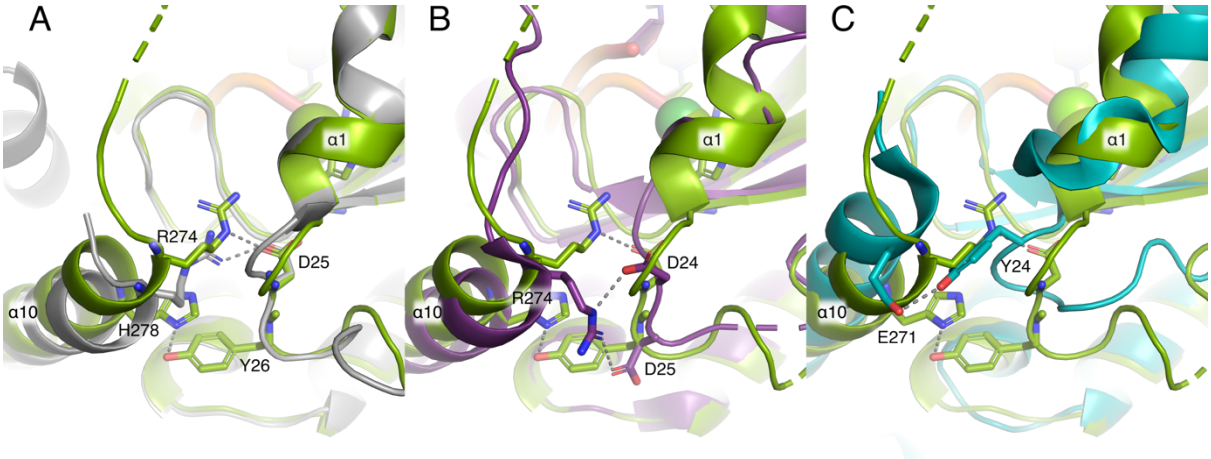


Figure S4. Overview of the interactions between the base of the thumb domain (α-helix 10) and the loop at the end of α-helix 1. In all panels, TraI is shown in in green, with important residues shown as sticks. (A) DNA-bound structure superimposed with the apo-structure (in grey) of TraI. Hydrogen bonds are formed between R274 and D25 in both structures, while an additional bond between H278 and Y26 is only present in the DNA-bound structure. (B) DNA-bound structure of $TraI_{pKM101}$ superimposed with $TrwC_{R388}$ (PDB: 1S6M, purple). In $TrwC_{R388}$ hydrogen bonds are formed between R274 and D24 or D25. (C) DNA-bound structure of $TraI_{pKM101}$ superimposed with $TraI_F$ (PDB: 2Q7T, turquoise). In $TraI_F$ hydrogen bonds are found between E271 and Y24.

Movie S1. Visual representation of the thumb domain movement during the transition from the apo- to the DNA-bound state of $TraI_{pKM101}$. The apo-structure is represented in light grey, the DNA-bound structure is shown as a cartoon with protein in green and DNA in purple and with the Mn^{2+} ion as a dark grey sphere. In both structures the following residues are highlighted by color to make it easier to follow the conformational change: K241 (red), R251 (blue) and Q256 (cyan).

610 **Supplementary tables**

611

612 **Table S1**. List of strains and plasmids used.

Strain, plasmid, or oligonucleotide	**Relevant features or sequences (5'-3')**	**References**
Strains (*E. coli*):		
TOP10 One Shot	Cloning host	Thermo-Fisher
BL21 (DE3)	Expression host	Thermo-Fisher
Plasmids:		
P7XC3H	FX cloning expression vector with a C-terminal 10-His-tag	Geertsma & Dutzler[16]
P7XNH3	FX cloning expression vector with a N-terminal 10-His-tag	Geertsma & Dutzler[16]
p7XC3H-TraI	P7XC3H expressing *traI-his*$_{10}$	This study
p7XNH3-TraI-H	P7XNH3 expressing *his*$_{10}$-*traI*$_{429-910}$	This study
p7XC3H-TraI-TE	P7XC3H expressing *traI*$_{1-299}$-*his*$_{10}$	This study
FX-cloning primers		
Fw-TraI	ATATATGCTCTTCTAGTCTTGATATAACCACGATTACCCGCCAG	Eurofins
Rv-TraI	TATATAGCTCTTCATGCGATTTCATGGCCCCCTTCTTCATGCTC	Eurofins
Rv-TraI-TE	TATATAGCTCTTCATGCTCCCTGCCATTCACGGTTATCAAAATC	Eurofins
Fw-TraI-H	ATATATGCTCTTCTAGTCTTAAAAAGACCAGCCACCGCGTCACT	Eurofins
Rv-TraI-H	TATATAGCTCTTCATGCGCTCACGCTCGATGCCAAAGATTTTTT	Eurofins
***ss*DNA oligomers**		
oriT57	TGTGATAGCGTGATTTATCGCGCTGCGTTAGGTGTATAGCAGGTTAAGGGATAAAAA	
F-*oriT*57	**FITC**-TGTGATAGCGTGATTTATCGCGCTGCGTTAGGTGTATAGCAGGTTAAGGGATAAAAA	Eurofins
F-*oriT*35	**FITC**-TGTGATAGCGTGATTTATCGCGCTGCGTTAGGTGT	Eurofins
*F-oriT*20	**FITC**-TATCGCGCTGCGTTAGGTGT	Eurofins
*oriT*11	GCGTTAGGTGT	Eurofins
*oriT*22-F	ATAGCAGGTTAAGGGATAAAAA-**FITC**	Eurofins

F-Random57	**FITC**-TCCGCCATGCAGACGAGACCAGTCGGAGATTACCGAGCATTCTATCAGGTCGGCGAC	Eurofins
F-Random35	**FITC**-CACTAGTGAGCTACTGGAGCCGAGGGGTAACCACG	Eurofins

613

614

615 **Table S2**. Data collection and refinement statistics.

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

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

617