Molecular Microbiology (2010) 78(6), 1539–1555 ■
doi:10.1111/j.1365-2958.2010.07423.x
First published online 14 October 2010

Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases

Silvia Lang,[1] Karl Gruber,[1] Sanja Mihajlovic,[1] Roland Arnold,[2] Christian J. Gruber,[1] Sonja Steinlechner,[1] Marc-Andre Jehl,[2] Thomas Rattei,[2,3] Kai-Uwe Fröhlich[1] and Ellen L. Zechner[1]*

[1]*Institute of Molecular Biosciences, University of Graz, Humboldtstrasse 50, 8010 Graz, Austria.*
[2]*Department of Genome Oriented Bioinformatics, Technische Universität München, Maximus-von-Imhof-Forum 3, 85354 Freising, Germany.*
[3]*Department of Computational Systems Biology, Ecology Centre, University of Vienna, Althanstr. 14, 1090 Vienna, Austria.*

Summary

In preparation for transfer conjugative type IV secretion systems (T4SS) produce a nucleoprotein adduct containing a relaxase enzyme covalently linked to the 5′ end of single-stranded plasmid DNA. The bound relaxase is expected to present features necessary for selective recognition by the type IV coupling protein (T4CP), which controls substrate entry to the envelope spanning secretion machinery. We prove that the IncF plasmid R1 relaxase TraI is translocated to the recipient cells. Using a Cre recombinase assay (CRAfT) we mapped two internally positioned translocation signals (TS) on F-like TraI proteins that independently mediate efficient recognition and secretion. Tertiary structure predictions for the TS matched best helicase RecD2 from *Deinococcus radiodurans*. The TS is widely conserved in MOB_F and MOB_Q families of relaxases. Structure/function relationships within the TS were identified by mutation. A key residue in specific recognition by T4CP TraD was revealed by a fidelity switch phenotype for an F to plasmid R1 exchange L626H mutation. Finally, we show that physical linkage of the relaxase catalytic domain to a TraI TS is necessary for efficient conjugative transfer.

Accepted 29 September, 2010. *For correspondence. E-mail ellen.zechner@uni-graz.at; Tel. (+43) 316 380 5624; Fax (+43) 316 380 9019.

Introduction

Bacterial type IV secretion systems (T4SS) translocate proteins and DNA across the cell envelope (Christie and Vogel, 2000; Alvarez-Martinez and Christie, 2009). One group of T4SS transfers effector proteins directly into the cytosol of eukaryotic cells (Fischer *et al*., 2002; Backert and Meyer, 2006; Schmid and Scheidegger, 2006; Juhas *et al*., 2008; Franco *et al*., 2009). Secretion of these proteins is often an essential part of the virulence mechanisms of pathogenic bacteria. A second group secretes DNA, covalently linked to protein, into a recipient cell. This group is responsible for DNA transfer by conjugation, and is important in the spread of antibiotic resistance (de la Cruz and Davies, 2000; Christie, 2001). A third group of T4SS transfers DNA into and from the extracellular milieu (Chen *et al*., 2005; Hamilton *et al*., 2005). This classification emphasizes the versatility and importance of T4SS, but the groups also overlap functionally: for example, systems involved in translocation of virulence proteins can often transfer plasmid DNA.

A conserved component called the type IV coupling protein (T4CP) identifies the appropriate macromolecules for secretion in most systems (Gomis-Rüth *et al*., 2001; Llosa *et al*., 2002; Schröder and Lanka, 2003; Alvarez-Martinez and Christie, 2009). The underlying mechanisms are not fully understood. Some knowledge of secretion protein recognition features is available (referred to hereafter as translocation signals, or TS). Data from the *Agrobacterium tumefaciens* VirB/VirD4 paradigm and the *Legionella pneumophila* Dot/Icm T4SS revealed that a surprisingly simple cluster of positively charged or hydrophobic residues at the C-termini constitute a functional TS (Nagai *et al*., 2005; Vergunst *et al*., 2005). More complex bipartite TS are shared by the seven Bartonella effector proteins, BepA-G. The charged C-terminal residues are combined with one or more copies of a second motif intrinsic to the conserved Bep intracellular delivery (BID) domain (Schulein *et al*., 2005). The CagA effector protein of *Helicobacter pylori* carries a 20-amino-acid C-terminal TS adjacent to a larger interaction domain for its secretion chaperone CagF (Hohlfeld *et al*., 2006; Pattis *et al*., 2007). Together, these features are sufficient to recruit CagA to the secretion apparatus. A subsequent recognition step involves an N-terminal domain (Hohlfeld *et al*.,

Table 1. TraI protein is functional after conjugative transport to a recipient cell.

	Donor strains 1 MS411	Recipient strain 2 MS411	Recipient strain 3 MS614 (Sm^R)	Mobilization frequency to strain 3 (Amp+Sm)	Conjugation frequency to strain 3 (Km+Sm)
I.	[R1-16] [pGZ119EH]	*oriT* plasmid (Amp^R)	–	$6.1 \times 10^{-3} \pm 0.001$	5.4 ± 0.7
II.	[R1-16Δ*traI*] [pGZ119+*traI*]	*oriT* plasmid (Amp^R)	–	$2.8 \times 10^{-3} \pm 0.001$	0.7 ± 0.2
III.	[R1-16Δ*traI*] [pGZ119EH]	*oriT* plasmid (Amp^R)	–	$<10^{-7}$	$<10^{-7}$

Three independent triparental matings (I, II, III) combined MS411 carrying a conjugative plasmid (strain 1), MS411 carrying a mobilizable plasmid (strain 2) and MS614 (strain 3). The transfer frequencies are given as number of transconjugants per donor cell. Values represent the mean of three experiments. Standard deviations are shown. Amp, ampicillin; Km, kanamycin; Sm, streptomycin.

2006). Modular architecture is a feature shared by the few other characterized T4 protein substrates (Alvarez-Martinez and Christie, 2009).

In Gram-negative bacterial conjugation a series of conserved reactions prepares the plasmid DNA to be transferred in single-stranded form (de la Cruz *et al*., 2010). A relaxase and one or more accessory proteins recognize the plasmid origin of transfer (*oriT*) and catalyse site- and strand-specific cleavage of the DNA strand destined for transfer (T-strand). The phosphodiester bond on the T-strand where cleavage occurs and duplex unwinding begins is called *nic*. The processed DNA–protein complex, or relaxosome, is linked to the mating channel via the T4CP. As originally proposed (Willetts and Wilkins, 1984), the nucleoprotein adduct of relaxase covalently linked to the 5′ end of the T-strand is then conveyed to the recipient cell via the T4 secretion machinery (Draper *et al*., 2005; Parker and Meyer, 2007). At present, however, we still understand little about the specific relaxosome/T4CP interactions that initiate nucleoprotein uptake by the transporter.

Our work investigates F-like (MOB_F) systems where the relaxase TraI is fused to a conjugative DNA helicase. Biochemical studies with T4CP TraD and MOB_F relaxosome proteins demonstrated physical interaction between TraD and TraM (Disque-Kochem and Dreiseikelmann, 1997; Beranek *et al*., 2004; Lu *et al*., 2008). TraD stimulates both the *nic*-cleavage and DNA helicase activities of TraI from plasmid R1 (Mihajlovic *et al*., 2009; Sut *et al*., 2009). Here we identify the TS of F-like TraI proteins and determine key functional residues. The TS are predicted to share structural similarity and are broadly conserved in MOB_F and MOB_Q relaxases. Moreover, the data imply that a covalent connection of TraI TS and catalytic domains is required for productive conjugative DNA processing.

Results

TraI protein is translocated to the recipient

Triparental mating experiments were performed to test whether TraI protein is physically transferred via bacterial conjugation to recipient cells. As summarized in Table 1 MS411 donor cells (strain 1) were compared that carried either R1-16 (line I) or a conjugation deficient R1-16Δ*traI* with (line II) or without (line III) *traI* complementation *in trans*. To monitor TraI protein transfer two distinct recipient strains were present. One harboured a mobilizable *oriT* Amp^R plasmid (strain 2). Strain 3 is resistant to streptomycin (Sm) and carried no plasmid. For triparental mating the individual donor strains were combined with both recipient types. Transfer of the conjugative plasmids to strain 3 was detected by selection on kanamycin (Km) and Sm. As expected, the R1-16 *traI* null derivative was transfer proficient only when the donor strain expressed *traI* (II). Selection for Amp^R Sm^R transconjugants of strain 3 identifies cells that acquired the *oriT* mobilizable plasmid. Transfer of this plasmid from strain 2 depends on the presence of TraI protein for conjugative DNA processing. Transconjugant cells that acquired the mutated R1-16 lack the *traI* gene, however. Thus, a subsequent mobilization of the *oriT* test plasmid from transconjugants with R1-16Δ*traI* should only be possible when functional TraI protein is transmitted from a donor to the recipient cell. Ampicillin-resistant transconjugants of strain 3 were observed with both R1-16 and the complemented R1-16Δ*traI* donor strains with equal efficiencies. We conclude therefore, that TraI is transmitted, either as a nucleoprotein complex or free protein during conjugation, and that the protein retains or refolds to a functional conformation in transconjugant cells. The related protein TrwC of plasmid R388 gives precedence for this finding (Draper *et al*., 2005).

Two functionally independent regions mediate TraI secretion

To map the TS in TraI protein we applied the Cre recombinase assay for translocation (CRAfT) (Vergunst *et al*., 2000; Vergunst *et al*., 2005). In this assay, Cre lacks features that enable it to be recognized and transported by a T4SS. Fusion of a *bona fide* secretion substrate, or its

Table 2. Bacterial strains used in this study.

Strain	Description and references[a]	Primer sequences 5′–3′[b]
E. coli DH5α	*endA1 recA1 gyrA96 thi-l hsdR17 supE44 λ⁻ relA1 deoR Δ(lacZYA-argF)- U169 φ80dlacZΔ(M15)* (Woodcock *et al.*, 1989)	
E. coli MS411	*ilvG rfb-50 thi* (M. Schembri; DTU, Denmark)	
E. coli MS614	Sm^R, *ilvG, rfb-50, thi, rpsL* (M. Schembri; DTU, Denmark)	
E. coli CSH26	*ara, Δ(lac-pro), thi* (Miller, 1972)	
E. coli CSH26crt	Cm^R, CSH26 *galK::cat::resC-tet-resC*, this study	
E. coli TOP10::LKL	Km^R, *loxP-Km-loxP* (Parker and Meyer, 2007)	
E. coli CSH26Cm::LKL	Km^R, CSH26 *galK::cat::loxP-Km-loxP*, this study	Fw_loxPCm: TTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTATAACTTCGTATAGCATA Rev_loxPCm: CTTTACGATGCCATTGGGATATATCAACGGTGGTATATCCGGATAACTTCGTATAATGTA
E. coli CHS26Cm::LTL	Tc^R, CSH26 *galK::cat::loxP-Tet-loxP*, this study	Fw_loxPTc: CCCGGGTGACTAACTAGGAGGAATAAATGGCTAAAATGAGACAAGAATTGCCGGCGGAT Rev_loxPTc: CCCGGGTCATTATGCCCTCCAGGTACTAAAACAATTCATCGGTATTTCACACCGCATAGC
E. coli MS614Cm	Cm^R, derivative of MG1655Str (Miranda *et al.*, 2004), this study	
E. coli DY330	W3110 *ΔlacU169 gal 490 ts λcl857 Δ(cro-bioA)* (Yu *et al.*, 2000)	

a. Antibiotic resistance: Cm^R, chloramphenicol; Km^R, kanamycin; Sm^R, streptomycin; Tc^R, tetracycline.
b. Underlined: homologous region.

TS regions, to Cre supports secretion of the fusion by the cognate T4SS. The indicator recipient strain (Table 2) harbours one antibiotic resistance gene interrupted by a second resistance cassette flanked by *loxP* sites. Recombination catalysed by the acquired Cre fusion at *loxP* restores functional expression of the disrupted resistance cassette. Protein transfer to a recipient strain is thus measured by the heritable change in antibiotic resistance phenotype. For this study, the full-length *traI* of R1 and a suite of sequential *traI* deletion variants were fused to the 3′ end of the Cre recombinase gene (Fig. 1). A normal level of recombination and, indirectly, the stability of all fusion proteins described in this work was first confirmed through transformation of the indicator strain with the Cre-plasmids and selection for recombinant cells (data not shown). In the CRAfT assay, donor cells harbour a derepressed conjugative plasmid, R1-16, to provide all of the essential components for substrate recognition, conjugative DNA processing and transport (including wild-type TraI protein). The translocation frequencies of the R1 fusion proteins were measured by scoring recombinants per donor cells (Fig. 1). R1-16 plasmid transfer occurs simultaneously and provided an internal standard for the conjugative transfer efficiency in every experiment. The summary illustration in Fig. 1B depicts two minimal TS identified in the R1 TraI protein (TSA and TSB, *grey*). The mapped positions span residues 530–816 (TSA) and 1255–1564 (TSB). As an alternative explanation the results could correspond to single-stranded DNA (ssDNA) binding domains associated with the TraI helicase, in which case the results would reflect an artefact of hitch-hiking on the T-strand. Recent mapping of the ssDNA binding activities for TraI of F and R1 excludes that possibility (Dostal and Schildbach, 2010).

To verify the functional assignment of these TS we performed a similar mapping with the closely related TraI protein of plasmid F. These experiments identified identical minimal TS (residues 530 to 816 and 1255 to 1564) that were competent for translocation by the F derivative pOX38 (Fig. 2, grey). The transfer efficiency of a set of F TraI variants with 31 residue insertions (i31) at defined locations in the protein (Haft *et al.*, 2006) was then compared. When TSA is separated from TSB, all i31 in TSA of F disrupted or significantly lowered (20- to 60-fold) protein transfer. Interestingly, however, the same i31 mutations in the full-length TraI were efficiently transferred. Thus, a single wild-type TS in the full-length protein was sufficient to complement the effect of a mutation in the second TS. We conclude that (i) the two TraI TS function independently, (ii) one TS is sufficient for recognition and uptake by the transporter, and (iii) dominant negative interference of these TS during the process is ruled out.

Conserved structure supports fold recognition for F-like TS

The existence of two regions of TraI, which are independently competent for translocation, implies that these

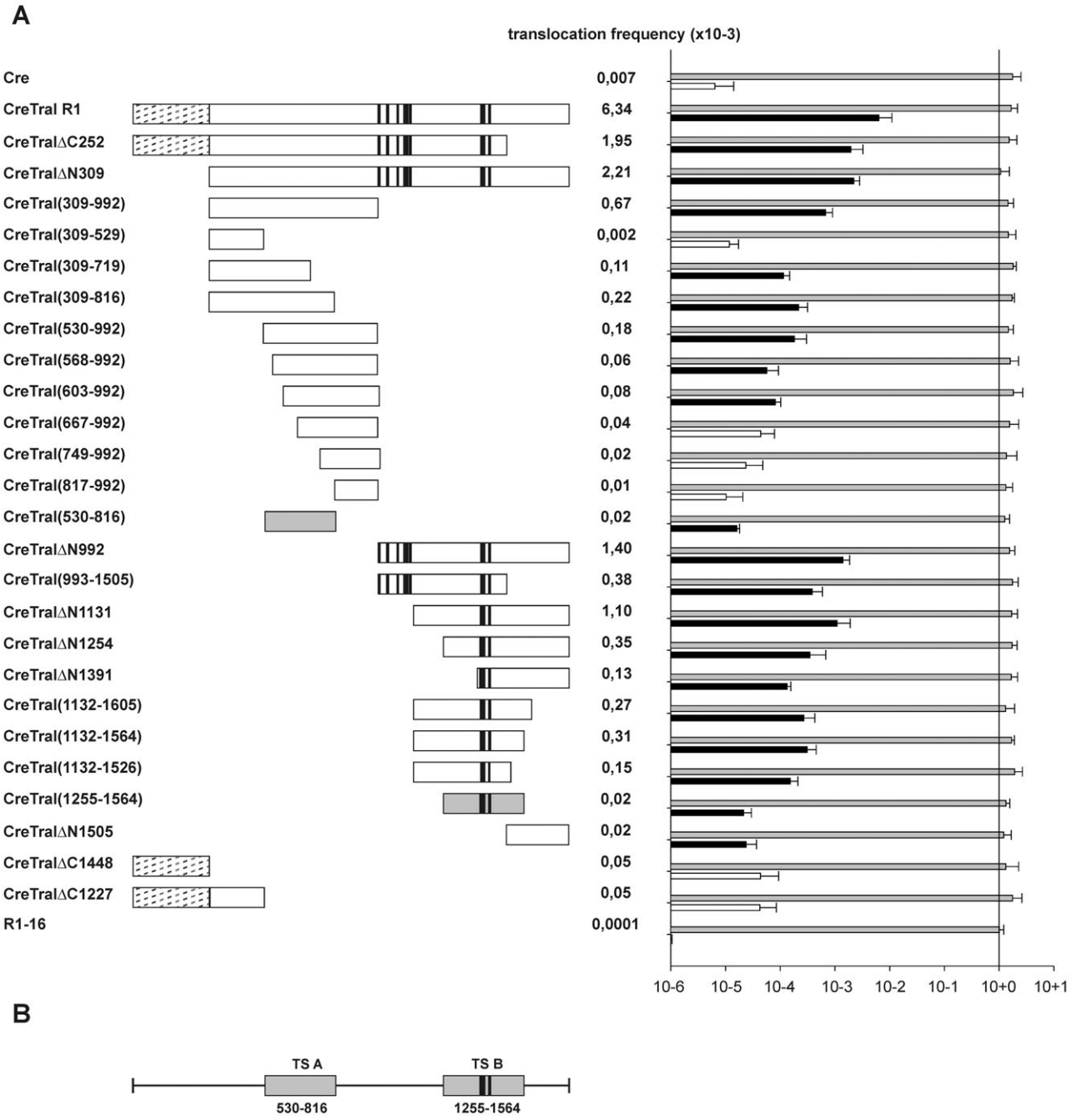

Fig. 1. Functional mapping of the TS of R1 TraI protein.
A. Schematic representations of full-length TraI and fragments fused to an N-terminal Cre recombinase (left). Patterned box at the TraI N-terminus represents the relaxase. Black vertical stripes indicate the position of conserved helicase motifs. The minimal TS identified are shown in grey. Vector expressing Cre recombinase alone serves as negative control. The translocation frequency is given as recombination events per donor (right). Bars represent frequencies of translocation with (black) or without (white) statistical significance compared with the negative control. Concomitantly measured R1-16 transfer frequencies are shown (grey bars). Values represent the mean of at least three experiments. Standard deviations are shown.
B. Summary of the mapped TraI TS A and B.

share characteristics important to T4 recognition. Although alignment of the primary sequences of the TS regions of TraI did not reveal much similarity it was reasonable to predict that the pair of signals may be structurally related. Tertiary structure prediction using the fold recognition server GenTHREADER (Jones, 1999; McGuffin and Jones, 2003) revealed the structure of an N-terminal truncation mutant of RecD2 from *Deinococcus*

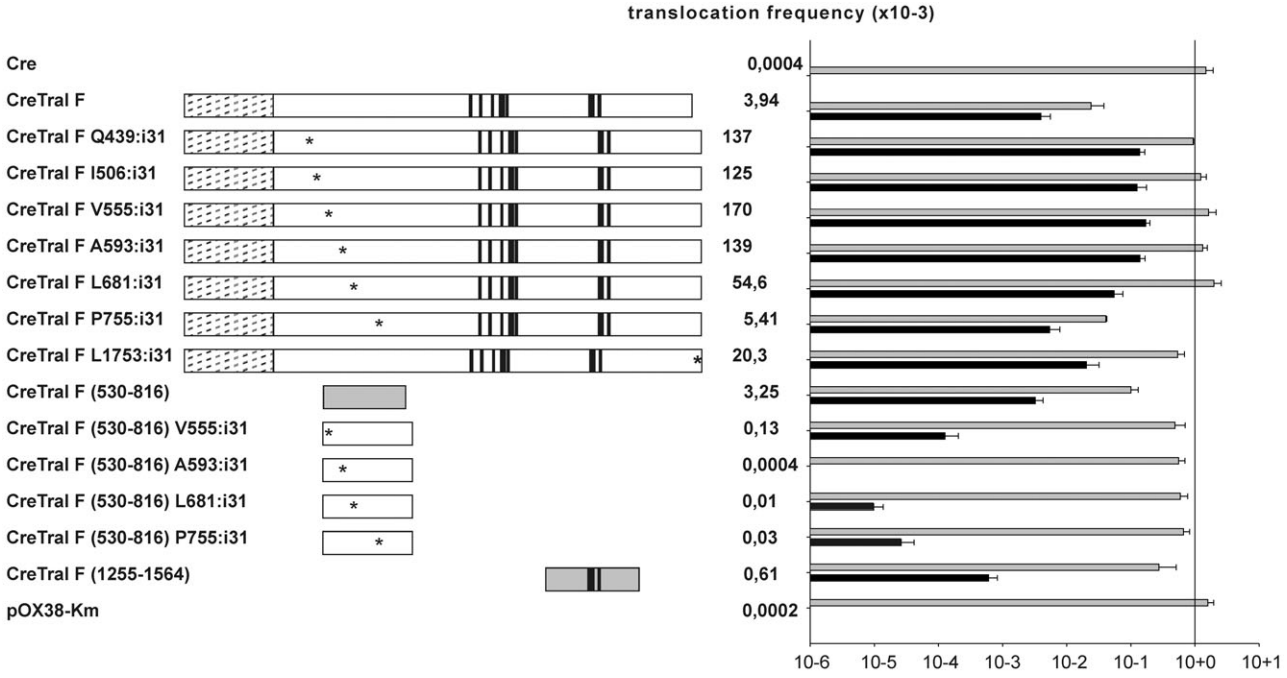

Fig. 2. Verification of the TS of F TraI. A set of TraI derivatives with 31 residue insertions at defined locations in the protein (*) was tested for translocation. TraI fragments are illustrated (left) with the minimal TS identified (grey). Frequencies (recombination events per donor) are shown (right). Bars represent frequencies of translocation with (black) or without (white) statistical significance compared with the negative control (Cre). Conjugation frequencies of pOX38 are indicated with grey bars. Values represent the mean of at least three experiments. Standard deviations are shown.

radiodurans (PDB-entry: 3e1s) (Saikrishnan *et al*., 2008) as the best template for both TSA and TSB (12% sequence identity). RecD2 is a 5′–3′ helicase belonging to the family of RecD-like proteins lacking RecB and RecC (Wang and Julin, 2004; Rocha *et al*., 2005). TSB also matched the structure of the C-terminal domain of F-TraI (PDB-entry: 3fld) (Guogas *et al*., 2009), as expected. The structurally conserved region common to TSA and TSB is shown in Fig. 3A. The sequence structure alignments of TSA and TSB to RecD2 obtained from GenTHREADER (shown in Fig. S1) revealed only about 10% sequence identity (20% similarity), but a cluster of conserved resi-

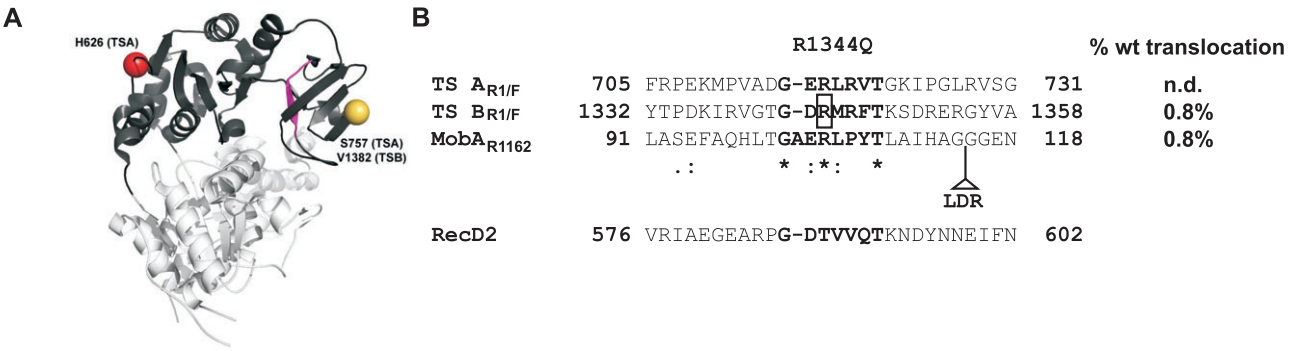

Fig. 3. Protein folding similarity is predicted for TSA and TSB with helicase RecD2 of *Deinococcus radiodurans*.
A. The structure of RecD2 is shown. The region conserved in TraI TS is indicated above (black). Variant amino acids in TSA of R1 versus F (S757) and in TSB of R1 versus F (V1382) colocalize (yellow). The second variant amino acid 626 in TSA (red) represents a fidelity determinant for efficient protein translocation. The conserved residues (G[E/D]R[L/M]R[V/F]T) similarly positioned in TSA and TSB are shown (pink β sheet).
B. Similarity to the consensus in a TS of $MobA_{R1162}$ is shown. A nearby insertion (LDR) in MobA, and R to Q mutations in R1 and F TSB (boxed), block translocation. Mutation of TSA was not done, n.d. The corresponding RecD2 sequence is shown underneath.

dues representing a β-strand (Fig. 3A, *pink*) was identified in both TS (Fig. 3B). Using the ClustalW program (Thompson *et al*., 2002) we searched for this cluster in $MobA_{R1162}$, which is the only other relaxase where TS have been mapped (Parker and Meyer, 2007). Two hits, one in each TS of $MobA_{R1162}$, were scored. Interestingly, mutations near this signature in the N-terminal TS of MobA disrupted protein transfer (Parker and Meyer, 2007), as did both R1344Q exchanges we created in R1/F TSB of TraI (Fig. 3B). Translocation of the mutant forms of TSB was barely detectable (3×10^{-6}), compared with the frequencies of wild-type TSB_{R1} and TSB_F ($2.9 \times 10^{-5} \pm 1.6 \times 10^{-5}$ and $3.9 \times 10^{-4} \pm 3.2 \times 10^{-4}$ respectively). We conclude that the TS present in TraI of IncF plasmids are structurally related to each other and to RecD2 and that conserved residues in these signals are important to function.

RecD/TS modules are conserved in relaxase families

Partial sequences of TSA and TSB (Fig. S1, bold letters) were used to assess the distribution of the TS/helicase module in the known relaxase families using PSI-BLAST (Altschul *et al*., 1997). Multiple alignments and subsequent phylogenies were constructed for the conserved regions identified among the relaxases (listed in Table S1). The maximum likelihood tree shown in Fig. 4 identifies five relatively well-resolved clades (indicated with coloured lines). The corresponding proteins overlap very closely with the MOB_Q and MOB_F groups defined previously (Garcillán-Barcia *et al*., 2009), as indicated with arcs in Fig. 4. Detailed information including bootstrap values and position of conservation for each entry protein is accessible in Fig. S2 and Table S1. Clades 1–4 (red, orange, yellow and green) represent conjugative relaxases harbouring only one TS/helicase module and correspond to MOB_Q, MOB_{F13}, MOB_{F11} and MOB_{F1} respectively (Garcillán-Barcia *et al*., 2009). Clade 5 (*blue*) is the largest and includes MOB_{F12} and other MOB_{F1} relaxases (Garcillán-Barcia *et al*., 2009). This clade can be divided into two branches (TSA, TSB; shaded) with each representing a different TS/helicase module in the same protein. The fact that TSA and TSB regions cluster together implies that they arose by an internal duplication event (arrow). Unexpectedly, the branching pattern and branch lengths for the two copies are very similar, with slightly longer branches (more diversification) for TSA. The constraints placed upon diversification of both TS-helicase modules imply that retention of the recognition properties for the T4CP is a key task for each domain. It is also important to note that although motifs important to the catalysis of DNA unwinding in F TraI are linked only to module TSB, the high affinity ssDNA binding activity associated with the helicase (residues 1 to 822) overlap with protein fragments carrying TSA (Dostal and Schildbach, 2010).

Linkage of TraI relaxase domain to TS is necessary for effective conjugative DNA processing

The domain structures of F-like TraI proteins and the related TrwC protein of plasmid R388 enable the N-terminal relaxase activity to be separated physically from the C-terminal helicase domain without compromising stability or catalytic activity *in vitro* (Llosa *et al*., 1996; Sut *et al*., 2009). Interestingly, however, the separated protein fragments support only very low frequencies of conjugative plasmid transfer from a donor bacterium (Llosa *et al*., 1996; Byrd *et al*., 2002). A defect in localization seems unlikely. The relaxase domain binds the *nic* region of the plasmid *oriT* with high affinity (Stern and Schildbach, 2001). The relaxosome is linked to the secretion apparatus by interactions of TraM or the R388 counterpart, TrwA, with both *oriT* and the T4CP (Schwab *et al*., 1991; Disque-Kochem and Dreiseikelmann, 1997; Moncalian *et al*., 1997; Beranek *et al*., 2004; Tato *et al*., 2007; Lu *et al*., 2008). Also the helicase fragments used in earlier work carry one (TrwC) or both (TraI) TS (Llosa *et al*., 1996; Byrd *et al*., 2002; M. Llosa, pers. comm.). Thus, recruitment of the relaxase catalytic domain based on *oriT* binding and the helicase domain via interactions with the DNA and the T4CP is apparently insufficient to support conjugative DNA processing. We therefore asked whether the extension of the N-terminal TraI relaxase domain to include TSA would correct this deficiency. Conjugative transfer of R1-16Δ*traI* was complemented to wild-type levels by full-length TraI expressed from a plasmid vector (Fig. 5). As expected from earlier studies, expression of the TraI helicase individually, or combined with the relaxase fragment (N-terminal 309 residues), failed to restore efficient conjugation (≤ 0.01% wild-type activity). By contrast, concomitant expression of the helicase with a longer TraI fragment, including both the relaxase domain and TSA (N-terminal 992 residues), from the same pair of vectors complemented transfer of R1-16Δ*traI* as efficiently as full-length TraI. These findings indicate that physical linkage of the TraI catalytic domains is not required for efficient plasmid transfer. Rather, we propose that each enzymatic domain requires additional docking interactions mediated by the TS regions for productive conjugative DNA processing.

Molecular determinants of substrate discrimination

Conjugative DNA transfer of a *traI* null derivative of R1-16 or pOX38 is complemented equally well by the *traI* gene of the heterologous system (Dostal and Schildbach, 2010; data not shown). The TraI proteins of R1 and F are 97% identical and the TraI binding sites surrounding *nic* are identical in these plasmids (Ostermann *et al*., 1984; Thompson *et al*., 1989). Given that the mapped TS of each system are predicted to fold similarly we expected that

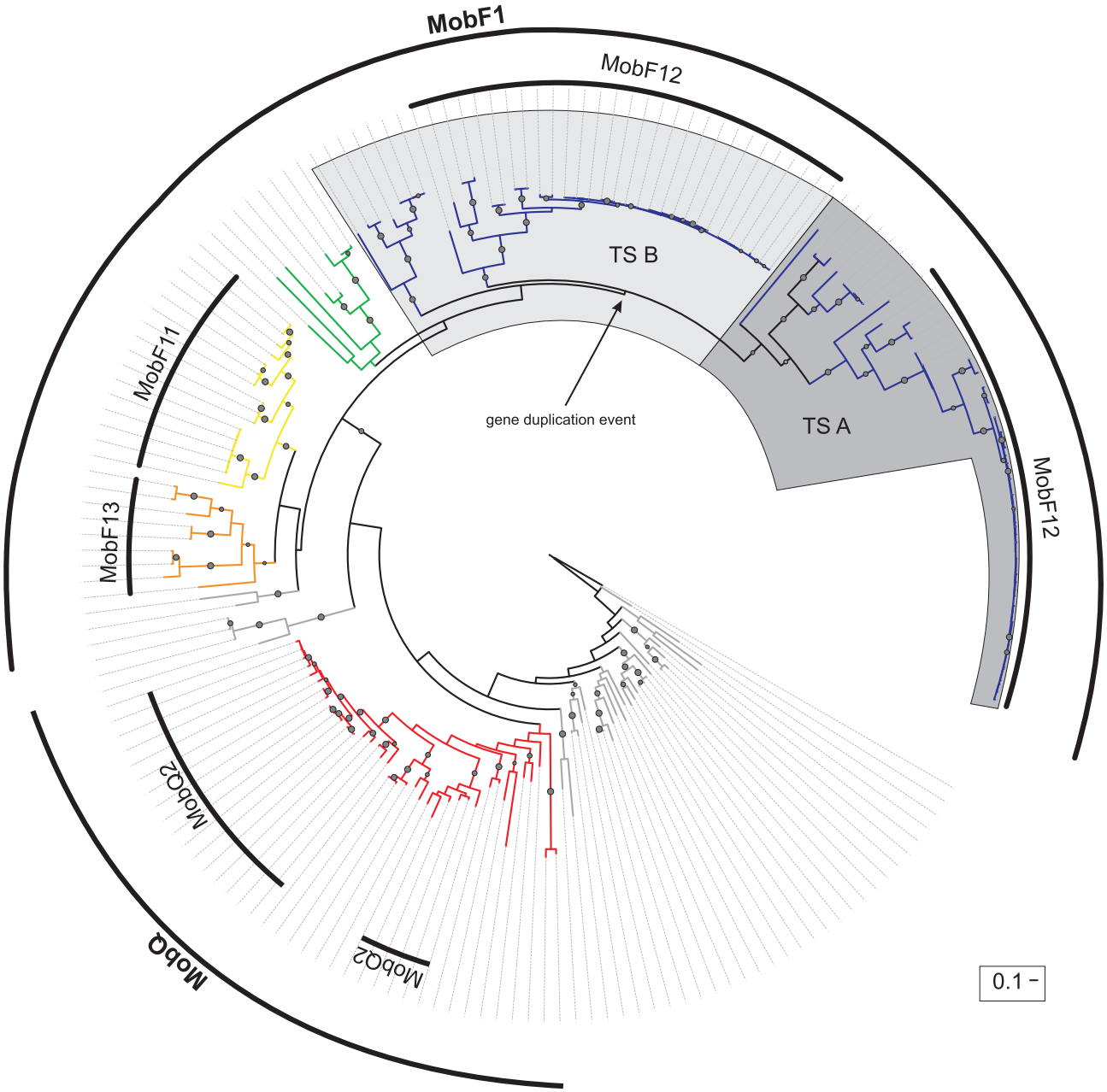

Fig. 4. Maximum likelihood tree of the helicase/TS modules. Bootstrap values ≥ 50% are indicated with circles and shown in Fig. S1. RecD2 (PDB 3e1s) was used as an outgroup. Major clades are emphasized in colour. Red, clade 1; orange, clade 2; yellow, clade 3; green, clade 4; blue, clade 5. The origin of the gene duplication creating the tandem modules is indicated with an arrow. TSA and TSB are indicated with grey boxes. Arcs designate the corresponding MOB classes (Garcillán-Barcia *et al.*, 2009).

protein translocation would also be supported by the related T4SS. Nonetheless, system-specific recognition of TraI was stringent (Fig. 6). Fusion proteins (full-length TraI or TS) of a given system were transferred at reduced (60- to 1000-fold lower) or undetectable frequencies by the heterologous system. Amino acid variation between R1 and F proteins is limited to two residues in TSA and nine in TSB. The positions of non-identical residues in R1 and F were evaluated relative to their predicted overlap with RecD2 (Fig. 3). Notably, two independent positions of heterogeneity (S757 of TSA and V1382 of TSB) colocalize relative to the RecD2 structure (Fig. 3A, yellow). This implied that mutation at the variant position would be tolerated in the substrate recognition process generally and not likely involved in the plasmid-specific discrimination for TraI observed here (Fig. 6). By contrast, the second position of natural variation in the TSA fragments of R1 and F (Fig. 3A, red) did not colocalize with heterogeneity in

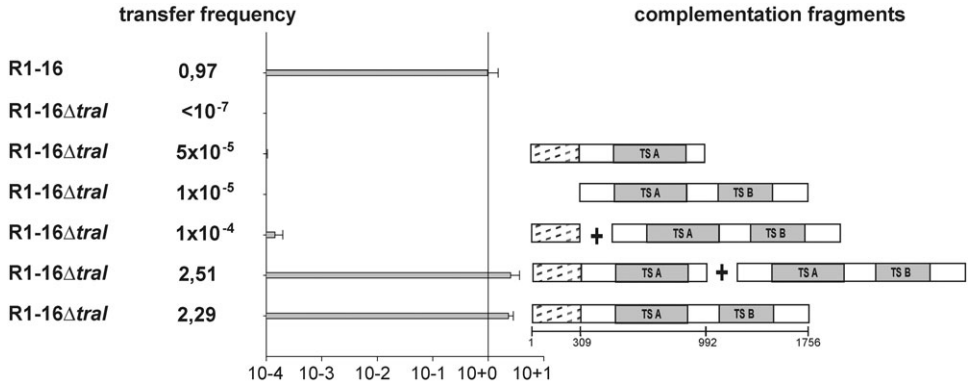

Fig. 5. Effective DNA transfer requires linkage of TraI relaxase domain to TS. Transfer frequencies are shown (left) for plasmid R1-16 or the R1-16Δ*traI* derivative when complemented with a wild-type *traI* allele or fragments thereof, as indicated schematically (right). Wild-type plasmid transfer was observed with the combination of TraI fragments $N_{1\text{-}992}$ (relaxase+TSA) and $N_{310\text{-}1756}$ (helicase). Values represent the mean of at least three experiments. Standard deviations are shown.

TSB. This residue was thus expected to contribute directly to specific recognition. To identify determinants underlying the observed substrate discrimination, we generated selected amino acid exchanges and tested their effect on protein translocation. Consistent with the model-based predictions we found that an F to R1 L626H exchange led to an F to R1 fidelity switch and a 12-fold gain of function for the heterologous R1 system (Fig. 7). This site would thus appear to represent a specificity determinant in IncF systems that cognate T4SS readily distinguish. Moreover, the TSA hybrid we generated with an F to R1 amino acid exchange S757T did not alter competence for translocation with pOX38 compared with the wild-type TSA_F (Fig. 7). Similarly, exchange E1382V in TSB_F reduced translocation

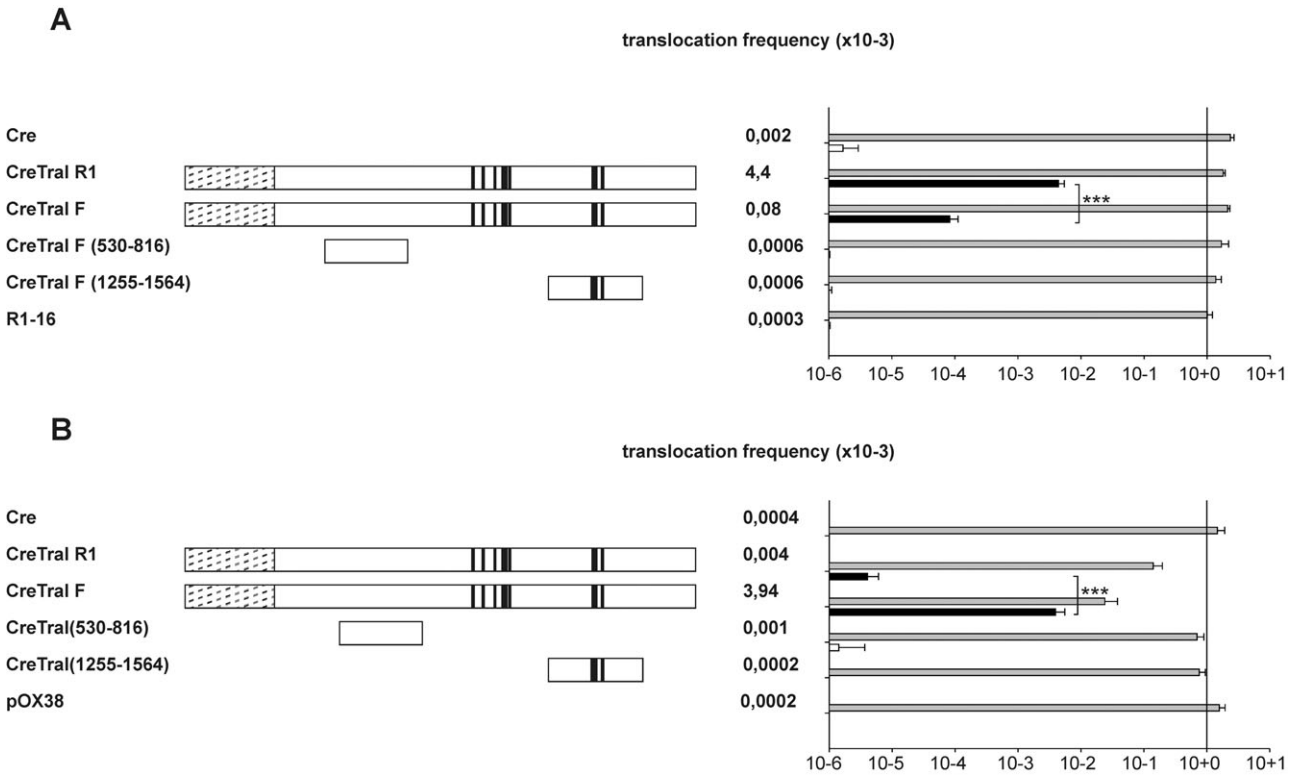

Fig. 6. IncFI and IncFII TraI TS are distinguished by the cognate T4SS. The translocation frequencies supported by R1-16 (A), or pOX38 (B) are shown (right) with schematic representations of the tested TraI fragments (left). Values represent recombination events per donor. Only black bars represent frequencies of translocation with statistical significance compared with the vector control. Conjugation frequencies are indicated with grey bars. Values represent the mean of at least three experiments. Standard deviations are shown (***$P < 0.001$).

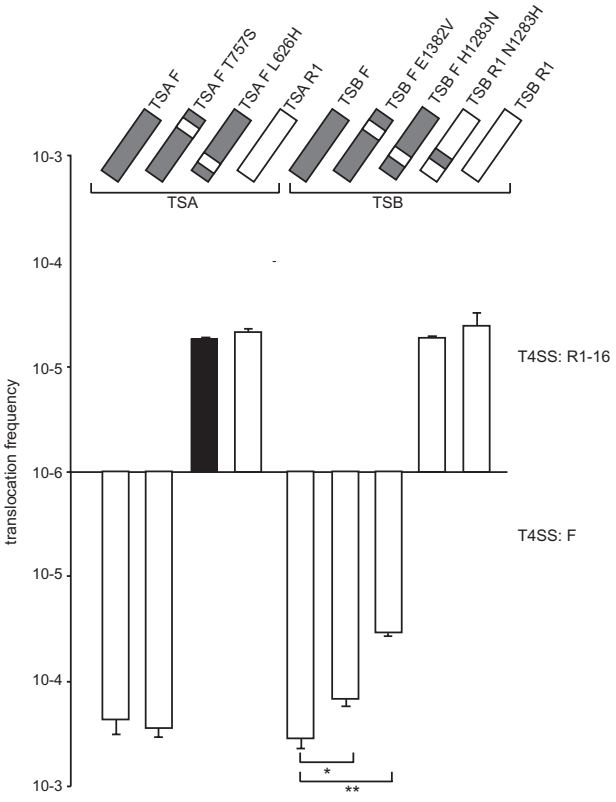

Fig. 7. Exchange of a variant TS residue switches system fidelity. The tested TraI TS fragments are illustrated above for F (grey) and R1 (white). Vertical stripes indicate a single exchange mutation. The translocation frequencies (recombination events per donor) supported by R1-16 (upper panel), or pOX38 (lower panel) are shown. No reported value indicates the absence of transfer by the heterologous system. All values shown were statistically significant compared with the vector control. The L626H exchange in TSA_F switches transfer fidelity from F to R1 (black bar). Values represent the mean of at least three experiments. Standard deviations are shown (*$P < 0.05$; **$P < 0.01$).

by pOX38 twofold but did not measurably enhance translocation by R1-16. As a final test of the proposed model of TraI protein recognition features we investigated variant residue 1283 in TSB. This position lies outside the region of predicted structural conservation (Fig. S1) and therefore may be dispensable for recognition. Consistent with that prediction, a pair of exchanges at this position were tolerated by the cognate system or led to 10-fold reduction in translocation, but did not measurably enhance activity in the heterologous system (Fig. 7).

Fidelity in protein translocation involves components of the relaxosome and/or the T4 channel

The role of the T4CP as the primary substrate receptor is well established (Cascales and Christie, 2004). Complementation of conjugative DNA transfer of *traD* null derivatives does not exhibit plasmid specificity (Willetts and Maule, 1986). We asked whether a *traD*-swap experiment would alter the fidelity of protein translocation. We used *traD* null derivatives of F and R1 to provide the T4 channel and the relaxosome. Plasmids carrying either wild-type *traD* alleles and/or the Cre–*traI* alleles were present *in trans*. We found no change in plasmid-specific discrimination of the cognate full-length TraI (Fig. 8A) based on the source of TraD.

In this experiment significantly lower frequencies of pOX*traD*411 DNA transfer were observed when a Cre–*traI* fusion allele was overexpressed *in trans* compared with Cre alone (Fig. 8B). Overexpression of the same fusion proteins *in trans* to R1-16Δ*traD* did not alter the conjugation frequency (Fig. 8A). A dominant negative effect of overexpressed F TraI on pOX38 DNA transfer was also observed in this study (Fig. 2) consistent with earlier results (Haft *et al*., 2006). The effect is not observed for R1-16 transfer (Fig. 1). An overview of pOX38 DNA transfer inhibition mediated by an excess of TraI in full-length or TS form is shown (Fig. 8C). Decreased pOX38 transfer can be observed from all donor cells expressing high levels of Cre fused to any fragment of $TraI_F$ or $TraI_{R1}$ that is competent for protein transfer. In the *traD* swap experiment we noted substantially increased transfer efficiencies for pOX*traD*411 DNA (23-fold) and Cre-$TraI_F$ protein (40-fold) when F *traD* was also expressed from a high copy vector (Fig. 8B). A similar enhancement can be observed for pOX38 transfer (sevenfold) and Cre-$TraI_F$ transfer (15-fold) with coexpression of $traD_F$ (Fig. 8C). Suppression of the negative effects of excessive TraI protein by increasing the intracellular levels of TraD emphasizes the importance of the balance of those factors. Importantly, however, coexpression of $traD_F$ failed to suppress the negative dominance of TraI from R1 [Fig. 8B and C (*arrows*)]. Also coexpression of $traD_{R1}$ and $traI_F$ did not improve transfer (Fig. 8B). These data imply that another factor important to both pOX38 transfer and protein translocation is limiting when excess TraI or TS is present. The limitation is apparently plasmid specific. Taken together, the data indicate that substrate discrimination observed for F-like plasmids is due to the T4CP and an additional component of the T4SS.

Plasmid-specific accessory proteins

F-like relaxosome components TraM and TraY bind effectively only to their cognate *oriT* (Klimke *et al*., 1998; Fekete and Frost, 2000; Zechner *et al*., 2000). Specificity on the DNA binding level presumably tags the correct plasmid for recognition by the transfer machinery. Moreover, TraM interactions have been mapped to the C-termini of both TraI and TraD (Sastre *et al*., 1998; Beranek *et al*., 2004; Ragonese *et al*., 2007; Lu *et al*., 2008). To test the prediction that TraM or TraY contributes

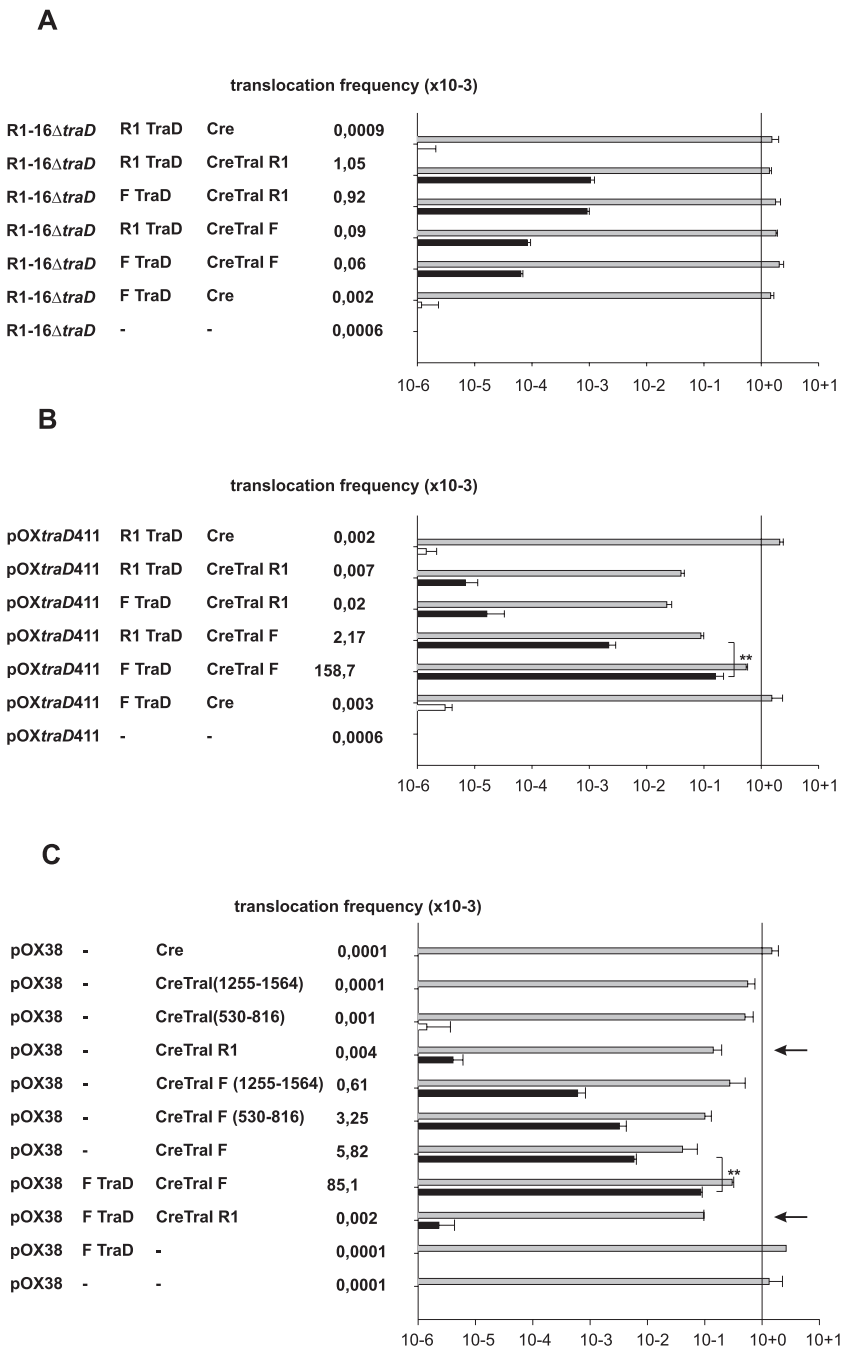

Fig. 8. Discrimination of TraI proteins involves more factors than the T4CP alone. *traD* null derivatives of R1-16 (A) and pOX38 (B) *traD* were complemented with the cognate or heterologous *traD* (left). (C) Summary of pOX38 transfer inhibition by overexpressed TraI and TS fusion proteins. The translocation frequencies of the tested TraI proteins are given as recombination events per donor (right). Bars represent frequencies of translocation with (black) or without (white) statistical significance compared with the vector control. Conjugation frequencies are indicated with grey bars. Values represent the mean of at least three experiments. Standard deviations are shown. Significant enhancement of protein transfer due to co-overexpression of $traD_F$ and $traI_F$ is shown (**$P < 0.01$). No enhancement is seen when $traD_F$ and $traI_{R1}$ are overexpressed (arrows).

to the fidelity of TraI protein recognition these F alleles were overexpressed together with the Cre–TraI fusions (Fig. 9). No change in phenotype was observed with $traM_F$ (Fig. 9A) or $traY_F$ (Fig. 9B) expressed *in trans*. Therefore, TraM and TraY proteins do not appear to be involved in competitive binding with TraI under these conditions.

Discussion

In this work we mapped two TS regions of F-like TraI proteins that are individually sufficient for transfer. The positions of TS in TraI are consistent with biochemical evidence that the T4CP stimulates both $TraI_{R1}$ relaxase and DNA helicase activities via interactions independent of the N-terminal 309 residues (Mihajlovic *et al.*, 2009; Sut *et al.*, 2009). In light of current models these positions are unorthodox in that they are neither located within the N-terminal relaxase domain nor at the protein C-terminus [(Alvarez-Martinez and Christie, 2009); Fig. 10]. They are also comparatively long. Substrates of the *A. tumefaciens* VirB T4SS are all marked for secretion by signals not longer than 50 amino acids, including relaxase

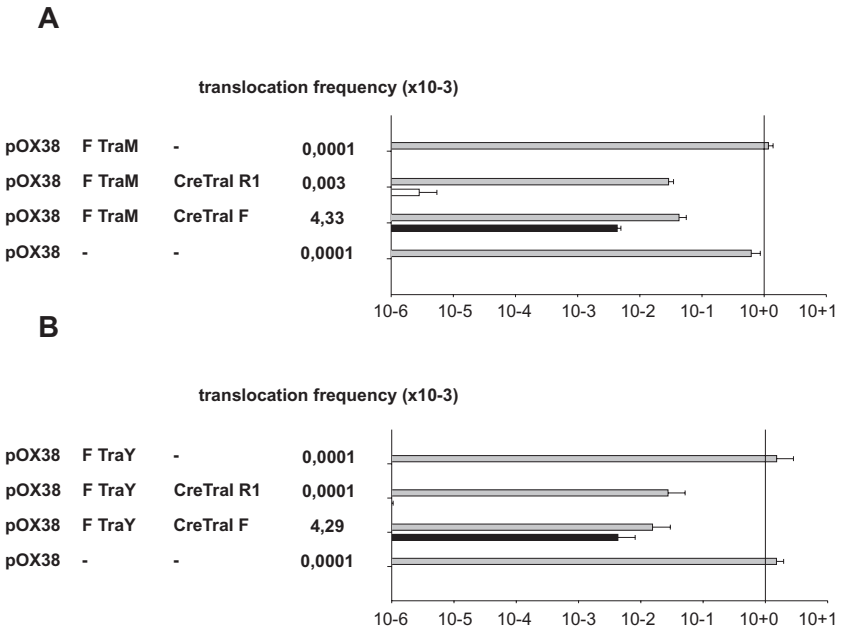

Fig. 9. Dominant negative effect of TraI F is not supressed by co-overexpression of TraM or TraY. Conjugative transfer of pOX38 (grey) and translocation frequencies of the indicated TraI fusion proteins were compared from donor cells overexpressing additionally (A) TraM or (B) TraY. Only black bars represent frequencies of translocation with statistical significance compared with the vector control. Values represent the mean of at least three experiments. Standard deviations are shown.

$MobA_{RSF1010}$. The Dot/Icm system of *L. pneumophila* transports numerous substrates via C-terminal tails ≤ 115 amino acids. In the Vir system of *Bartonella henselae*, recognition of basic C-termini is enhanced by required internal signals within the BID of the Bep substrates (Schulein *et al.*, 2005). The C-terminal signal of CagA protein is necessary but also insufficient without the additional activity of a larger N-terminal domain (Hohlfeld *et al.*, 2006). The ancestral helicase module in TraI that carries the TS underwent duplication but the pair of signals clearly did not acquire a bipartite dependence. Not only is each signal individually sufficient for transfer but moreover, the presence of one wild-type copy in the full-length protein was observed to complement the translocation defects of insertion mutations in the second TS.

The TraI TS do share features with the sig1 and sig2 TS defined for MobA relaxase (Parker and Meyer, 2007). The IncQ plasmids R1162 and the nearly identical RSF1010 can be translocated by a number of different conjugative and effector T4SS. A MobA TS recognized by the *A. tumefaciens* VirB system was first defined as a cluster of positive charge at the C-terminus (Vergunst *et al.*, 2005). By contrast, the P-like transfer system specified by plasmid R751 required much larger MobA peptide fragments exhib-

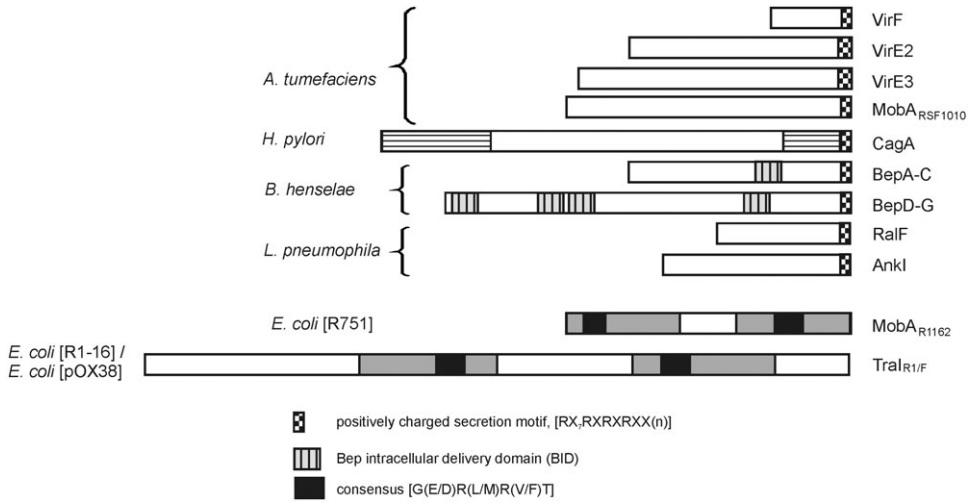

Fig. 10. Experimentally determined recognition domains for T4SS substrates. Proteins examined for recognition and translocation features are illustrated. The test T4SS for effective translocation is identified on the left. Where investigated, peptide fragments assigned an independent TS (grey) or bipartite TS function are shown (stripes, *BID*). Conserved signals are described (below) including the signature motif of the F-like TraI proteins. This consensus is present in both mapped TS of $MobA_{R1162}$.

iting comparably more secondary structure for effective recruitment and translocation (Parker and Meyer, 2007). The two signals, sig1 and sig2, mapped by these authors were also dependent on the accessory factor MobB.

Fold recognition using GenTHREADER (Jones, 1999; McGuffin and Jones, 2003) predicted that the F-like TS share structural properties most closely resembling a portion of the three dimensional structure of helicase RecD2 of *D. radiodurans*. Bioinformatic analyses of the evolution of the TS of TraI and other relaxases reveal that an ancient fusion of RecD2 to the 3′ end of a progenitor gene took place. Accordingly, conservation of a helicase/TS module was detected in different families of relaxases defined previously (Garcillán-Barcia *et al*., 2009). The current data indicate also that this early fusion was followed in the MOB_{F12} group by a later internal duplication event linking a tandem copy of the TS/helicase module downstream. Based on what we know for MOB_{F12} TraI proteins of F and R1 (Dostal and Schildbach, 2010), one of the modules (TSB) has retained helicase activity whereas the other (TSA) has retained just ssDNA binding activity associated with DNA unwinding. Usually, a functional loss reduces the constraint for conservation, resulting in a much faster diversification (Birschmann *et al*., 2005). Given that diversification of domain TSA remained modest, the additional task of mediating recognition interactions with the T4CP and/or other T4 components is apparently demanding. That conclusion is supported by the demonstration that separate fragments of TraI each carrying just one catalytic domain necessary for conjugative DNA processing need not be physically linked as long as each carries additionally at least one TS region. By contrast, the TrwC relaxase-helicase of R388 belongs to MOB_{F11} and contains a single conserved TS module. We predict that physical separation of the helicase and relaxase activities is not tolerated in that system due to the lack of a second TS.

Plasmid R1 is similar enough to F that both the heterologous T4CP and the full-length TraI proteins support efficient conjugative transfer of the cognate relaxosome. Although the DNA sequences at and around *nic* that are required for relaxase recognition are identical in F and R1, the remaining *oriT* sequences diverge such that productive interactions of auxiliary relaxosome factors TraM and TraY are plasmid specific (Everett and Willetts, 1980; Willetts and Maule, 1986; Fekete and Frost, 2000). Natural variation in the primary sequences of the full-length TraI of F and R1 revealed a stringent plasmid specificity for protein translocation. We exploited the plasmid-specific aspects of the related systems to follow two important lines of investigation. Extrapolation of the structural data for RecD2 to the F-like TS followed by single residue exchange mutations revealed that position 626 fills a particularly key function in the fidelity of substrate selection. The L626H exchange for this molecular determinate resulted in a fidelity switch and gain of function for the heterologous R1 system. Other positions of natural heterogeneity were tolerated by both systems.

The hypersensitive plasmid-specific responses to amino acid variation in TS appear to be related to the relaxosome. A single mutation that lowers the affinity of TraI TS and an interaction partner may still be compatible with function if TraI is integrated in the relaxosome. In that case many additional protein–protein and protein–DNA interactions would contribute to stable and productive contacts with the T4CP. In the CRAfT assay wild-type TraI protein encoded by the conjugative plasmid is always present to support normal conjugative DNA processing. This is necessary since the majority of TraI truncations should disrupt their activity as enzymes. We observed generally lower protein translocation efficiencies as the fusion proteins were shortened. It is reasonable to propose that as the length of TraI is decreased, it will be less likely to be able to support a normal range of protein and DNA interactions. Creating hybrid TS $_{F/R1}$ in the minimum length fragment thus resulted in a hypersensitive response and a comparably clear view of structure and function within that determinant.

In support of this explanation the data also indicate that other factors in addition to substrate receptor TraD are involved in substrate recognition. This possibility has been suggested before (Beranek *et al*., 2004; Lu *et al*., 2008). Raising the concentration of predicted interaction partners TraM and TraY had no effect in our experiments. At present this aspect of the recognition and recruitment pathway remains unclear. Accessory factors with various secretion chaperone-like mechanisms of action are known for T3 and T4SS. These act via direct interactions to stabilize the substrates, prevent aggregation, induce productive docking conformations in the substrate proteins and/or help traffic the substrates to the secretion apparatus assembly sites (Parsot *et al*., 2003; Ghosh, 2004; Alvarez-Martinez and Christie, 2009).

In current models interactions between relaxosome and T4CP are also expected to control conjugative DNA processing (as reviewed in detail; de la Cruz *et al*., 2010). Importantly, knowledge of the locations of the TraI TS provides new options to investigate that regulation. We reconstituted effective DNA transfer for the first time with separate TraI catalytic domains by extending the relaxase-containing fragment to include TSA. In these complementation experiments TSA and TSB are both present on the helicase fragment of TraI (residues 309–1756). The relaxase domain alone (residues 1–309), or the domain with TSA, bind to DNA at *nic*. Therefore, we expect that both the catalytic domains should be present at the relaxosome /T4CP interface whether or not the

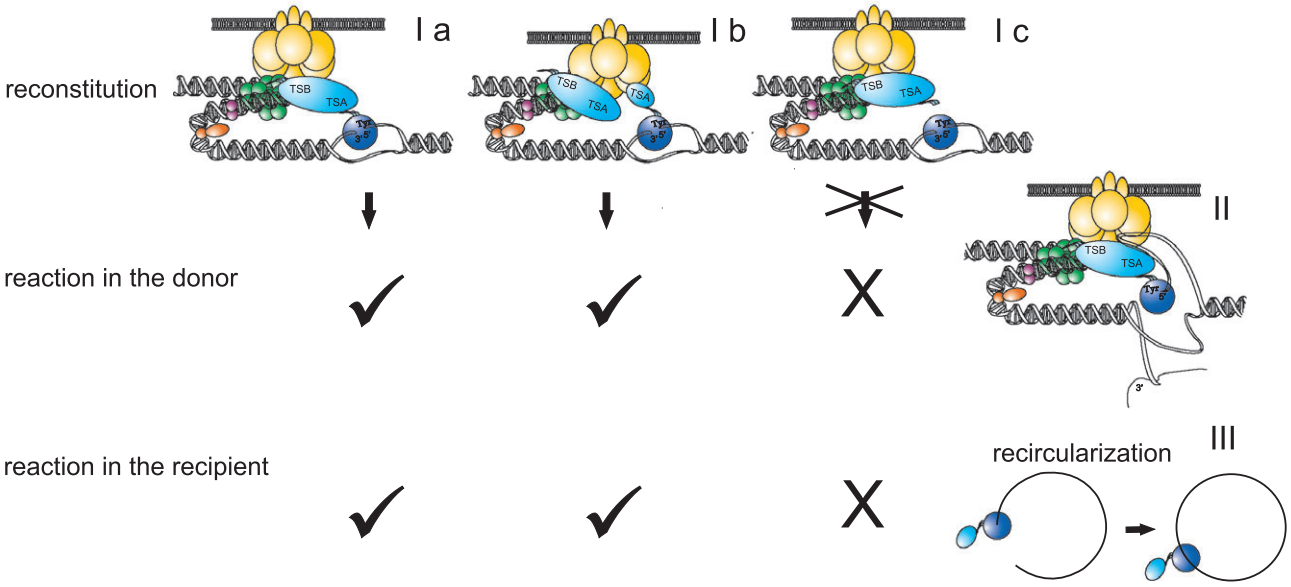

Fig. 11. Reconstitution of conjugative DNA processing requires a covalently linked relaxase-TSA. Wild-type TraI (Ia) and the combination of helicase domain plus a relaxase-TSA fragment (Ib) may support essential stages of DNA processing reactions in the donor (II). In that case, TSA mediated control of relaxase domain is necessary for progression. Transmission of these regulatory cues does not occur if relaxase is severed from TSA (Ic). In a second possibility, productive DNA processing may be terminated in recipient cells (III) only when relaxase is present there. In that case, transfer of the truncated relaxase domain in reconstitution Ic is apparently insufficient for recircularization and stabilization of the plasmid DNA. Protein components are illustrated as follows: the bifunctional TraI relaxase (blue) with active site tyrosine (Tyr) and TraI helicase domain (light blue), IHF (orange), TraY (purple), TraM (green) and the T4CP TraD (yellow).

relaxase fragment is also linked to a TS. Productive conjugative DNA processing occurred only in the latter case. This finding suggests two possible explanations that are not necessarily exclusive (Fig. 11). In the first scenario, the reaction cascade controlling T-strand production and delivery to the T4CP requires TS-mediated docking contacts with recognition factors at the conjugative pore (Fig. 11, Ia, Ib and Ic). The key link in effective reconstitution physically joins TraI relaxase activity and TSA (Fig. 11, Ia and Ib). Regulatory signals may be conveyed over TSA to the relaxase domain to trigger a key step, e.g. to permit productive loading of the DNA helicase at *nic*. In this model a truncated relaxase domain that is disconnected from a TS module (Fig. 11, Ic) cannot be activated to produce the necessary ssDNA loading platform or to permit access to that ssDNA as required for TraI helicase activation. Conjugative DNA processing is blocked in the donor cells in that case (Fig. 11, II). In the second possibility, recircularization of the transferred T-strand is catalysed by the relaxase domain of TraI in recipient cells (Fig. 11, III). The link of TSA and relaxase performs the simple task of ensuring that sufficient protein reaches the recipient to perform efficiently. Both explanations may apply. Importantly, advances in the functional map of TraI and insights to the structure of a T4 secretion signal provide new tools to investigate both conjugative DNA processing and selective uptake by the secretion channel.

Experimental procedures

Strains and plasmids

All *E. coli* K-12 strains used in this study are described in Table 2. Plasmids are described in Table S2.

DNA preparation and PCR amplification

Plasmid DNA was purified from *E. coli* cells with the QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany). Restriction endonucleases and DNA modifying enzymes were obtained from Fermentas GmbH (St. Leon-Rot, Germany). DNA fragments for cloning were amplified using Phusion High-Fidelity DNA Polymerase (Finnzymes Oy, Espoo, Finland) or the Taq-Polymerase (New England Biolabs, Beverly, MA, USA). The correct DNA sequence of all amplified fragments used in cloning was verified. Enzymes were used according to manufacturers' recommendations.

Construction of traI *and* traD *null derivatives*

Primer sequences are shown in Table S2. To generate pOX38Δ*traI*, R1-16Δ*traI* and R1-16Δ*traD*, the primer pairs ar042 and DtraIDcas, UtraIUcas and DtraIDcas, or TraDUcas_fw and TraDDcas_rev were used to amplify a TetRA-resistance cassette from pAR183 (Reisner *et al.*, 2006). The amplified fragments were introduced into *E. coli* DY330 [pOX38] or *E. coli* DY330 [R1-16] and integrated via homolo-

gous recombination as previously described (Reisner *et al.*, 2002).

Construction of expression plasmids

Cre fusion plasmids were constructed by inserting PCR-amplified DNA between the KpnI or KpnI/SalI sites of the CFP B plasmid (Parker and Meyer, 2007). The R1 *traI* inserts were PCR amplified from R1-16, pHP2, p99I+ or p99traI derivates. CreTraI F derivatives were constructed by ligating the KpnI/SalI fragment from p99I+ or a p99traI derivative with CFP B. To generate streptomycin (Sm) resistant derivatives CFP B Sm, CreTraI(3-1756) Sm and CreTraI F Sm the *aadA* gene from pAH144 (Haldimann and Wanner, 2001) was amplified with primers FW_CFPSm and Rev_CFPSm. This fragment was introduced into *E. coli* DY330 harbouring CFP B, CreTraI(3-1756) or CreTraI F to replace the *bla* gene via homologous recombination as previously described (Reisner *et al.*, 2002). Two-step PCR was used to generate all R1/F hybrid TS plasmids. In the first step primer sets 1 and 2 (Table S2) were used to amplify two fragments, which both carried the desired point mutation. In the second step these two fragments were annealed and amplified with primer set 3. The inserts for pMSTraD_wt, pMSTraMF, pMSTraY, pTraIrel and pCG02 were amplified with the indicated primers (Table S2) from R1-16 or pOX38 and ligated with pMS119EH or pGZ119EH.

Construction of recipient strains

The indicator *E. coli* strains CSH26Cm::LKL and CSH26Cm::LTL, used in the CRAfT experiments, were created in sequential steps. To generate *E. coli* CSH26crt, the *cat::resC-tet-resC* cassette of *E. coli* K10636 was inserted into *galK* of CSH26 via P1 phage transduction. Primers (Table 2) Fw_loxPCm and Rev_loxPCm were used to amplify the *loxP*-Km-*loxP* cassette from *E. coli* TOP10::LKL. The 5′-overlapping termini of this primer pair contain sequence homologous to the *cat* gene of *E. coli* CSH26crt. The temperature-sensitive pKD46 was introduced into CSH26crt and the *resC-tet-resC* cassette was replaced with the *loxP*-Km-*loxP* cassette via homologous recombination as described (Datsenko and Wanner, 2000). The same strategy was used to create CSH26Cm::LTL using targeting DNA made with primers Fw_loxPTc and Rev_loxPTc and template pAR183.

MS614Cm is a chloramphenicol-resistant derivative of MG1655Str (Miranda *et al.*, 2004) that was created by bacteriophage λ integrase-mediated insertion of pAR185 in $attB_{\lambda}$ using the CRIM helper plasmid pInt-ts as described previously (Haldimann and Wanner, 2001). pAR185 carries a functional *cat* gene and is a derivative of pCAH63 (Haldimann and Wanner, 2001) lacking the *Pst*I fragment encoding the 3′ end of the *uidAF* gene.

Triparental mating assay

An overnight culture of donor *E. coli* MS411 carrying a conjugative plasmid and *traI* expression vector (Table 1, strain 1) was subcultured in drug-free LB for 1 h at 37°C to an A_{600} 0.02. A 10-fold excess of both recipient strains *E. coli* MS411 carrying mobilizable plasmid pMM-M0 (strain 2) and *E. coli* MS614 (strain 3) was added and the mixture incubated for 2.5 h at 37°C without shaking. DNA transfer was stopped by vortexing for 1 min and rapid cooling on ice. Transconjugants were selected by plating serial dilutions on LB agar plates with ampicillin (100 μg ml^{-1}) and streptomycin (25 μg ml^{-1}) or kanamycin (40 μg ml^{-1}) and streptomycin (25 μg ml^{-1}). Donor cell counts were determined with kanamycin (40 μg ml^{-1}) and chloramphenicol (10 μg ml^{-1}) selection.

CRAfT (Cre recombinase assay for translocation)

The Cre fusion reporter assay (Vergunst *et al.*, 2000) was applied with *E. coli* MS614 harbouring pOX38 or R1-16 to provide the respective T4SS. T4-mediated plasmid transfer and secretion of the fusion protein occurred under the conditions described for mating above, except that a single recipient, CSH26Cm::LTL, was present. Transconjugants were identified by plating serial dilutions on LB agar with kanamycin (40 μg ml^{-1}) and X-Gal (50 μg ml^{-1}). Recombinants created by uptake of the Cre-protein fusion were selected on plates containing chloramphenicol (10 μg ml^{-1}). Donors were selected on kanamycin (40 μg ml^{-1}) or kanamycin (40 μg ml^{-1}) and ampicillin (100 μg ml^{-1}). The conjugation and protein translocation frequencies are calculated as transconjugants or recombinants per donor respectively. Prior to mating experiments, a functional test for Cre recombinase and indirectly, the stability of each fusion protein was performed by transforming the recipient strain with plasmid DNA and selecting for chloramphenicol resistant recombinants.

The complementation assays for the *traI* and *traD* null mutants were performed similarly, except that *E. coli* MS411 carrying the plasmids of interest was used as donor. Conjugative transfer in this case was measured in a parallel experiment using MS614Cm as recipient and selection of transconjugants on chloramphenicol (10 μg ml^{-1}) and kanamycin (40 μg ml^{-1}). The presence of all plasmids in the donor was confirmed by PCR before and after the CRAfT was performed.

Computational annotation of TS domains and phylogenetic reconstruction

Putative occurrences of helicase/TS domains were determined by PSI-BLAST (Altschul *et al.*, 1997). The TSA and TSB amino acid subsequences of TraI were searched in the NCBI non-redundant database (Sayers *et al.*, 2009) using five iterations and an e-value threshold of 1e-04. From the resulting hits, a position specific scoring matrix was built. This matrix has been used to screen a data set of conjugative relaxases compiled from the review of Garcillán-Barcia *et al.* (2009) with one iteration of PSI-BLAST using the same e-Value cut-off. For all resulting local alignments, the subsequences corresponding to the aligned regions of the database proteins were retrieved from the database. A multiple alignment of these sequences was calculated using MUSCLE (Edgar, 2004). A phylogenetic tree was reconstructed by RAXML (Stamatakis, 2006) using the PROTGAMMAJT model and the fast bootstrap algorithm. For estimating

the statistical support of branches, 1000 bootstrap replicates were calculated. The most likely tree and the bootstrap values were visualized using the iTOL server (Letunic and Bork, 2007).

Acknowledgements

This work is supported by the Austrian FWF through grants P18607-B12 and W901-B05 DK: Molecular Enzymology, by the German DFG through grant RA 1719/1-1 and by ERA-NET PathoGenoMics project PATHOMICS. R.J. Meyer, L. Frost and B. Traxler are gratefully acknowledged for providing plasmids or strains. We also thank M. Mally, K. Schilcher, A. Wöran and M. Balog for their contributions to this study.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73:** 775–808.

Backert, S., and Meyer, T.F. (2006) Type IV secretion systems and their effectors in bacterial pathogenesis. *Curr Opin Microbiol* **9:** 207–217.

Beranek, A., Zettl, M., Lorenzoni, K., Schauer, A., Manhart, M., and Koraimann, G. (2004) Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. *J Bacteriol* **186:** 6999–7006.

Birschmann, I., Rosenkranz, K., Erdmann, R., and Kunau, W.H. (2005) Structural and functional analysis of the interaction of the AAA-peroxins Pex1p and Pex6p. *FEBS J* **272:** 47–58.

Byrd, D.R., Sampson, J.K., Ragonese, H.M., and Matson, S.W. (2002) Structure-function analysis of *Escherichia coli* DNA helicase I reveals non-overlapping transesterase and helicase domains. *J Biol Chem* **277:** 42645–42653.

Cascales, E., and Christie, P.J. (2004) Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science* **304:** 1170–1173.

Chen, I., Christie, P.J., and Dubnau, D. (2005) The ins and outs of DNA transfer in bacteria. *Science* **310:** 1456–1460.

Christie, P.J. (2001) Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Mol Microbiol* **40:** 294–305.

Christie, P.J., and Vogel, J.P. (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol* **8:** 354–360.

de la Cruz, F., and Davies, J. (2000) Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol* **8:** 128–133.

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.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97:** 6640–6645.

Disque-Kochem, C., and Dreiseikelmann, B. (1997) The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD in vitro. *J Bacteriol* **179:** 6133–6137.

Dostal, L., and Schildbach, J.F. (2010) Single-stranded DNA binding by F TraI relaxase and helicase domains is coordinately regulated. *J Bacteriol* **192:** 3620–3628.

Draper, O., Cesar, C.E., Machon, C., de la Cruz, F., and Llosa, M. (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc Natl Acad Sci USA* **102:** 16385–16390.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 113.

Everett, R., and Willetts, N. (1980) Characterisation of an in vivo system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J Mol Biol* **136:** 129–150.

Fekete, R.A., and Frost, L.S. (2000) Mobilization of chimeric oriT plasmids by F and R100-1: role of relaxosome formation in defining plasmid specificity. *J Bacteriol* **182:** 4022–4027.

Fischer, W., Haas, R., and Odenbreit, S. (2002) Type IV secretion systems in pathogenic bacteria. *Int J Med Microbiol* **292:** 159–168.

Franco, I.S., Shuman, H.A., and Charpentier, X. (2009) The perplexing functions and surprising origins of *Legionella pneumophila* type IV secretion effectors. *Cell Microbiol* **11:** 1435–1443.

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.

Ghosh, P. (2004) Process of protein transport by the type III secretion system. *Microbiol Mol Biol Rev* **68:** 771–795.

Gomis-Rüth, F.X., Moncalián, G., Pérez-Luque, R., González, A., Cabezón, E., de la Cruz, F., and Coll, M. (2001) The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* **409:** 637–641.

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.

Haft, R.J., Palacios, G., Nguyen, T., Mally, M., Gachelet, E.G., Zechner, E.L., and Traxler, B. (2006) General mutagenesis of F plasmid TraI reveals its role in conjugative regulation. *J Bacteriol* **188:** 6346–6353.

Haldimann, A., and Wanner, B. (2001) Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J Bacteriol* **183:** 6384–6393.

Hamilton, H.L., Dominguez, 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.

Hohlfeld, S., Pattis, I., Puls, J., Plano, G.V., Haas, R., and Fischer, W. (2006) A C-terminal translocation signal is necessary, but not sufficient for type IV secretion of the *Helicobacter pylori* CagA protein. *Mol Microbiol* **59:** 1624–1637.

Jones, D.T. (1999) GenTHREADER: an efficient and reliable

protein fold recognition method for genomic sequences. *J Mol Biol* **287:** 797–815.
Juhas, M., Crook, D.W., and Hood, D.W. (2008) Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol* **10:** 2377–2386.
Klimke, W.A., Fekete, R., Manchak, J., Anthony, K., and Frost, L.S. (1998) Plasmid specificity and interaction: the similarities and differences between the transfer regions of two compatible plasmids, F and R100-1. In *Horizontal Gene Transfer*. Syvanen, M., and Kado, C.I. (eds). London: Chapman and Hall, pp. 25–39.
Letunic, I., and Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23:** 127–128.
Llosa, M., Grandoso, G., Hernando, M.A., and de la Cruz, F. (1996) Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J Mol Biol* **264:** 56–67.
Llosa, M., Gomis-Rüth, F.X., Coll, M., and de la Cruz, F. (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* **45:** 1–8.
Lu, J., Wong, J.J., Edwards, R.A., Manchak, J., Frost, L.S., and Glover, J.N. (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70:** 89–99.
McGuffin, L.J., and Jones, D.T. (2003) Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* **19:** 874–881.
Mihajlovic, S., Lang, S., Sut, M.V., Strohmaier, H., Gruber, C.J., Koraimann, G., *et al*. (2009) Plasmid R1 conjugative DNA processing is regulated at the coupling protein interface. *J Bacteriol* **191:** 6877–6887.
Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
Miranda, R.L., Conway, T., Leatham, M.P., Chang, D.E., Norris, W.E., Allen, J.H., *et al*. (2004) Glycolytic and gluconeogenic growth of *Escherichia coli* O157:H7 (EDL933) and *E. coli* K-12 (MG1655) in the mouse intestine. *Infect Immun* **72:** 1666–1676.
Moncalian, G., Grandoso, G., Llosa, M., and de la Cruz, F. (1997) *oriT*-processing and regulatory roles of TrwA protein in plasmid R388 conjugation. *J Mol Biol* **270:** 188–200.
Nagai, H., Cambronne, E.D., Kagan, J.C., Amor, J.C., Kahn, R.A., and Roy, C.R. (2005) A C-terminal translocation signal required for Dot/Icm-dependent delivery of the Legionella RalF protein to host cells. *Proc Natl Acad Sci USA* **102:** 826–831.
Ostermann, E., Kricek, F., and Hogenauer, G. (1984) Cloning the origin of transfer region of the resistance plasmid R1. *EMBO J* **3:** 1731–1735.
Parker, C., and Meyer, R.J. (2007) The R1162 relaxase/primase contains two, type IV transport signals that require the small plasmid protein MobB. *Mol Microbiol* **66:** 252–261.
Parsot, C., Hamiaux, C., and Page, A.L. (2003) The various and varying roles of specific chaperones in type III secretion systems. *Curr Opin Microbiol* **6:** 7–14.
Pattis, I., Weiss, E., Laugks, R., Haas, R., and Fischer, W. (2007) The *Helicobacter pylori* CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein. *Microbiology* **153:** 2896–2909.
Ragonese, H., Haisch, D., Villareal, E., Choi, J.H., and Matson, S.W. (2007) The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at *oriT* through an interaction with TraI. *Mol Microbiol* **63:** 1173–1184.
Reisner, A., Molin, S., and Zechner, E.L. (2002) Recombinogenic engineering of conjugative plasmids with fluorescent marker cassettes. *FEMS Microbiol Ecol* **42:** 251–259.
Reisner, A., Höller, B.M., Molin, S., and Zechner, E.L. (2006) Synergistic effects in mixed *E. coli* biofilms: conjugative plasmid transfer drives biofilm expansion. *J Bacteriol* **188:** 3582–3588.
Rocha, E.P., Cornet, E., and Michel, B. (2005) Comparative and evolutionary analysis of the bacterial homologous recombination systems. *PLoS Genet* **1:** e15.
Saikrishnan, K., Griffiths, S.P., Cook, N., Court, R., and Wigley, D.B. (2008) DNA binding to RecD: role of the 1B domain in SF1B helicase activity. *EMBO J* **27:** 2222–2229.
Sastre, J.I., Cabezón, E., and de la Cruz, F. (1998) The carboxyl terminus of protein TraD adds specificity and efficiency to F-plasmid conjugative transfer. *J Bacteriol* **180:** 6039–6042.
Sayers, E.W., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., *et al*. (2009) Database resources of the national center for biotechnology information. *Nucleic Acids Res* **37:** D5–15.
Schmid, M.C., and Scheidegger, F. (2006) A translocated bacterial protein protects vascular endothelial cells from apoptosis. *PLoS Pathog* **2:** e115.
Schröder, G., and Lanka, E. (2003) TraG-like proteins of type IV secretion systems: functional dissection of the multiple activities of TraG (RP4) and TrwB (R388). *J Bacteriol* **185:** 4371–4381.
Schulein, R., Guye, P., Rhomberg, T.A., Schmid, M.C., Schröder, G., Vergunst, A.C., *et al*. (2005) A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. *Proc Natl Acad Sci USA* **102:** 856–861.
Schwab, M., Gruber, H., and Högenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol Microbiol* **5:** 439–446.
Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.
Stern, J.C., and Schildbach, J.F. (2001) DNA recognition by F factor TraI36: highly sequence-specific binding of single-stranded DNA. *Biochemistry* **40:** 11586–11595.
Sut, M.V., Mihajlovic, S., Lang, S., Gruber, C.J., and Zechner, E.L. (2009) Protein and DNA effectors control the TraI conjugative helicase of plasmid R1. *J Bacteriol* **191:** 6888–6899.
Tato, I., Matilla, I., Arechaga, I., Zunzunegui, S., de la Cruz, F., and Cabezón, E. (2007) The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J Biol Chem* **282:** 25569–25576.
Thompson, T.L., Centola, M.B., and Deonier, R.C. (1989) Location of the nick at oriT of the F plasmid. *J Mol Biol* **207:** 505–512.
Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2002) Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* Chapter 2: Unit 2.3.

Vergunst, A.C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C.M., Regensburg-Tuink, T.J., and Hooykaas, P.J. (2000) VirB/D4-dependent protein translocation from Agrobacterium into plant cells. *Science* **290:** 979–982.

Vergunst, A.C., van Lier, M.C., den Dulk-Ras, A., Stuve, T.A., Ouwehand, A., and Hooykaas, P.J. (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. *Proc Natl Acad Sci USA* **102:** 832–837.

Wang, J., and Julin, D.A. (2004) DNA helicase activity of the RecD protein from *Deinococcus radiodurans*. *J Biol Chem* **279:** 52024–52032.

Willetts, N., and Maule, J. (1986) Specificities of IncF plasmid conjugation genes. *Genet Res* **47:** 1–11.

Willetts, N., and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol Rev* **48:** 24–41.

Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., *et al*. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res* **17:** 3469–3478.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA* **97:** 5978–5983.

Zechner, E.L., de la Cruz, F., Eisenbrandt, R., Grahn, A.M., Koraimann, G., Lanka, E., *et al*. (2000) Conjugative DNA transfer processes. In *The Horizontal Gene Pool*. Thomas, C.M. (ed.). Amsterdam: Harwood Academic Publishers, pp. 87–174.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.