Trans-kingdom DNA transfer: Comparing the Ti and RP4 systems

Wittleben, F.T.

Chapter 4

Identification of a C-terminal translocation signal in the TraI relaxase protein of the RP4 plasmid

Felix T. Wittleben, Silvia Lang, Paul J.J. Hooykaas and Bert J. van der Zaal

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Abstract

The conjugative plasmid RP4 with its type-IV secretion system ($T4SS_{RP4}$) is widely used in the laboratory to transfer DNA constructs between organisms. These DNA molecules are transferred to recipient cells in a single-strand DNA form with the RP4 relaxase $TraI_{RP4}$ covalently attached to the 5'-end. Although the main molecular elements of this conjugative transfer system have been described, no adequate descriptions of potential translocation signals in the relaxase have been published. In this study, the translocation signal (TS) of the DNA present in $TraI_{RP4}$ was identified using a modified Cre Reporter Assay for Translocation (CRAfT). The TS was determined to be located in the very C-terminus of the $TraI_{RP4}$ relaxase. The TS shares similarity with several previously published TS used by related T4SS; it has characteristic arginine residues and a similar hydropathic profile. The addition of the isolated $TraI_{RP4}$ translocation signal to a heterologous protein was sufficient to facilitate their delivery via the $T4SS_{RP4}$, from *Escherichia coli* to *E. coli*, and from *Agrobacterium tumefaciens* to *E. coli* even in the absence of conjugal DNA transfer.

Introduction

Conjugation allows the transfer of genetic material between bacteria. One type of conjugational transfer system employs the type-IV secretion system (T4SS) as a conduit. The T4SS can transfer DNA and (effector) proteins to other bacteria and, in some cases to eukaryotes. DNA molecules are transferred because they are covalently bound by a relaxase protein at their 5'-end (van Kregten *et al.*, 2009). In general, a conjugational system consists of the following components: a *trans*-membrane spanning protein complex for mate pair formation (Mpf) and facilitating nucleoprotein transfer, most commonly the T4SS, ending in an extracellular pilus structure; a DNA processing machinery (Dtr) encompassing a relaxase/nickase facilitating plasmid processing leading to formation of a transferred nucleoprotein complex; and the 'gatekeeper' between both components, the coupling protein (CP) binding both to the substrates as well as the Mpf/T4SS and thus selecting the substrates that are transferred. The CP recognizes a translocation signal (TS) located within the relaxase and other effector proteins that are transferred. Conjugation systems may have evolved from protein translocation systems which use a T4SS and CP for protein translocation, but lack the Dtr machinery (Alvarez-Martinez & Christie 2009; Grohman *et al.*, 2018).

One of the more thoroughly studied transfer systems is the inducible $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ system found in the natural plant genetic engineer *Agrobacterium tumefaciens*, encoded on its tumor-inducing (Ti) plasmid (Hooykaas & Beijersbergen 1994; Alvarez-Martinez & Christie, 2009). Part of the Ti plasmid, called transfer DNA (T-DNA), is transferred via this $CP_{Ti}/T4SS_{Ti}$ system into recipient plant cells. The proteins making up the $T4SS_{Ti}$ are encoded by 11 virulence B genes ($virB_{Ti}$). Induction and function of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ system are triggered by specific environmental conditions. In particular, vir_{Ti} gene expression is triggered by phenolic compounds like acetosyringone and mediated by a $VirA_{Ti}/VirG_{Ti}$ 2-component system (Brencic & Winans, 2005). Besides low nutrient availability, temperatures around 20° C and a slightly acidic pH level are important (Melchers *et al.*, 1989; Turk *et al.*, 1991). The T-DNA is defined by specific sequences called the right border (RB_{Ti}) and left border (LB_{Ti}). The relaxase $VirD2_{Ti}$, part of the inducible Dtr_{Ti} system, can recognize the border sequences and create nicks there, which is thought then to lead to displacement of one single-strand of the T-DNA called the T-strand. During the nicking reaction, $VirD2_{Ti}$ remains bound to the 5'-end at the RB_{Ti}. The nucleoprotein complex (T-strand with bound relaxase), as well as other so-called effector Vir_{Ti} proteins, are then recognized by the CP_{Ti} $VirD4_{Ti}$ and transferred via the $T4SS_{Ti}$ into the recipient cell. The T-strand bound to $VirD2_{Ti}$ with the aid of the effector Vir_{Ti} protein $VirE2_{Ti}$ is then imported via the host endogenous importing system into the nucleus (Ziemienowicz *et al.*, 2001). In the nucleus the T-strand is randomly integrated into the host genome largely determined by host endogenous factors (Bundock *et al.*, 1995). The T-DNA encoded genes converting the plant cells into tumor cells, starting unregulated cellular growth and synthesis of specific compounds called opines for the survival of *A. tumefaciens*, in effect causing disease within the affected organism (Escobar & Dandekar, 2003). The Ti plasmid is also able to mediate T-DNA transfer to other eukaryotes, for example, filamentous fungi (de Groot *et al.*, 1998) and *Saccharomyces cerevisiae* (Bundock *et al.*, 1995).

A mechanistically similar type of transfer systems can be found encoded on the conjugative plasmid RP4. The main function of the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ is the conjugative transfer of the plasmid itself and other mobilizable plasmids to recipient cells. RP4, a 60,099 bp large IncPα plasmid with a broad host range, is able transfer and replicate in a variety of Gram-negative hosts including *Escherichia coli* and *A. tumefaciens* (Jain & Srivastava, 2013). Further, RP4 can facilitate plasmid transfer to Gram-positive bacteria as well as to the eukaryote *S. cerevisiae* (Trieu-Cuot *et al.*, 1987; Bates, Cashmore & Wilkins, 1998). Under laboratory conditions, transfer of RP4 itself or mobilization of other plasmids does not require a specific induction of the transfer system. One of the key elements of the RP4 mediated conjugative

transfer is the relaxase $TraI_{RP4}$ (**Figure 1A**), part of the Dtr_{RP4} system, encoded in the *tra1*$_{RP4}$ region (Pansegrau *et al.*, 1994). Guided by several system-specific auxiliary proteins $TraH_{RP4}$, $TraJ_{RP4}$ and $TraK_{RP4}$ (Pansegrau & Lanka; 1996A), $TraI_{RP4}$ recognizes a 6 bp sequence within the *oriT*$_{RP4}$ (origin of transfer) of the RP4 plasmid, where it introduces a single-strand nick and binds covalently to the 5'-end via a tyrosine residue present at position 22 by means of a transesterification reaction (**Figure 1A**; Pansegrau, Schröder & Lanka, 1993; Pansegrau, Schröder & Lanka, 1994; Balzer, Pansegrau & Lanka, 1994). After the nick at the *oriT*$_{RP4}$, a single-strand of the plasmid is presumably displaced via rolling-circle-type processing (Balzer, Pansegrau & Lanka 1994) and then recognized by the CP_{RP4} $TraG_{RP4}$ and transferred via the RP4 pilus structure $T4SS_{RP4}$ (Pansegrau *et al.*, 1994). Further, $TraI_{RP4}$ contains a possible $TraH_{RP4}$ interaction domain on the C-terminus (Pansegrau & Lanka, 1996B). $TraH_{RP4}$ is thought of stabilizing the protein-protein interactions of the relaxosome, thereby aiding in the nicking of the border sequences, without binding to DNA itself (Pansegrau & Lanka, 1996B).

Despite structural similarity and ability to mediate *trans*-kingdom transformation, there are several differences between the transfer as mediated by Dtr/CP/T4SSs of the Ti plasmid and the RP4 plasmid. In the Ti system, the substrate T-DNA contains two border sequences for relaxase nicking, defining a specific linear sequence. The linear T-strand, once it is transferred, has no intrinsic feature for double-strand formation. There are also no T-strand specific effector proteins present to initiate double-strand formation. The T-strand does not contain an origin of vegetative replication (*oriV*), and thus can only be stably maintained once it is integrated into the host genome (Gordon & Christie, 2014). In the RP4 system, the transferred DNA is a complete plasmid containing an *oriV* allowing it to replicate in a wide range of bacterial hosts (Pansegrau *et al.*, 1994). RP4 DNA contains only one motif for relaxase nicking, the *oriT*$_{RP4}$, as well as a DNA primase binding site for the primase $TraC_{RP4}$, initiating DNA double-strand formation in the recipient cell (Lanka & Fürste, 1984). The primase $TraC_{RP4}$ is also transferred via the $CP_{RP4}/T4SS_{RP4}$ into the recipient cell (Rees & Wilkins, 1990). As mentioned, both systems share a relatedness regarding the composition of their Dtr, CP and T4SS, in particular, their analogous relaxases (Scheiffele, Pansegrau & Lanka, 1995; Garcillán-Barcia, Francia & de La Cruz, 2009). The relaxase $VirD2_{Ti}$ contains an N-terminal relaxase domain (Vogel & Das, 1992) and has a TS located at the C-terminal region (van Kregten *et al.*, 2009). The relaxase $TraI_{RP4}$ also has the catalytic domain located at the N-terminus (**Figure 1A**; Pansegrau & Lanka, 1996A), as well as the TS in the C-terminus (**Figure 1A**; Cole, Lanka & Guiney, 1993; Cole & Guiney, 1995). It is to note that both relaxases $VirD2_{Ti}$ and $TraI_{RP4}$ are essential for DNA transfer via their respective CP/T4SS (Balzer, Pansegrau & Lanka, 1994; van Kregten *et al.*, 2009). Translocation of various proteins independent of DNA transfer has been shown to occur with the $CP_{Ti}/T4SS_{Ti}$ (Vergunst *et al.*, 2000; Schrammeijer *et al.*, 2003; Vergunst *et al.*, 2005; van Kregten *et al.*, 2009). With the $CP_{RP4}/T4SS_{RP4}$, protein translocation into the recipient cell has thus far only been inferred for the primase $TraC_{RP4}$ (Rees & Wilkins, 1990).

One striking phenomenon that was discovered for the $CP_{Ti}/T4SS_{Ti}$ was its ability to translocate proteins without any concomitant DNA transfer. This occurs naturally in specialized transfer systems also containing a CP/T4SS that are known to only translocate effector proteins, for example, the CP/T4SSs of specialized human pathogens as *Heliobacter pylori* and *Legionella pneumophilia* (Nagai & Roy, 2003; Cascales & Christie, 2004). Naturally, as described above, the $CP_{Ti}/T4SS_{Ti}$ is transferring T-DNA combined with effector proteins as well as effector proteins mechanistically independent of DNA transfer. In 2000, Vergunst *et al.* were able to prove that the effector protein $VirF_{Ti}$ could be transferred via the $CP_{Ti}/T4SS_{Ti}$ independent of T-DNA transfer into plant cells. It was also shown that the TS of the effector proteins $VirD2_{Ti}$, $VirF_{Ti}$, and $VirE2_{Ti}$ could be used for translocation *in vivo* (Schrammeijer *et al.*, 2003; van Kregten *et al.*, 2009). Until now, the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ has only been studied under the aspect of being a conjugative DNA delivery system (Ditta *et al.*,

1980; Simon, 1984; Wiater, Marra & Shuman, 1994; Paget *et al.*, 1999; Babic, Guérout & Mazel 2008, Strand *et al.*, 2014).

This study explored potential similarities between the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ and $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ and thereby investigated whether the relaxase $TraI_{RP4}$ did possess a discrete TS that directed its transport via the $CP_{RP4}/T4SS_{RP4}$. Evidence for a $TraI_{RP4}$ TS was found located at the relaxase C-terminus, and in addition it could be demonstrated for the first time that protein translocation between bacteria via the $CP_{RP4}/T4SS_{RP4}$ was possible independent of DNA transport. The implications and possible use of these observed phenomena are discussed within.

Results

$TraI_{RP4}$ TS is situated between the amino acids 625 and 732 of the protein

So far, no TS domain has been defined for the relaxase $TraI_{RP4}$; therefore, this study aimed to locate and characterize it. The position of the TS on the relaxase $TraI_{RP4}$ was determined by a modified Cre Reporter Assay for Translocation (CRAfT) assay. In this assay $TraI_{RP4}$ segments possibly containing a TS were expressed as a C-terminal fusion with a Cre recombinase protein encoded on an expression plasmid. When a CP_{RP4}/$T4SS_{RP4}$ compatible TS would be present, the fusion protein could then be transported from *E. coli* DH5α cells containing an RP4 plasmid into the recipient strain *E. coli* CSH26Cm:LTL, harbouring a Cre reporter target cassette integrated into the genome. The Cre reporter target cassette contains a tetracycline resistance cassette (*tet*) flanked by *loxP* sites, which is embedded within a chloramphenicol resistance cassette (*cat*). After Cre-mediated excision, the *loxP* sites are re-ligated, leaving only a 6 bp in-frame scar, and thereby creating a functional chloramphenicol resistance cassette. Successful fusion protein transfer can therefore be detected by a change from tetracycline to chloramphenicol antibiotic resistance. Several parts of the 732 amino acids (AA) $TraI_{RP4}$ protein sequence were fused to the C-terminus of the Cre recombinase. A first indicative CRAfT assay implied that a $TraI_{RP4}$ TS was located in the C-terminus, between 532-732 AA and a second one between 243-515 AA (**Figure S2**). Subsequent CRAfT assays with further truncations of the C-terminus were performed to narrow down the precise location of the TS (**Figure 1B**). In these experiments the fusion protein Cre::$TraI_{RP4}$ 625-732 AA demonstrated the most efficient translocation, indicating that a TS is located within the C-terminal 625-732 AA sequence. The translocation efficiency, as measured by the CRAfT assay, was higher than observed after fusion to the full-length $TraI_{RP4}$ 1-732 AA sequence, or the 243-732 AA sequence (lacking the relaxase domain) fused to Cre. The fusion constructs with the $TraI_{RP4}$ 1-531 AA fragment containing the relaxase domain as well as the middle part of the protein, but excluding the $TraI_{RP4}$ C-terminal part, exhibited a low but clearly detectable translocation frequency, indicative of the presence of a second TS between 243-415 AA (**Figure 1B**). The Cre::$TraI_{RP4}$ 599-679 AA and Cre::$TraI_{RP4}$ 532-632 AA fragments showed no evidence for transfer (**Figure 1B**), as the numbers of resistant colonies did not exceed those in negative controls (Cre protein only) of the comparable CRAfT assays in **figure S2**. When comparing the truncation 599-679 AA (no transfer) with the truncation 625-732 AA (transfer), it becomes apparent that the TS is likely to be located in the part of 680-732 AA. These would be the last C-terminal 53 AA. In conclusion, the CRAfT assay revealed the presence of a TS in the C-terminus of $TraI_{RP4}$ between 625-732 AA (last 107 AA of the protein), and a putative TS with lower activity located between 243-415 AA just beyond the N-terminal relaxase domain.

The $TraI_{RP4}$ TS shares a consensus arginine motif, charge and similar hydropathic profile with other T4SS TSs

The 680-732 AA C-terminal sequence (last 53 AA) of $TraI_{RP4}$ theoretically containing the translocation motif (**Figure 1B**) was compared to the C-termini of other relaxases and a primase encoded by RP4 (Miele *et al.*, 1991; Pansegrau *et al.*, 1994; Schröder & Lanka, 2005; Vergunst *et al.*, 2005; Garcillán-Barcia, Francia & de La Cruz, 2009). The following proteins were used for the C-terminal comparison: the relaxase $MobA_{RSF1010}$ of the mobilizable RSF1010 plasmid, which can be mobilized by RP4 (Haase *et al.*, 1995), with a C-terminal TS, which is recognized by the RP4 CP_{RP4} $TraG_{RP4}$. Importantly, RSF1010 can also be transferred from *Agrobacterium* by the plasmid Ti and its CP_{Ti} $VirD4_{Ti}$ (Beijersbergen *et al.*, 1992), which makes the comparison of the $MobA_{RSF1010}$ C-terminus containing the TS particularly interesting.

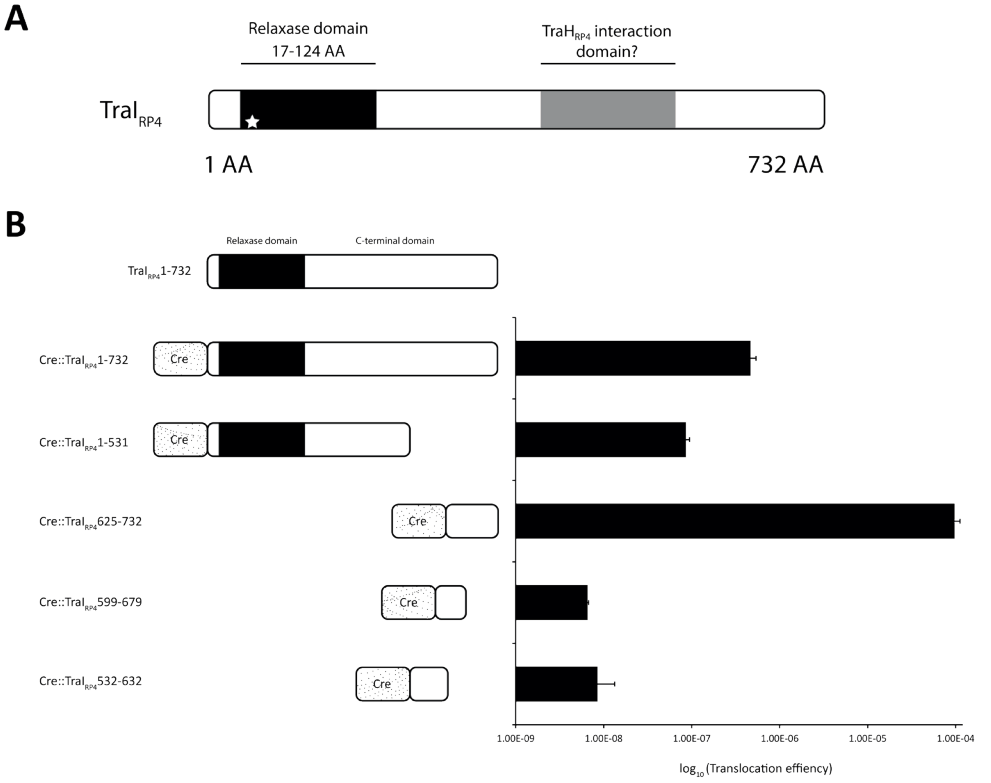


Figure 1. A: Schematic overview of the relaxase $TraI_{RP4}$. The full-length $TraI_{RP4}$ protein 1-732 AA, contains a relaxase domain situated between 17-124 AA at the N-terminus. The conserved essential tyrosine 22 forming the covalent bond with DNA at the *ori*T_{RP4}, is marked with a white star (Pansegrau & Lanka, 1996A). The $TraH_{RP4}$ interaction domain is potentially located at the C-terminus (Pansegrau & Lanka, 1996B). **B:** Characterization of the TS of the relaxase $TraI_{RP4}$ using a CRAfT assay. Cre recombinase was fused to $TraI_{RP4}$ truncations (indicated on the left) and was tested for transfer via RP4 CP_{RP4}/$T4SS_{RP4}$ from *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL. The x-axis (log_{10} scale) depicts the translocation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins, and indicates the used plasmids. The Cre protein box is not in size relation to the $TraI_{RP4}$ depiction. The $TraI_{RP4}$ truncations are in relative size to each other. The translocation frequencies are provided in detail in **figure S1**. Error bars depict SEM (n = 3).

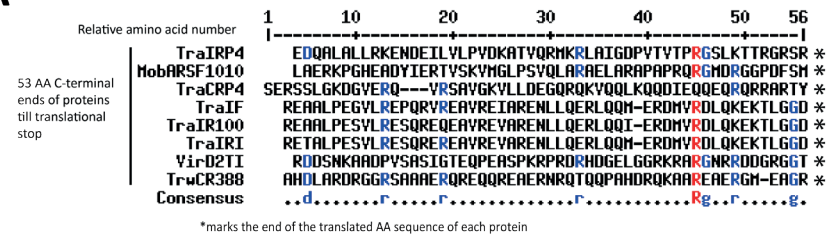
The TS of $MobA_{RSF1010}$ has been described in detail by Meyer (2015). Additionally, the relaxase $VirD2_{Ti}$ was compared, encoded on the Ti plasmid, described and characterized in more detail by Vergunst *et al.* (2005) and van Kregten *et al.* (2009). Furthermore, the relaxase of the conjugative plasmid R388 $TrwC_{R388}$ (Fernández-López *et al.*, 2006), the related relaxases $TraI_F$ of plasmid F, $TraI_{R100}$ of plasmid R100 and $TraI_{R1}$ of plasmid R1 (Cox & Schildbach, 2017) and the RP4 primase $TraC_{RP4}$ were included in the comparison to the C-terminus of $TraI_{RP4}$ (Miele *et al.*, 1991). An earlier study (Vergunst *et al.*, 2005), looking at the characteristics of the virulence proteins translocated via the CP_{Ti}/$T4SS_{Ti}$, described conservation of arginine residues in the TS, which were located in the last 30 AA of $VirF_{Ti}$, $VirE2_{Ti}$, $VirE3_{Ti}$, $VirD2_{Ti}$ and $VirD5_{Ti}$. When comparing the $TraI_{RP4}$ C-terminal end with the C-terminal ends (last 53 AA) of the relaxases and the primase mentioned above (**Figure 2A**), an arginine residue (R721) of $TraI_{RP4}$ appears to be conserved in the other relaxases but not in the primase. A second less conserved arginine (R709) in $TraI_{RP4}$ appears to be in consensus with $MobA_{RSF1010}$ and

$VirD2_{Ti}$ **(Figure 2A)**. Another important characteristic of CP/T4SS TSs was the net positive charge of the presumed TS domains, primarily due to the present arginine residues (Vergunst *et al.*, 2005; Alperi *et al.*, 2013; Meyer 2015). In the case of $TraI_{RP4}$, the C-terminal sequence contains 13.2% arginine residues (**Figure 2B**), which seems to be an enrichment over a normal occurrence of 5%, when each of the 20 AA residues would occur with equal chance. A rather similar arginine enrichment was found in other C-terminal sequences of related relaxases (**Figure 2B**). To further illustrate this, the frequency of occurrence of arginine residues in the C-terminal 30 AA sequence of $VirF_{Ti}$ TS sequence as well for $VirD2_{Ti}$ TS are around 20% to 30% compared to an average arginine frequency in *A. tumefaciens* of around 6.64% (Vergunst *et al.*, 2005). For 53 AA, statistically significant enrichment over an expected 5% to 6.64% using a chi-squared test is already reached when observing 7 to 8 arginine residues, which is the case with all C-terminal sequences shown (**Figure 2B**). Moreover, as another characteristic feature Vergunst *et al.* (2005) argued for the importance of a total positive charge of the TS for efficient translocation of the protein. This is also the case for the TS sequence of $TraI_{RP4}$ as it exhibited an all-over positive charge, just as its closest related relaxases $MobA_{RSF1010}$, $VirD2_{Ti}$, and $TrwC_{R388}$ as well as for the primase $TraC_{RP4}$ (**Figure 2B**). The more distantly related relaxases TraI of F, R1 and R100, which are in between themselves very related, exhibited a negative total charge in their C-terminal sequence (**Figure 2B**). Another factor assumed important for translocation is the hydrophilicity of the TS; reduction of hydrophilic properties of the $CP_{Ti}/T4SS_{Ti}$ TS strongly decreased translocation (Vergunst *et al.*, 2005). The 53 AA $TraI_{RP4}$ C-terminal sequence containing the TS is slightly hydrophilic, although mainly at the very terminal end. Comparative analysis of hydrophobicity plots (**Figure 2C**) showed the same trend for the $TraI_{RP4}$ C-terminal as for the other relaxase C-terminal ends as well as the primase C-terminus; the last 14 AAs of the terminal ends are all hydrophilic, with the conserved arginine R721 of $TraI_{RP4}$ being part of this sequence. Summarizing the results, the 53 AA C-terminal sequence of $TraI_{RP4}$ showed clear characteristics of a previously described relaxases $VirD2_{Ti}$ and $MobA_{RSF1010}$ TS motif that seems also to be present in other related relaxases as well as in the primase of RP4.

Protein translocation via the $CP_{RP4}/T4SS_{RP4}$ is independent of DNA transfer and is determined by the C-terminal TS motifs of $TraI_{RP4}$

When using the Ti plasmid for transfer, the $VirD2_{Ti}$ covalently bound to DNA is transferred into the recipient cells. Nevertheless, it was shown that $VirD2_{Ti}$ itself, in the absence of DNA transfer, could be translocated into recipient cells (Vergunst *et al.*, 2005). **Figure S2** provided evidence that the $TraI_{RP4}$ 243-732 AA fragment, without the relaxase domain and therefore not able to bind to DNA, could be translocated from cells containing RP4. It was subsequently tested using a CRAfT assay if the $TraI_{RP4}$ fragment could also be translocated from cell carrying RP4Δ*traI*, which is not transferred itself due to a deleted relaxase. In this experiment, the donor strain *E. coli* DH5α containing the RP4Δ*traI* plasmid and an expression plasmid expressing a Cre::$TraI_{RP4}$ 243-733 AA were used. The latter protein sequence has a deletion of the relaxase domain, which is located within the N-terminal 1-242 AA and therefore DNA transfer is prohibited from this donor. As a recipient strain *E. coli* CSH26Cm:LTL was used. While no DNA transfer occurred to the recipient as expected (**Figure S4B**), evidence for translocation of the Cre::$TraI_{RP4}$ 243-733 AA protein into recipient cells was found (**Figure 3A**). To further corroborate lack of RP4Δ*traI* plasmid transfer in the absence of any complementing full-length $TraI_{RP4}$, 5 to 10 randomly picked chloramphenicol-positive colonies were tested in each assay and always found to be sensitive to the RP4Δ*traI* based kanamycin marker, proving that RP4Δ*traI* plasmid transfer had not occurred. To demonstrate the functionality of RP4Δ*traI*, Cre fused to a full-length $TraI_{RP4}$ was used to substitute for the relaxase function. This fusion protein was translocated (**Figure 3A**), and it complemented the RP4Δ*traI* plasmid

A



B

TS (53 AA total)	Total charge	Arginine composition
$TraI_{RP4}$	+4	13.2% (7*)
$TraC_{RP4}$	+3	14.6% (8*)
$VirD2_{Ti}$	+3	18.9% (10*)
$MobA_{RSF1010}$	+2	13.2% (7*)
$TrwC_{R388}$	+3	22.6% (12*)
$TraI_{F}$	-1	17.0% (9*)
$TraI_{R100}$	-3	15.1% (8*)
$TraI_{R1}$	-2	17.0% (9*)

*total number of arginine residues in the 53 AA sequence

C

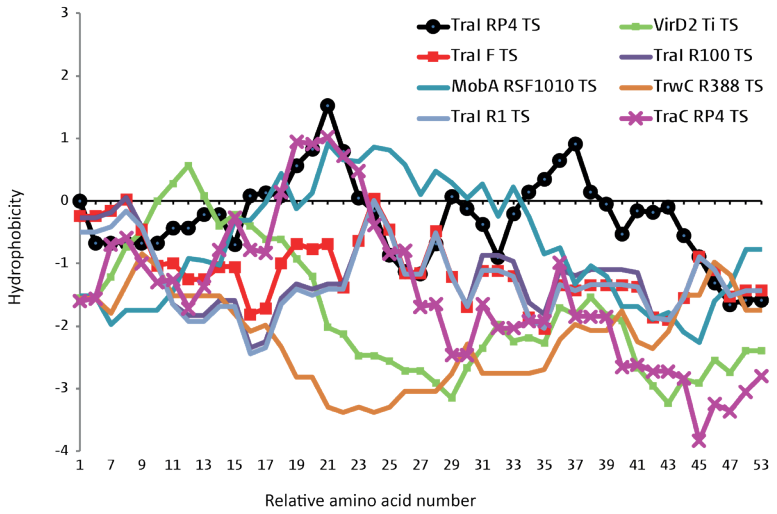


Figure 2. A: Alignment of the end C-terminal sequences (last 53 AA) of different CP/T4SS mobilizable proteins. The C-terminal 53 AA of $TraI_{RP4}$, potentially containing the TS, was compared to the C-terminal region of the same length of the relaxase $MobA_{RSF1010}$, $VirD2_{Ti}$, $TraI_{F}$, $TrwC_{R388}$, $TraI_{R100}$ and $TraI_{R1}$, and the primase $TraC_{RP4}$. High consensus AA are highlighted in a capital read letter, and lower consensus AA are highlighted in blue lowercase letters. The end of the translated AA sequence is indicated with an asterisk. The protein alignment was performed using the online Multalin software (http://multalin.toulouse.inra.fr/multalin). **B:** Comparison of the total charge of the 53 AA C-terminal residues of different relaxases and a primase translocated by a CP/T4SS and their arginine composition. The arginine composition in percent is based on the relative normal occurrence of all different essential 20 AAs in the sequence with equal chance. **C:** Kyte-Doolittle hydrophobicity plot. Comparison of the 53 AA C-termini of different relaxases and a primase translocated by a CP/T4SS. The x-axis shows the hydrophobicity index of the plot. Positive values indicate a potential hydrophobic part of the sequence (Kyte & Doolittle, 1985). The y-axis shows the 53 AAs of each C-terminus. Calculation of the Kyte-Doolittle index numbers for each C-terminal residue was performed using the software CLC Genomics Workbench 7 (https://www.qiagenbioinformatics.com).

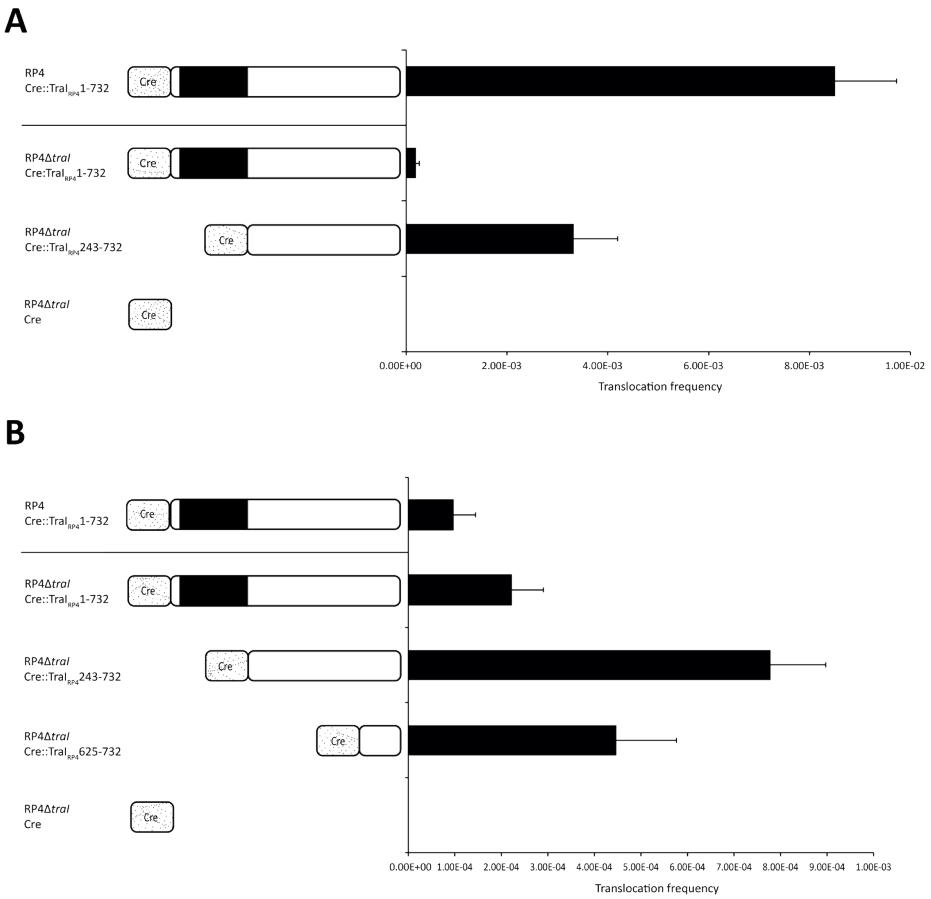


Figure 3. A: Protein translocation activity mediated by the TS of $TraI_{RP4}$ between E. *coli* via $CP_{RP4}/T4SS_{RP4}$ without conjugative DNA transfer using a CRAfT assay. The donor strain was *E. coli* DH5α and the recipient strain was *E. coli* CSH26Cm:LTL. The x-axis depicts the translocation frequency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids. The Cre protein box is not in size relation to the $TraI_{RP4}$ depiction. The $TraI_{RP4}$ truncations are in relative size to each other. The translocation frequencies are provided in detail in **figure S1.** Error bars depict SEM (n = 3). **B:** Protein translocation activity mediated by the C-terminal 625-732 AA of $TraI_{RP4}$ between *E. coli* via $CP_{RP4}/T4SS_{RP4}$ without conjugative DNA transfer using a CRAfT assay. As donor strain *E. coli* DH5α and recipient strain *E. coli* CSH26Cm:LTL was used. The x-axis depicts the transformation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids. The Cre protein box is not in size relation to the $TraI_{RP4}$ depiction. The $TraI_{RP4}$ truncations are in relative size to each other. The translocation frequencies are provided in detail in **figure S1.** Error bars depict SEM (n = 3).

for full plasmid transfer. As negative control, the Cre protein by itself was expressed in the strain containing RP4Δ*traI*, but from that strain, neither protein translocation nor plasmid transfer was detected (**Figure 3A**). To narrow down the TS motif requirements for protein-only translocation via CP_{RP4}/$T4SS_{RP4}$, these experiments were repeated and included a fusion protein where Cre was fused with the C-terminal TS of $TraI_{RP4}$ described in **figure 1B**. This 625-732 AA C-terminal fragment of $TraI_{RP4}$ fused to Cre was indeed also transferred from strains containing RP4Δ*traI* (**Figure 3B**). The frequency of transfer of Cre::$TraI_{RP4}$ 625-732 AA was comparable to transfer of the larger fusion protein only lacking the relaxase domain (**Figure 3B**). In summary, protein translocation independent of DNA transfer is possible via the CP_{RP4}/$T4SS_{RP4}$, and only the TS present in the small C-terminal part consisting of 625-732 AA of $TraI_{RP4}$ is sufficient for effective protein translocation.

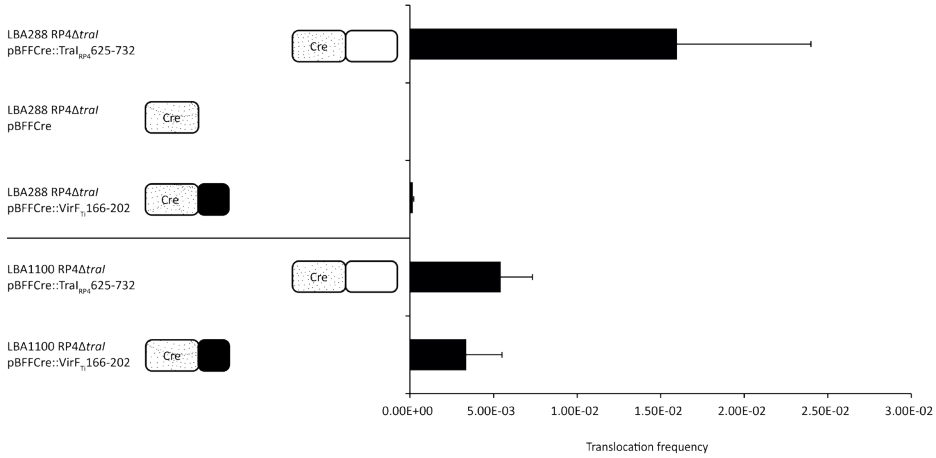


Figure 4. Protein translocation activity mediated by the C-terminal 625-732 AA of $TraI_{RP4}$ and the TS of $VirF_{Ti}$ via the CP_{RP4}/$T4SS_{RP4}$ from *A. tumefaciens* to *E. coli* without conjugative DNA transfer using a CRAfT assay. As donor strains, *A. tumefaciens* LBA288 (no Ti plasmid) and LBA1100 (containing a Ti plasmid) and recipient strain *E. coli* CSH26Cm:LTL were used. The x-axis depicts the translocation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids and donor strains. The grey dotted box is the Cre protein, the white box is the C-terminal part of $TraI_{RP4}$, and the black box is the TS for $VirF_{Ti}$. The size of the boxes is not in relation to the real size of the protein constructs and to each other. The translocation frequencies are provided in detail in **figure S1.** The error bars depicted are SEM (n = 3).

RP4 can deliver proteins, independent of DNA transfer, from *A. tumefaciens to E. coli*

The conjugative plasmid RP4 has a broad host range and is able to transfer into a large variety of bacteria. To explore the ability of RP4 to translocate proteins independent of DNA transfer into different bacterial species, translocation from *A. tumefaci*ens to *E. coli* was tested with CRAfT assays using Cre proteins fused with the short 625-732 AA $TraI_{RP4}$ sequence that was shown to contain a functional TS motif. Transfer requires not only the $T4SS_{RP4}$, but also the CP_{RP4} $TraG_{RP4}$, which selects the proteins to be transferred by their specific TS. In order to check for the selectivity of the CP_{RP4} $TraG_{RP4}$ not only to transfer protein constructs with the RP4 TS, but also protein constructs with the C-terminal 166-202 AA TS motif of the Ti

plasmid effector protein $VirF_{Ti}$ (Vergunst *et al.*, 2000; van Kregten *et al.*, 2009) was evaluated. Transfer of $VirF_{Ti}$ to recipient cells via the $T4SS_{Ti}$ normally requires the CP_{Ti} $VirD4_{Ti}$. The expression plasmid for Cre-fusion protein was introduced into *A. tumefaciens* donor strain LBA288 containing RP4Δ*traI* but lacking a Ti plasmid**.** In this set-up, no DNA transfer is possible and protein transfer can only occur via the CP_{RP4}/$T4SS_{RP4}$. As recipient strain *E. coli* CSH26Cm:LTL was used. Strains were co-cultivated at 30° C on rich LB medium. For expression of Cre fusion proteins in *A. tumefaciens*, the previously assembled expression plasmid pBFF (**Table 2**) was used. The expression cassette on pBFF is under the control of a $virF_{Ti}$ promoter, and protein sequences are fused with an N-terminal FLAG tag and SV40 NLS (van Kregten *et al.*, 2009). Although the natural $virF_{Ti}$ promoter is activated only under inducing conditions for the $VirA_{Ti}$/$VirG_{Ti}$ signalling pathway (Melchers *et al.*, 1990), van Kregten *et al.* (2009) already described that the $virF_{Ti}$ promoter construct in pBFF was active in rich LB medium at 30° C, which was verified for this promoter (**Figure S3**). CRAfT assays showed that Cre::$TraI_{RP4}$ 625-732 AA was transferred from *A. tumefaciens* LBA288 containing RP4Δ*traI* to *E. coli* (**Figure 4**). In contrast, from the same donor strain Cre::$VirF_{Ti}$ TS transfer was not seen, revealing the selectivity of the CP_{RP4}.

Discussion

This study focused on the characterization of the TS of the relaxase $TraI_{RP4}$ of the conjugative plasmid RP4. While the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ is a highly efficient system to transfer large size plasmids between cells, the related $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ of *A. tumefaciens* Ti plasmid can be used for either DNA delivery or protein only transfer (Schrammeijer *et al.*, 2003, Vergunst *et al.*, 2005). Here the position of a discrete TS motif of the relaxase $TraI_{RP4}$ is described. In addition, the TS motif fused to a Cre protein was shown to be able to mediate Cre translocation via the $CP_{RP4}/T4SS_{RP4}$, independent of transfer of a DNA substrate.

The main TS motif of the relaxase $TraI_{RP4}$ was shown to be situated on its C-terminus between 680-732 AA **(Figure 1B)**. This is congruent with the TS position of other relaxases, including $VirD2_{Ti}$, (Vergunst *et al.*, 2005; van Kregten *et al.*, 2009) and $MobA_{RSF1010}$ (Parker & Meyer, 2007; Meyer, 2015). The 108 AA $TraI_{RP4}$ C-terminal sequence when fused to a Cre protein was able to mediate translocating of the Cre protein into a recipient cell via the $CP_{RP4}/T4SS_{RP4}$ (**Figures 1B, 3A, 3B and 4**). The present study also indicated that there might be a second and much weaker TS motif in the middle of the $TraI_{RP4}$ protein, between 243 and 531 AA (**Figures 1B and 3B**). Evidence for the presence of a second TS in the middle of the protein adjacent to the relaxase domain was seen in several relaxases studied so far, including $MobA_{RSF1010}$, $VirD2_{Ti}$, $TraI_R$, $TraI_F$ (Parker & Meyer, 2007; Lang *et al.*, 2010; Meyer, 2015). This is in contrast to the translocated effector proteins, which only have a C-terminal TS (Vergunst *et al.*, 2005). The TS of $TraI_{RP4}$ was further characterized and compared to other C-terminal ends of relaxases as well of the C-terminal end of the RP4 primase. Previous studies proposed that functional specificity of TSs for mediating transport via CPs and cognate T4SSs does not entirely rely on the precise primary AA sequence, but may be mainly determined by the charge and hydrophilicity of the signal (Miele *et al.*, 1991; Vergunst *et al.*, 2005 Thomas & Hecht, 2007); different substrates (effector proteins), but all with a similar profile, can therefore bind to the CP for further translocation. Yet, some AA residues seem to be shared. A conserved arginine motif was found to be present in the C-termini of several T4SS transfer system relaxases of the IncP, IncQ and IncF plasmid families, and a second arginine residue more downstream is conserved also in $MobA_{RSF1010}$ and $VirD2_{Ti}$ (**Figure 2A**). The different relaxase TSs, as well as the primase C-terminus, share a similar hydrophilic profile, particularly for the last 14 AA that are all hydrophilic **(Figure 2C)**. Furthermore, the total charge comparison emphasized the characteristics of the TSs as well as that of the primase C-terminus of the IncP and IncQ plasmid family **(Figure 2B)**. It thus seems that any specific charge, specific AA, or other characteristics cannot be a determining factor for making a C-terminal sequence suitable as CP/T4SS TS, but that it rather will be a combination of the specific local charge, hydrophilic pattern, and still a certain sequence specificity. This combination will create a TS with specificity to the cognate CP of the translocation machinery. However, even within related translocation machineries, the specificity varies and may not allow transfer of heterologous substrates, *e.g.* the $VirF_{Ti}$ TS is not transferred via $CP_{RP4}/T4SS_{RP4}$ (**Figure 4**).

As mentioned, the 108 AA sequence of $TraI_{RP4}$ clearly mediated reporter protein translocation via the $CP_{RP4}/T4SS_{RP4}$. Such translocation was observed from *E. coli* to *E. coli* (**Figures 1B, 3A and 3B**), as well as also from *A. tumefaciens* to *E. coli* (**Figure 4**). This is in line with earlier observations that truncations of $TraI_{RP4}$ without C-terminus, reduced transfer of RP4 drastically (Cole, Lanka & Guiney, 1993; Cole & Guiney, 1995). In both studies (Cole, Lanka & Guiney, 1993; Cole & Guiney, 1995) the C-terminally truncated $TraI_{RP4}$ proteins consisted of the first 471 or 514 AAs that still mediated a reduced transfer of RP4, probably because they still contain the weaker secondary TS motif (**Figure 1B**).

Co-cultivation experiments determining conjugative transfer or protein translocation, all exhibit a large *inter*-experimental variation. For independent experiments using the same strains, constructs and parameters, a large variation in their measured

transformation efficiency was observed. Several repeats can manifest discrepancies of several magnitudes of the measured transformation efficiency, whereas measured phenomena consistently manifest, but are also bound to this variation. It is difficult to pinpoint the exact cause of this variation (*e.g.* temperature fluctuations, nutrient fluctuation, growth inhibition), but it is a common occurrence within this technique (author personal communication). The variation does not affect the main conclusion of each experiment. One example would be the discrepancy between the relative translocation efficiency of the full-length $TraI_{RP4}$ fused with Cre transferred with RP4 between **figure 1B** (translocation efficiency: $(4.6 \pm 1.4) \times 10^{-7}$) and **figures 3A** and **3B** (transformation efficiency: $(8.5 \pm 1.2) \times 10^{-3}$ and $(9.6 \pm 4.8) \times 10^{-5}$ respectively), that could be explained by the all-over fluctuations of a given conjugative CRAfT assay. The seemingly lower full-length $TraI_{RP4}$ translocation efficiency in **figure 1B**, should not be used in this case as an indication for its all-over transfer rate, but only should be viewed as indication for the presence of a TS within the $TraI_{RP4}$ sequence.

Some CP/T4SSs have the ability to translocate a relaxase and/or effector proteins independent of DNA transfer (Vergunst *et al.*, 2000). Experimentally, this was demonstrated by using CRAfT assays in which the candidate proteins were fused to the Cre protein and translocation could be monitored through a site-specific recombination event on a floxed substrate in the recipient cell. Thus it was shown that the TS of the relaxase $VirD2_{Ti}$ and effector proteins such as $VirF_{Ti}$ was enough to translocate the Cre protein independently of DNA transfer via the $CP_{Ti}/T4SS_{Ti}$ (Vergunst *et al.*, 2000; Vergunst *et al.*, 2005); the relaxase $TrwC_{388}$ could be translocated without DNA transfer via the $CPR_{388}/T4SS_{R388}$ of the IncW plasmid family (Draper *et al.*, 2005); also the primase Sog_{Colb} of the IncI plasmid ColIb could be transferred independently of DNA via the $CP_{Colb}/T4SS_{Colb}$ (Wilkins & Thomas, 2000). However, in other cases *e.g.* with the $CP_{R1}/T4SS_{R1}$ of plasmid R1, protein transfer independent of DNA transfer was not seen (Lang & Zechner, 2012). It was indirectly observed that a DNA substrate coupled to the relaxase plays a role in the energetical activation of the CP as well as specific parts of the T4SS for translocation of the DNA-protein complex (Cascales & Christie, 2004; Jakubowski *et al.*, 2009). When the DNA substrate is not present, relaxase protein translocation seems inhibited in some systems, as shown for $TraI_{R1}$, which is only translocated via its $CP_{R1}/T4SS_{R1}$when bound to DNA (Lang & Zechner, 2012). In the present study, it was shown for the first time that Cre fused with $TraI_{RP4}$ **(Figures 3A and 3B)** as well as Cre fused to the $TraI_{RP4}$ C-terminal sequence 625-732 AA **(Figure 3B)** was translocated via the $CP_{RP4}/T4SS_{RP4}$ independent of DNA transfer. The protein only translocation was also shown to be functional from *A. tumefaciens* to *E. coli* **(Figure 4)**. Further *trans*-kingdom delivery of proteins via the $CP_{RP4}/T4SS_{RP4}$, *e.g.* from *A. tumefaciens* to *S. cerevisiae*, or *E. coli* to *S. cerevisiae* still needs to be tested.

The recorded phenomenon of protein only delivery via $CP_{RP4}/T4SS_{RP4}$ has potential biotechnological applications using a modified RP4 plasmid as protein delivery tool. Currently, no comparable RP4 based protein delivery system exists, and so far, little literature is available on alterations and specifications using RP4 as a potential biotechnological tool beyond just DNA delivery (Babic, Guérout & Mazel, 2008). RP4 transfer is highly efficient in *E. coli,* and can be easily implemented and handled in laboratory conditions. RP4 can be easily isolated and modified, using standard molecular biological methods. In addition, the $T4SS_{RP4}$ pilus structure does not require specific induction conditions, and can polymerize within a wide temperature range. Further, loose contact of an *E. coli* donor cell to a recipient cell is sufficient for translocation of a protein (Samuels, Lanka & Davies, 2000). No induction and loose cell contact means, that simple and easy to use co-cultivation protocols can be established for protein translocation. In order to create an RP4-based delivery tool, the RP4 plasmid can modified. For example, when $oriT_{RP4}$ is be removed inhibiting the self-transfer of RP4, in combination with a second plasmid for fusion protein expression a versatile and powerful *inter*-bacterial T4SS protein delivery system is available.

Material and methods

Organisms and constructs

All strains used are listed in **table 1**, and plasmid constructs in **table 2**. *E. coli* DH5α, DH10β, CSH26CM::LTL and MS614 were made competent and transformed using the protocol of Inoue *et al.* (1990). *A. tumefaciens* strains were created using electroporation (den Dulk-Ras & Hooykaas, 1995) or tri-parental mating (Ditta *et al.*, 1980). All *E. coli* and *A. tumefaciens* strains used in this chapter were cultured on the appropriate antibiotic selection **(Tables 1 and 2)** on LB agar or in LB medium (10 g/L bacto tryptone, 5 g/L bacto yeast extract, 8 g/L NaCl, pH 7.0). *E. coli* was cultured at 37° C and *A. tumefaciens* at 29° C.

Molecular cloning

Standard molecular cloning procedures were performed according to established laboratory methods (Sambrook & Russell, 2001). Plasmids were contained and amplified in *E. coli* DH5α and DH10β. The plasmid RP4 and its derivatives were isolated from *E. coli* using the plasmid DNA purification kit Nucleobond PC100 (Macherey-Nagel), and from *A. tumefaciens* using the plasmid isolation kit GeneJet Plasmid Midi-prep Kit (Fermentas).

CRAfT-assay plasmids

The used plasmids and the primers used to generate inserts are described in **table 2** and the sequences of the primers are given in **table 3**.

For *E. coli* based protein expression, the following variations of the parental plasmid CFPBSm (Lang *et al.*, 2010) were created. Full-length and truncated versions of *traI*$_{RP4}$ DNA were amplified via PCR from an RP4 template, and digested with KpnI/SalI. The primers contained an opal TGA stop codon, the natural occurring stop codon of TraI$_{RP4}$, after each TraI$_{RP4}$ encoding fragment. Those fragments were then inserted in between the KpnI/SalI sites of the CFPBSm plasmid (Lang *et al.*, 2010). The constructs were verified by restriction digestion and sequencing.

For *A. tumefaciens* based expression, the parental plasmid pBFFCre(N) (van Kregten *et al.*, 2009) was used, containing a full-length Cre coding region fused at the N-terminus with a FLAG tag and SV40 NLS and fused at the C-terminus to a VirF$_{Ti}$ 166-202 AA TS. The following variations of the plasmid were created. The TraI$_{RP4}$ 625-732 AA fragment DNA sequence contained an opal TGA stop codon at its end and was amplified via PCR from an RP4 plasmid template and digested with SalI/XhoI. This fragment was then inserted into the XhoI digested plasmid pBFFCre(N) creating the plasmid pBFFCre::TraI625-732. The plasmid pBFFCre::VirF$_{Ti}$TS, was created by digesting the plasmid pBFFCre(N) with XhoI, removing a 90 bp sequence containing an N-terminal polylinker between Cre and the VirF$_{Ti}$ TS. The plasmid pBFFCre contains an oligo insertion (5′- to 3′-direction; GGC CGC TAT AGT GAT AAT TAG TAA CTA GGA ATT CAG CTG TT) in the NotI site of pBFFCre(N) supplying 3 in-frame stop codons shortly after the Cre coding sequence, preventing read through into the *virF*$_{Ti}$ TS. The constructs were verified by restriction digestions and sequencing.

CRAfT-assay

A Cre Reporter Assay for Translocation (CRAfT) assay (Vergunst *et al.*, 2000) as modified by Lang *et al.* (2010) was used to determine the translocation sequence of $TraI_{RP4}$, and to investigate protein translocation via $CP_{RP4}/T4SS_{RP4}$. The modifications to the original CRAfT assay by Lang *et al.*, (2010) were only minimal: a single recipient strain was used and different antibiotic selections due to the different selection markers present in the strains. The donor strain *E. coli* DH5α contained the expression vector CFPBSM coding for the fusion of the Cre protein and protein of interest (**Table 2**), as well as either the wild-type RP4 or the RP4Δ*traI* plasmid (see experiments). The donor strains of *A. tumefaciens* (LBA288 and LBA1100) contained the expression plasmid pBFF encoding Cre fusion proteins (**Table 2**) as well as the RP4 plasmid. See **table 1** for details of the used strains in this assay.

When *E. coli* DH5α was used as a donor with the recipient *E. coli* CSH26Cm::LTL or MS614, both strains were cultured in liquid LB medium with appropriate antibiotics, depending on constructs used, for selection for 24 h at 37° C. Then 100 µL of donor strain was transferred into 3 mL fresh LB medium without antibiotic selection and cultured for around 1 h at 37° C until the strains reached an OD_{600} of 0.2. For the recipient strain, after overnight growth, a volume with an OD_{600} of 0.2 was centrifuged to remove the medium containing antibiotics. The recipient cells were mixed with OD_{600} of 0.2 of the selected donor strain, providing a mixing ratio of 1:1. The combined cell suspension was centrifuged to remove the medium and then re-suspended in 30 µL fresh LB and transferred onto nitrocellulose filters (0.48 µm pore size) on solid LB. The co-cultivation was carried out at 37° C for 3 h.

When *A. tumefaciens* (LBA288 or LBA1100) was used as a donor for the recipient strain *E. coli* CSH26Cm::LTL, different cultivation temperature needed to be used. *A. tumefaciens* strains were cultured in liquid LB medium on rifampicin (10 µg/mL) selection for 24 h at 30° C. After incubation, 100 µL each of donor strain was transferred into 3 mL LB medium without selection until the strains reached an OD_{600} of 0.2. For the recipient strain, after overnight growth, a volume with an OD_{600} of 0.2 was centrifuged to remove the medium containing antibiotics. The recipient cells were mixed with OD_{600} of 0.2 of the selected donor strain, providing a mixing ratio of 1:1. The cell suspension was centrifuged to remove the medium, and then re-suspended in 30 µL fresh LB to be transferred onto nitrocellulose filters (0.48 µm pore size) on solid LB. The co-cultivation was carried out at 30° C for 24 h.

After co-cultivation, the filter was transferred from the agar surface into reaction tubes containing 1 mL LB, which were vortexed until the biofilm was re-suspended. A dilution series was plated onto LB agar plates with chloramphenicol (25 µg/mL) selection to determine the amount of recombinants cells. To only determine the recipient cell output, a dilution series was plated on LB agar plates with tetracycline (10 µg/mL) selection. All plates were cultivated at 37° C and scored after 24 h. The protein translocation frequencies were expressed as transconjugants per total number of recipient cells output after the co-cultivation.

RP4 and RP4Δ*traI* self-transfer was tested for each construct by transferring chloramphenicol resistant colonies onto a fresh LB agar plate containing kanamycin (100 µg/mL) and culturing it at 37° C for 24 h. The RP4 plasmid was considered as transferred when a re-streaked colony was able to grow on a kanamycin (100 µg/mL) selection plate. For RP4Δ*traI*, if a re-streaked colony was not able to grow, it was considered that the plasmid was not transferred.

C-terminal amino acid sequences of relaxases and analysis

The last 53 C-terminal amino acid sequences of the relaxases were taken from sequences from publicly available repositories: $TraI_{RP4}$ (GenBank L27758.1), $TraC1_{RP4}$ referred to as $TraC_{RP4}$ in this chapter, as the last 53 AA sequence of $TraC2_{RP4}$ is homologous to $TraC1_{RP4}$ (GenBank BN000925.1), $MobA_{RSF1010}$ (GenBank M28829.1), $VirD2_{Ti}$ (GenBank AF242881.1), $TrwC_{R388}$ (GenBank NC_028464.1), $TraI_{F}$ (GenBank NC_002483), $TraI_{R100}$ (GenBank NC_002134.1), and $TraI_{R1}$ (GenBank AY423546.1). To align and compare the protein sequences the Multalin software (http://multalin.toulouse.inra.fr/multalin) was utilized, and for the calculation of the Kyte-Doolittle index, the total protein charge calculation and arginine composition the software CLC Genomics Workbench 7 (https://www.qiagenbioinformatics.com) was used.

RP4 modification

To test protein translocation via the $CP_{RP4}/T4SS_{RP4}$ independent of DNA transfer, the relaxase $traI_{RP4}$ was removed from the RP4 plasmid using the λ-Red homologous recombination method based on Datsenko & Wanner (2000). The linear targeting DNA, used to knock-out $traI_{RP4}$ on RP4, contained an apramycin cassette flanked by *loxP* sites for later removal of the resistance cassette using Cre recombination, followed by homologous sequences of $traI_{RP4}$. To create the $traI_{RP4}$ target DNA**,** the apramycin resistance cassette (*aac(3)IV*) was amplified via PCR from plasmid pSET152 with primer ApraFW and ApraRV introducing SalI and SpeI restriction sites, respectively. The PCR product was ligated into the plasmid pJET1.2 (Fermentas), and transformants exhibiting apramycin resistance were selected. The SalI/SpeI apramycin fragment was removed from the positive selected pJET1.2 plasmids with the PCR insert, and then inserted into XhoI/SpeI digested p2G2L plasmid. The p2G2L plasmid contained a SalI/SpeI cloning site that is flanked by *LoxP*; this added the sites to be later removed by the Cre recombination. This created an apramycin resistance cassette flanked by *LoxP* sites creating the plasmidp2G2LApra. Primers that were used to amplify the *LoxP* flanked apramycin cassette, contained $traI_{RP4}$ homologous sequences; the amplified cassettes were then used as linear target DNA. Due to the length of the $traI_{RP4}$ homology, a two-step amplification protocol using two overlapping primer pairs was applied. The PCR protocol was as follows: 4 cycles of pre-amplification using the primer pair TraIFW1 and TraIRV1 at 72° C for the annealing step. Subsequently, 1/10 of the initial PCR product was transferred into a new PCR mix containing the second primer pair TraIFW2 and TraIRV2, initially with 4 cycles using a 60° C annealing temperature, which was increased to 72° C for another 31 cycles. The PCR product was purified by agarose gel electrophoresis using the Zymoclean gel DNA recovery kit. The purified PCR product was then used for disrupting the $traI_{RP4}$ locus on the RP4 plasmid. For this, the temperature-sensitive pKD46 plasmid containing the λ-Red recombinase was inserted into *E. coli* MS411 containing RP4. Of the PCR amplified targeting DNA (from plasmid p2GSLApra), 500 ng were transformed into this strain allowing for homologous recombination according to Datsenko & Wanner (2000). Insertion of the disruption sequence was checked with primers TraICheckFW, TraICheckRV, ApraCheckFW and ApraCheckRV. The apramycin cassette, now present at the $traI_{RP4}$ locus, was excised using Cre recombination. RP4 plasmids containing the disruption sequence were inserted into *E. coli* MS411 pKD-*cre*, as both plasmids contained chloramphenicol resistance. Chloramphenicol- and kanamycin-resistant transformants were selected at 30° C after 48 h, and re-streaked four times at 37° C with kanamycin selection only, to lose the pKD-*cre* vector. This created the plasmid RP4Δ*traI*. The knock-out of $traI_{RP4}$ was checked with primers TraICheckFW and TraICheckRV (**Figure S4A**). The plasmid RP4Δ*traI* was transformed into the donor *E. coli* DH5α and checked for lack of self-transfer to *E. coli* recipient MS614. No transconjugants were observed (**Figure S4B**). When RP4Δ*traI* was inserted into *E. coli* DH5α containing the

plasmid pGZ119EHTraI (containing a *traI*$_{RP4}$), plasmid transfer (pGZ119EHTraI) as well as self-transfer (RP4Δ*traI*) was observed (**Figure S4B**). For all experiments, the RP4Δ*traI* clone #1 was used (**Figure S4A**). The plasmid pGZ119EHTraI was created as follows. The full-length *traI*$_{RP4}$ was PCR amplified from plasmid RP4 containing SalI digestion sites on both ends. The SalI *traI*$_{RP4}$ fragment was inserted into SalI digested pGZ119EH plasmid, creating the plasmid pGZ119EHTraI. The plasmid was checked by restriction digest and sequencing.

Table 1. Organisms used in this study.

Strain	Description	Resistance	Reference
A. tumefaciens			
LBA288	based on WT C58 (LBA201); Rif; Nal; cured of pTiC58; cryptic plasmid pAtC58	Rif, Nal	Hooykaas *et al.*, 1980
LBA1100	LBA 288 with pAL1100 or octopine pTiB6: ΔT_L, ΔT_R, Δ*tra*, Δ*occ*; pAtC58	Rif, Nal, Spec	Beijersbergen *et al.*, 1992
E. coli			
DH5α	*F-*; *Φ80lacZΔM15*; *Δ(lacZYA-argF)*; *U169*; *recA1*; *endA1*; *hsdR17 (rK–, mK+)*; *phoA*; *supE44*; *λ– thi-1*; *gyrA96*; *relA1*		Grant *et al.*, 1990
DH10β	*F-*; *araDJ39*; *A(ara, leu)7697*; *AlacX74*; *galU*; *galK*; *rpsL*; *deoR*; *Φ80dlacZAM15*; *endAI*; *nupG*; *recA1*; *mcrA*; *A(mrr hsdRMS*, *mcrBC)*		Grant *et al.*, 1990
CSH26CM::LTL	CSH26 galK::cat::loxP-Tet-loxP	Tc	Lang *et al.*, 2010
MS411	*ilvGI*; *rfb-50*; *thi*		Lang *et al.*, 2010
MS614	*ilvG*; *rfb-50*; *thi*; *rpsL*	Sm	Lang *et al.*, 2010

Nal: Nalidixic acid; Rif: Rifampicin; Sm: Streptomycin; Spec: Spectinomycin; Tc: Tetracycline; WT: Wild-type

Table 2. Genetic constructs used in this study.

Plasmid	*oriV*	Description	Resistance	Reference
CFPBSm	*pMB1*	Cre fusion plasmid	Sm	Lang *et al.*, 2010
Cre::$TraI_{RP4}$1-732	*pMB1*	CFPBSm: $TraI_{RP4}$ 1-732 AA Primer: CreTraIfw01/CreTraIrev01	Sm	This study
Cre:: $TraI_{RP4}$1-531	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 1-531 AA Primer: FTW151/FTW152	Sm	This study
Cre:: $TraI_{RP4}$243-732	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 243-732 AA Primer: CreTraIfw02/ CreTraIrev01	Sm	This study
Cre:: $TraI_{RP4}$625-732	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 625-732 AA Primer: FTW149/FTW150	Sm	This study
Cre:: $TraI_{RP4}$599-679	*pMB1*	CFPBSm: Truncation $TraI_{RP4}$ 599-679 AA Primer: FTW147/FTW148	Sm	This study
Cre:: $TraI_{RP4}$532-632	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 532-632 AA Primer: FTW145/FTW146	Sm	This study
pBFFCre	*pBBR1*	ZN245; full-length Cre with N-terminal fused FLAG tag and SV40 NLS	Gm	van Kregten *et al.*, 2009
pBFFCre(N)	*pBBR1*	ZN239; parental plasmid for cloning; $virF_{Ti}$ promoter; contains full-length Cre with fused N-terminal FLAG tag and SV40; NLS; polylinker; $VirF_{Ti}$ 166-202 AA translocation signal	Gm	van Kregten *et al.*, 2009
pBFFCre::$VirF_{Ti}$TS	*pBBR1*	Translocation signal $VirF_{Ti}$ 166-202 AA fused to Cre, with polylinker XhoI fragment removed	Gm	BJ van der Zaal, unpublished
pBFFCre::$TraI_{RP4}$625-732	*pBBR1*	Full-length Cre fused to truncation $TraI_{RP4}$ 625-732 AA Primer: TraI625FW and TraI732RV	Gm	This study
RP4	*Inc*Pα	mob_{RP4}; tra_{RP4}; $oriT_{RP4}$; isolated from strain *A. tumefaciens* LBA2210	Km, Tc, Cb	Pansegrau *et al.*, 1994
RP4Δ*traI*	*Inc*Pα	Δ*traI*	Km, Tc, Cb	This dissertation, Chapter 5
pKD46	*pSC101*	contains the λ-Red recombinase; r*epA101ts* is temperature sensitive	Amp	Datsenko & Wanner, 2000
pKD-*cre*	*pSC101*	based on pKD46; encoding Cre recombinase	Amp	Lang *et al.*, 2014
p2G2L	*ColE1, F1*	Target DNA plasmid for λ-Red knock-out; pBSSK based; cloning site XhoI and SpeI flanked by *loxP* sites	Amp	S Lang, unpublished
p2G2LApra	*ColE1, F1*	apramycine *aac(3)IV* cassette flanked with *loxP* sites	Amp	This study
pSET152	*pMB1*	lacZαMCS; $oriT_{RP4}$; $Int^{\Phi C31}$	Apra	Bierman, Logan & O'Brian, 1992
pGZ119EH	*ColD*	$lacI^{Q}$ regulated *pTAC* promoter	Cam	Lessl *et al.*, 1992
pGZ119EHTraI	*ColD*	$lacI^{Q}$ regulated *pTAC* promoter; full-length $traI_{RP4}$	Cam	This study
pJET1.2	*pMB1*	*bla* (Ap^{R}); *eco47IR*; P_{lacUV5}; *pT7*; supplied as cloning vector blunt linearized with EcoRV		Fermentas, GenBank EF694056

oriV: Origin of vegetative replication (synonymously with origin of replication *oriRep* or *ori*); Amp: Ampicilin; Apra: Apramycin; Cam: Chloramphenicol; Cb: Carbenicillin; Gm: Gentamycin; Km: Kanamycin; Sm: Streptomycin; Tc: Tetracycline

Table 3. Oligonucleotide sequences used in this study.

FTW145 GCAGGTACCATTCGACCTGCCGCAGCCAGAAC
FTW146 GCAGTCGACTCATCTTCCTGGCCCAGAAACTC
FTW147 GCAGGTACCGCTGAAATGCTATTGCCGCGTGA
FTW148 GCAGTCGACTCAGACGTTCCGGGTGCCTGCATA
FTW149 GCAGGTACCGGAGTTTCTGGGCCAGGAAGAGG
FTW150 GCAGTCGACTCATCTACTCCTACCTCGGGTAG
FTW151 GCAGGTACCGTGATTGCCAAGCACGTCC
FTW152 GCAGTCGACTCAGCCTCCATCGCCGCCGCGA
CreTraIfw01 GCAGGTACCGTGATTGCCAAGCACGTCC
CreTraIrev01 GCAGTCGACTCATCTACTCCTACCTCGGGTAGTTTTAAGGG
CreTraIfw02 GCAGGTACCCCGAAGCTCGAAGCCCGATTCGG
TraI625FW AATTCTCGAGGGAGTTTCTGGGCCAGGAA
TraI732RV AATTGTCGACTCATCTACTCCTACCTCGGG
ApraFW: GGAAGTCGACTACGCGCAGAAAAAAAGGATCT
ApraRV: GGAAACTAGTCGGTGAGTTCAGGCTTTTTCAT
TraIFW1:CAAAGGTGATTGCCAGACGATAACTTCGTATAGCATACATTATACGAAGTTATCTCGA-CTACGCGCAGAAAAAAAGGATCTCAA
TraIFW2: CTCCTTCGTCAGTAGTTCTTGACGGCGGCGCTCAAGGGCGGCGTCGTCAAAGGTGATTGCCAGACG
TraIRV1: ACGAACTGCGAGGCAAACATAACTTCGTATAATGTATGCTATACGAAGTTATACTAGTC-GGTGAGTTCAGGCTTTTTCATATCTCATT
TraIRV2: ATCACCGACGAGCAAGGCAAGACCGAGCGCCTGGGTCACGTGCGCGTCACGAACTGCGAGGCAAAC
TraICheckFW: GTGGCGTTGAGGGTGGTTCTG
TraICheckRV: GGCGAAGATCGAAGAGAAGCAG
ApraCheckFW: ACGAATGGCGAAAAGCCGAG
ApracCheckRV: GGCGGATGCAGGAAGATCAAC

*Underscored bases are opal stop codons; all oligonucleotides are depicted in 5′-3′direction

Acknowledgements

This work is part of the research programme Green Genetics with project number 828.11.001, which is (partly) financed by the Dutch Research Council (NWO).

References

Alperi A, Larrea D, Fernández-González E, Dehio C, Zechner EL & Llosa M (2013) A translocation motif in relaxase TrwC specifically affects recruitment by its conjugative type IV secretion system. *J. Bacteriol.* **195:** 4999–5006

Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* **73:** 775–808

Babic A, Guérout AM & Mazel D (2008) Construction of an improved RP4 (RK2)-based conjugative system. *Res. Microbiol.* **159:** 545–549

Balzer D, Pansegrau W & Lanka E (1994) Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. *J. Bacteriol.* **176:** 4285–4295

Bates S, Cashmore AM & Wilkins BM (1998) IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*: involvement of the Tra2 mating system. *J. Bacteriol.* **180:** 6538–6543

Beijersbergen AGM, den Dulk-Ras AD, Schilperoort RA & Hooykaas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256:** 1324–1327

Bierman M, Logan R & O'Brien K (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces spp*. Gene **116**: 43–49

Brencic A & Winans SC (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol. Mol. Biol. Rev.* **69:** 155–194

Bundock P, den Dulk-Ras AD, Beijersbergen AGM & Hooykaas PJJ (1995) *Trans*-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J.* **14:** 3206–3214

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

Cole SP, Lanka E & Guiney DG (1993) Site-directed mutations in the relaxase operon of RP4. *J. Bacteriol.* **175:** 4911–4916

Cole SP & Guiney DG (1995) Site-specific mutations in the TraI relaxase and upstream region of plasmid RP4. *Plasmid* **34:** 236–239

Cox KEL & Schildbach JF (2017) Sequence of the R1 plasmid and comparison to F and R100. *Plasmid* **91:** 53–60

Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97:** 6640–6645

de Groot MJA, Bundock P, Hooykaas PJJ & Beijersbergen AGM (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol.* **16:** 839–842

Ditta G, Stanfield S, Corbin D & Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* **77:** 7347–7351

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

den Dulk-Ras AD & Hooykaas PJJ (1995) Electroporation of *Agrobacterium tumefaciens*. *Plant Cell Electroporation Electrofusion Protoc.* **55:** 63–72

Escobar MA & Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci.* **8:** 380–386

Farrand S, Hwang I & Cook D (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* **178:** 4233–4247

Fernández-López R, Garcillán-Barcia MP, Revilla C, Lázaro M, Vielva L & de la Cruz F (2006) Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. *FEMS Microbiol. Rev.* **30:** 942–966

Garcillán-Barcia MP, Francia MV & de La Cruz F (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* **33:** 657–687

Gordon JE & Christie PJ (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Grant SG, Jessee J, Bloom FR & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 4645–4649

Grohmann E, Christie PJ, Waksman G & Backert S (2018) Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* **107:** 455–471

Haase J, Lurz R, Grahn AM, Bamford DH & Lanka E (1995) Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. *J. Bacteriol.* **177:** 4779–4791

Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, von Bodman SB, Lanka E, Ream W & Farrand SK (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* **182:** 1541–1548

Hooykaas PJJ & Beijersbergen AGM (1994) The virulence system of *Agrobacterium tumefaciens*. *Annu. Rev. Phytopathol.* **32:** 157–181

Hooykaas PJJ, den Dulk-Ras H, Ooms G & Schilperoort RA (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J. Bacteriol.* **143:** 1295–1306

Inoue H, Nojima H & Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96:** 23–28

Jakubowski SJ, Kerr JE, Garza I, Krishnamoorthy V, Bayliss R, Waksman G & Christie PJ (2009) *Agrobacterium* VirB10 domain requirements for type IV secretion and T pilus biogenesis. *Mol. Microbiol.* **71:** 779–794

Jain A & Srivastava P (2013) Broad host range plasmids. *FEMS Microbiol. Lett.* **348:** 87–96

Kyte J & Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132

Lang S, Gruber CJ, Raffl S, Reisner A & Zechner EL (2014) Common requirement for the relaxosome of plasmid R1 in multiple activities of the conjugative type IV secretion system. *J. Bacteriol.* **196:** 2108–2121

Lang S, Gruber K, Mihajlovic S, Arnold R, Gruber CJ, Steinlechner S, Jehl MA, Rattei T, Fröhlich KU & Zechner EL (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol. Microbiol.* **78:** 1539–1555

Lang S & Zechner EL (2012) General requirements for protein secretion by the F-like conjugation system R1. *Plasmid* **67:** 128–138

Lanka E & Fürste JP (1984) Function and properties of RP4 DNA primase. In *Proteins Involved in DNA Replication* pp 265–280. Boston, MA: Springer US

Mantis NJ & Winans SC (1992) The *Agrobacterium tumefaciens vir* gene transcriptional activator VirG is transcriptionally induced by acid pH and other stress stimuli. *J. Bacteriol.* **174:** 1189–1196

Melchers LS, Maroney MJ, den Dulk-Ras AD, Thompson DV, van Vuuren HAJ, Schilperoort RA & Hooykaas PJJ (1990) Octopine and nopaline strains of *Agrobacterium tumefaciens* differ in virulence; molecular characterization of the *virF* locus. *Plant Mol. Biol.* **14:** 249–259

Melchers LS, Regensburg-Tuïnk AJG, Bourret RB, Sedee NJ, Schilperoort RA & Hooykaas PJJ (1989) Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. *EMBO J.* **8:** 1919–1925

Meyer R (2015) Mapping type IV secretion signals on the primase encoded by the broad-host-range plasmid R1162 (RSF1010). *J. Bacteriol.* **197:** 3245–3254

Miele L, Strack B, Kruft V & Lanka E (1991) Gene organization and nucleotide sequence of the primase region of IncP plasmids RP4 and R751. *DNA Seq.* **2:** 145–162

Nagai H & Roy CR (2003) Show me the substrates: modulation of host cell function by type IV secretion systems. *Cell. Microbiol.* **5:** 373–383

Paget MSB, Chamberlin L, Atrih A, Foster SJ & Buttner MJ (1999) Evidence that the extracytoplasmic function sigma factor σ(E) is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **181:** 204–211

Pansegrau W & Lanka E (1996A) Mechanisms of initiation and termination reactions in conjugative DNA processing. *J. Biol. Chem.* **271:** 13068–13076

Pansegrau W & Lanka E (1996B) Enzymology of DNA transfer by conjugative mechanisms. *Prog. Nucleic Acid Res. Mol. Biol.* **54:** 197–251

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncPα plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Pansegrau W, Schröder W & Lanka E (1993) Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90:** 2925–2929

Pansegrau W, Schröder W & Lanka E (1994) Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J. Biol. Chem.* **269:** 2782–2789

Parker C & Meyer RJ (2007) The R1162 relaxase/primase contains two, type IV transport signals that require the small plasmid protein MobB. *Mol. Microbiol.* **66:** 252–261

Rees CED & Wilkins BM (1990) Protein transfer into the recipient cell during bacterial conjugation: studies with F and RP4. *Mol. Microbiol.* **4:** 1199–1205

Sambrook J & Russell DW (2001) Molecular cloning: a laboratory manual 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Samuels AL, Lanka E & Davies JE (2000) Conjugative junctions in RP4-mediated mating of *Escherichia coli*. *J. Bacteriol.* **182:** 2709–2715

Scheiffele P, Pansegrau W & Lanka E (1995) Initiation of *Agrobacterium tumefaciens* T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of superhelical T-border DNA *in vitro*. *J. Biol. Chem.* **270:** 1269–1276

Schrammeijer B, den Dulk-Ras AD, Vergunst AC, Jácome EJ & Hooykaas PJJ (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. *Nucleic Acids Res.* **31:** 860–868

Schröder G & Lanka E (2005) The mating pair formation system of conjugative plasmids - a versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54:** 1–25

Simon R (1984) High frequency mobilization of Gram-negative bacterial replicons by the *in vitro* constructed Tn 5-Mob transposon. *MGG Mol. Gen. Genet.* **196:** 413–420

Strand TA, Lale R, Degnes KF, Lando M & Valla S (2014) A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PLoS One* **9:** 1–6

Thomas J & Hecht DW (2007) Interaction of *Bacteroides fragilis* pLV22a relaxase and transfer DNA with *Escherichia coli* RP4-TraG coupling protein. *Mol. Microbiol.* **66:** 948–960

Trieu-Cuot P, Carlier C, Martin P & Courvalin P (1987) Plasmid transfer by conjugation from *Escherichia coli* to Gram-positive bacteria. *FEMS Microbiol. Lett.* **48:** 289–294

Turk SCJ, Melchers LS, den Dulk-Ras H, Regensburg-Tuïnk AJG & Hooykaas PJJ (1991) Environmental conditions differentially affect *vir* gene induction in different *Agrobacterium* strains. Role of the VirA sensor protein. *Plant Mol. Biol.* **16:** 1051–1059

van Kregten M, Lindhout BI, Hooykaas PJJ & van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Mol. Plant-Microbe Interact.* **22:** 1356–1365

Vergunst AC, Schrammeijer B, den Dulk-Ras AD, De Vlaam CMT, Regensburg-Tuink AJG & Hooykaas PJJ (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science.* **290:** 979–982

Vergunst AC, van Lier MCM, den Dulk-Ras AD, Grosse Stuve TA, Ouwehand A & Hooykaas PJJ (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. U.S.A.* **102:** 832–837

Vogel AM & Das A (1992) Mutational analysis of *Agrobacterium tumefaciens* VirD2: tyrosine 29 is essential for endonuclease activity. *J. Bacteriol.* **174:** 303–308

Wilkins BM & Thomas AT (2000) DNA-independent transport of plasmid primase protein between bacteria by the I1 conjugation system. *Mol. Microbiol.* **38:** 650–657

Ziemienowicz A, Merkle T, Schoumacher F, Hohn B & Rossi L (2001) Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. *Plant Cell* **13:** 369

Supplemental information

Supplemental Table S1. Tables showing the average translocation frequencies of CRAfT assays within figures.

Figure 1 protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Constructs used	Average translocation efficiency (n = 3)
RP4 / Cre::$TraI_{RP4}$1-732	$(4.6 \pm 0.8) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$1-531	$(8.4 \pm 1.0) \times 10^{-8}$
RP4 / Cre::$TraI_{RP4}$625-732	$(9.5 \pm 1.7) \times 10^{-5}$
RP4 / Cre::$TraI_{RP4}$599-697	$(6.4 \pm 0.3) \times 10^{-9}$
RP4 / Cre::$TraI_{RP4}$532-732	$(8.3 \pm 1.9) \times 10^{-9}$

Figure 3A protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Constructs used	Average translocation efficiency (n = 3)
RP4 / Cre::$TraI_{RP4}$1-732	$(8.5 \pm 1.2) \times 10^{-3}$
RP4Δ*traI* / Cre::$TraI_{RP4}$1-732	$(1.9 \pm 0.5) \times 10^{-4}$
RP4Δ*traI* / Cre::$TraI_{RP4}$243-732	$(3.3 \pm 0.8) \times 10^{-3}$
RP4Δ*traI* / Cre	0

Figure 3B protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Constructs used	Average translocation efficiency (n = 3)
RP4 / Cre::$TraI_{RP4}$1-732	$(9.6 \pm 4.8) \times 10^{-5}$
RP4Δ*traI* / Cre::$TraI_{RP4}$1-732	$(2.2 \pm 0.7) \times 10^{-4}$
RP4Δ*traI* / Cre::$TraI_{RP4}$243-732	$(7.8 \pm 1.2) \times 10^{-4}$
RP4Δ*traI* / Cre::$TraI_{RP4}$625-732	$(4.5 \pm 1.2) \times 10^{-4}$
RP4Δ*traI* / Cre	0

Figure 4 protein translocation frequencies from the *A. tumefaciens* LBA288 and LBA1100 donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Strains and constructs used	Average translocation efficiency (n = 3)
LBA288 / RP4Δ*traI* / pBFFCre::$TraI_{RP4}$625-732	$(1.6 \pm 0.8) \times 10^{-2}$
LBA288 / RP4Δ*traI* / pBFFCre	0
LBA288 / RP4Δ*traI* / pBFFCre::$VirF_{Ti}$166-202	$(1.8 \pm 0.1) \times 10^{-8}$
LBA1100 / RP4Δ*traI* / pBFFCre::$TraI_{RP4}$625-732	$(5.4 \pm 1.9) \times 10^{-3}$
LBA1100 / RP4Δ*traI* / pBFFCRE::$VirF_{Ti}$166-202	$(3.4 \pm 2.1) \times 10^{-3}$

Table continued on the next page.

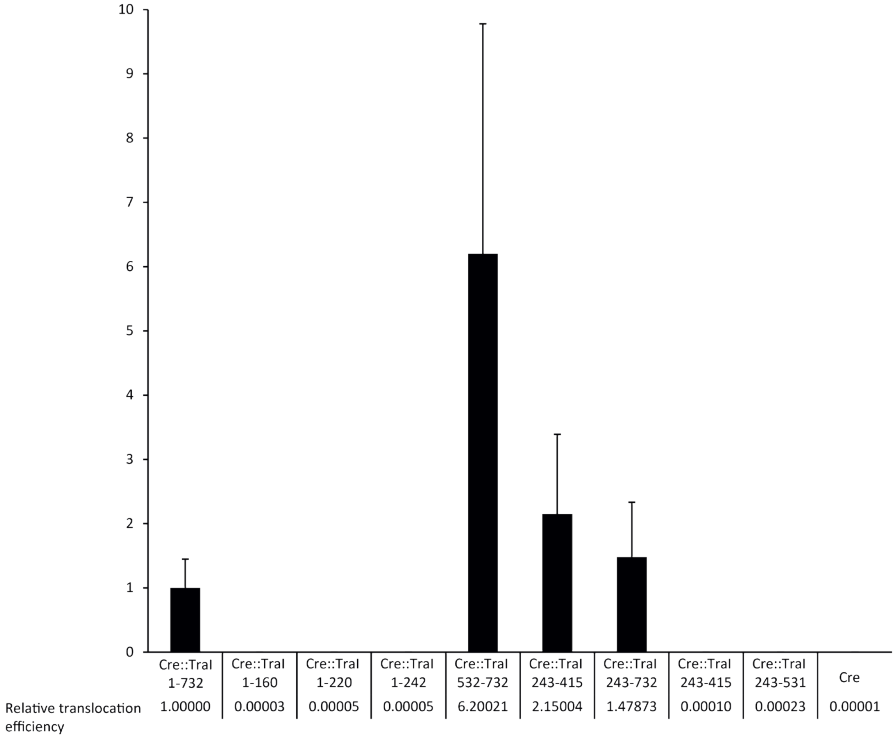
Figure S2 protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Experiment #1

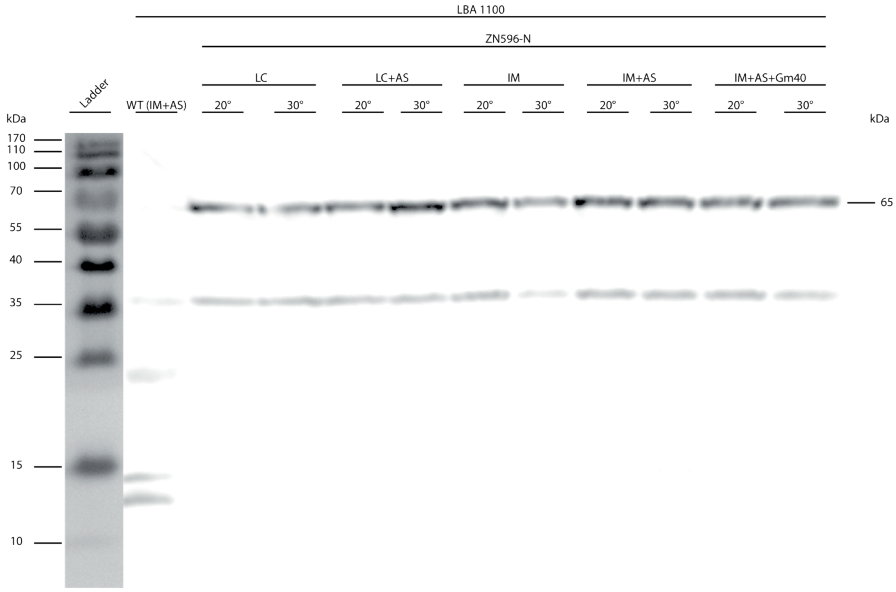
Constructs used	**Average translocation efficiency (n = 3)**
RP4 / Cre::$TraI_{RP4}$1-732	$(5.3 \pm 1.3) \times 10^{-3}$
RP4 / Cre::$TraI_{RP4}$1-160	$(1.4 \pm 0.2) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$1-220	$(2.9 \pm 0.4) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$1-242	$(2.5 \pm 0.3) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$532-732	$(3.3 \pm 1.0) \times 10^{-2}$
RP4 / Cre	$(5.9 \pm 2.5) \times 10^{-8}$

Experiment #2

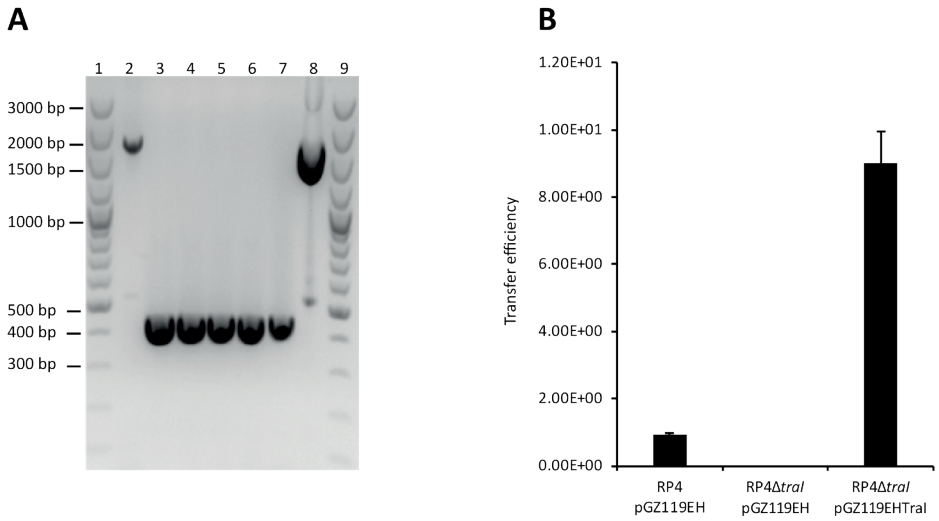
Constructs used	**Average translocation efficiency (n = 3)**
RP4 / Cre::TraIRP41-732	$(3.1 \pm 3.1) \times 10^{-3}$ (n=2)
RP4 / Cre::TraIRP4243-415	$(1.0 \pm 0.1) \times 10^{-2}$
RP4 / Cre::TraIRP4243-415	$(6.9 \pm 1.8) \times 10^{-3}$
RP4 / Cre::TraIRP4243-531	$(4.7 \pm 2.4) \times 10^{-7}$
RP4 / Cre	$(1.6 \pm 0.1) \times 10^{-8}$



Supplemental Figure S2. Preliminary characterization of the TS of the relaxase $TraI_{RP4}$ using a CRAfT assay. This CRAfT assay showed the normalized relative averaged translocation efficiency of a Cre fusion proteins compared to the full-length $TraI_{RP4}$ translocation efficiency from *E. coli* DH5α to *E. coli* CSH26Cm:LTL. The constructs used were as follows: Cre::TraI (full-length $TraI_{RP4}$ sequence 1-732 AA); Cre::TraI 1-160 ($TraI_{RP4}$ truncation from 1-160 AA); Cre::TraI 1-220 ($TraI_{RP4}$ truncation from 1-220 AA), Cre::TraI 1-242 ($TraI_{RP4}$ truncation from 1-242 AA); Cre::TraI 532-732 ($TraI_{RP4}$ truncation from 532-732 AA); Cre::TraI 243-415 ($TraI_{RP4}$ truncation from 243-415 AA); Cre::TraI 243-732 ($TraI_{RP4}$ truncation from 243-732 AA); Cre::TraI 243-415 ($TraI_{RP4}$ truncation from 243-415 AA); Cre::TraI 243-531 ($TraI_{RP4}$ truncation from 243-531 AA) and Cre (full-length Cre protein). In this experiment, there was an issue detected with constructs Cre::$TraI_{RP4}$ 1-160 AA, $TraI_{RP4}$ 1-220 AA, $TraI_{RP4}$ 1-242 AA, $TraI_{RP4}$ 243-415 AA, and $TraI_{RP4}$ 243-531 AA, as there was no stop codon present to terminate translation at the desired position within $TraI_{RP4}$ domains; translational stop occurred after incorporation of an additional and arbitrary C-terminal 60 AA. All other construct used had the intended stop codon present at the end. This series of CRAfT assays clearly demonstrated that the TS was most likely located in the last C-terminal 200 AA. Two independent experiments were normalized to the translocation efficiency of full-length $TraI_{RP4}$ fused to the Cre protein that was averaged and set to 1 for the basal translocation rate. All other translocation efficiencies were averaged and divided by the Cre::$TraI_{RP4}$ basal rate of its independent experiment. The y-axis depicts the normalized relative translocation frequency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The Error bars depict SEM (n = 3; except for Cre::TraI with n = 5; and Cre with n = 6). The transformation efficiencies of the experiments can be found in **figure S1**.



Supplemental Figure S3. Protein expression of *A. tumefaciens* expression plasmid pBFF under the control of $virF_{Ti}$ promoter in different conditions. This western blot shows the ZN596-N protein expression in different temperature and media conditions. The only Cre construct with a FLAG tag available for testing at the time point was the ZN596-N protein. ZN696-N is a Cre recombinase, fused with a 6 zinc-finger DNA binding motif and a $VirF_{Ti}$ TS as well as a FLAG tag, encoded on a pBFF plasmid (van Kregten *et al.*, 2009) which is present in *A. tumefaciens* LBA1100 and cultured in induction medium (IM; **Chapter 2**) or Luria broth (LB), either with or without 200 μM acetosyringone (AS) and 40 μg/mL gentamycin (Gm40). For detection, a monoclonal FLAG tag antibody was used (Sigma). The size of the ZN596-N protein is around 65 kDa. The bands detected around 36 kDa could be due to degradation of the recombinant protein and/or due to unspecific binding of the antibody. *A. tumefaciens* LBA1100 not containing the construct, in lane 2, shows minimal unspecific binding around those sizes as well.



Supplemental Figure S4. **A**: Gel electrophoresis of PCR products indicative of RP4Δ*traI* deletion plasmids. All plasmids were isolated from *E. coli* DH5α. Primers TraICheckFW and TraICheckRV were used. The expected fragment size of $traI_{RP4}$ with those primers was to be expected around 1700 bp and with a deletion of $traI_{RP4}$ the fragment size was to be expected around 400 bp. Lanes 1 and 9: contain the 100 bp plus DNA ladder (Thermo Scientific); Lane 2: colony preparation of wild-type RP4; Lane 3: RP4Δ*traI* clone 1; Lane 4: RP4Δ*traI* clone 2; Lane 5: RP4Δ*traI* clone 3; Lane 6: RP4Δ*traI* clone 4; Lane 7: RP4Δ*traI* clone 5; Lane 8: midi-prep of wild-type RP4. **B**: Conjugation experiment to determine the transfer of the RP4Δ*traI* knock-out plasmids with and without the complementation of the relaxase $TraI_{RP4}$. As donor strain *E. coli* DH5α and as recipient strain *E. coli* MS614 was used. The wild-type RP4 was used as a positive control to show self-transfer in the presence of the plasmid pGZ119EH. RP4Δ*traI* was not able to transfer itself in the presence of the plasmid pGZ119EH. RP4Δ*traI* was able to transfer itself when the complementing $traI_{RP4}$ was present on plasmid pGZ119EHTraI. Conjugation time was 2 h. Transfer efficiency (on y-axis) was expressed as the fraction of transconjugant cells with kanamycin (50 μg/mL) and streptomycin (25 μg/mL) resistance over the number of donor cells with kanamycin (50 μg/mL) and chloramphenicol (10 μg/mL) resistance. In transconjugant cells, the kanamycin resistance marker would be present on the transferred RP4 plasmid, and the streptomycin resistance marker is present in the genome of the recipient strain *E. coli* MS614. The donor cells contain a kanamycin marker on the RP4 plasmid and the chloramphenicol marker on the plasmid pGZ119EH or pGZ119EHTraI. The error bar depicts the SEM (n = 3).