1 INTRODUCTION

Type IV secretion is an energy-driven mechanism for delivery of defined substrates from a bacterial donor cell to a recipient cell. It enables the transfer of DNA between bacteria (bacterial conjugation) and is used by pathogenic bacteria to transfer toxic proteins or other virulence factors into eukaryotic hosts. A set of genes is conserved among type IV secretion systems (T4SS). Most of these genes are responsible for the formation of a membrane-spanning protein complex and for biosynthesis of a pilus (19). According to the nomenclature of the *Agrobacterium tumefaciens* T-DNA transfer system, this set of genes is called the VirB operon. An additional component conserved among T4SS is the VirD4-like protein (TraG-like protein, coupling protein).

1.1 Bacterial conjugation

Bacterial conjugation is responsible for the spread of genetic traits among a broad range of bacterial species. It is the primary mechanism for dissemination of antibiotic resistances among human pathogens (106). Bacterial conjugation was discovered in 1946 by Lederberg and Tatum (52), who observed genetic recombination between auxotrophic derivatives of *Escherichia coli* K12. The sexual factor required for this ability was later identified as the fertility factor (F factor, F plasmid, ref. 15, 34). Nearly all functions required to mediate bacterial conjugation are encoded by conjugative plasmids like the F plasmid. Such plasmids frequently contain antibiotic resistance genes (R factor plasmids) and several copies of insertional elements (IS) that allow homologous recombination with the chromosome. Bacterial conjugation is most frequently observed between Gram-negative, as well as between Gram-positive bacteria. Depending on the host range, a conjugative plasmid can replicate and be stably maintained in the host cell after successful transfer. Broad host range plasmids like RP4 can replicate in a wide range of Gram-negative bacteria. Furthermore, such plasmids are transferable from Gram-negative into Gram-positive bacteria (101) or even into yeast (35) and mammalian cells (107), although at a much lower transfer rate. Conjugative plasmids are grouped into incompatibility (Inc) groups, since plasmids belonging to the same Inc group are incompatible, i.e. these plasmids cannot coexist within one host.

The mechanism of DNA transfer was initially studied with the F plasmid (IncF). A series of other conjugative plasmids, like R1 (IncF), RP4 (IncP), ColIb (IncI), pKM101 (IncN) and R388 (IncW) were discovered later. Despite some differences in sequence, host range and compatibility, these conjugation systems share high functional similarity with the F plasmid system. The investigations carried out with these conjugative plasmids each contributed in unraveling some of the fundamental mechanisms of bacterial conjugation (Figure 1.1).

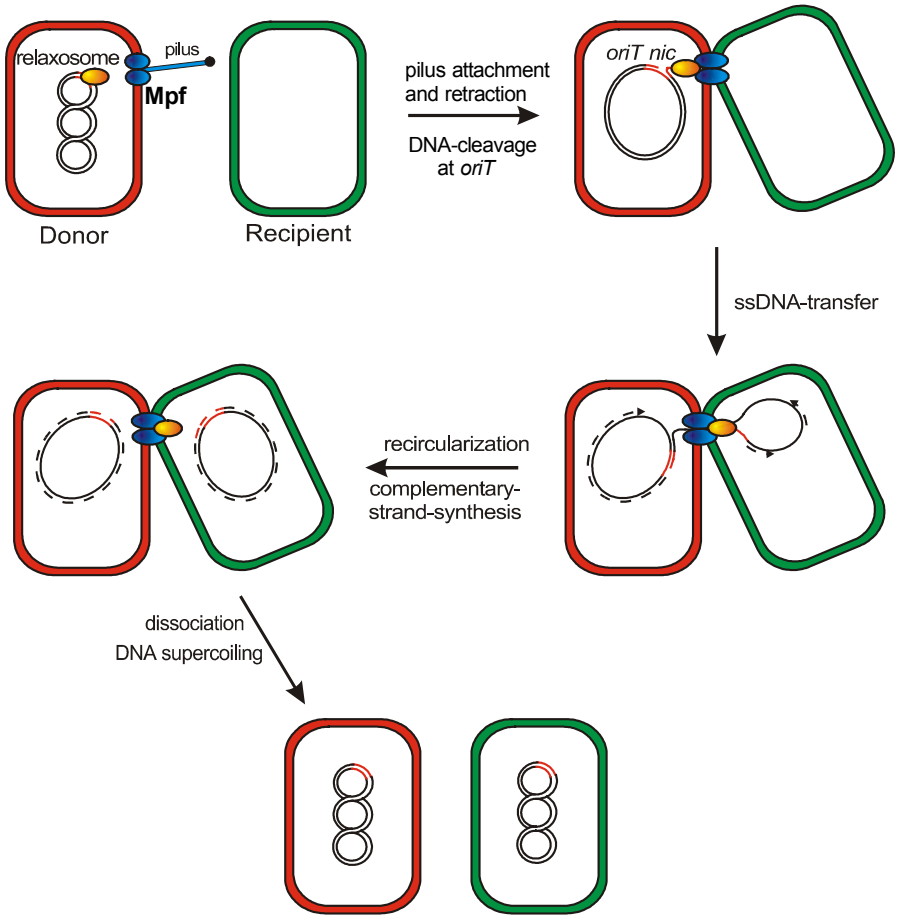

Figure 1.1. General mechanism of bacterial conjugation. Explanations are given in the text.

The plasmid-encoded transfer (Tra) proteins are grouped into functional classes defined as those involved in mating pair formation (Mpf) and DNA processing (DNA transfer and replication, Dtr). The Mpf system comprises the subunits for pilus biogenesis and a set of proteins that probably form a large, membrane-spanning complex. In the mechanism of bacterial conjugation, the pilus is thought to mediate the initial contact with a recipient cell and to retract or disassemble after cell adsorption (Figure 1.1). Donor and recipient cells then come into intimate physical contact and a mating bridge is formed. As a next step, the Dtr system

intitiates processing of the plasmid DNA by formation of the relaxosome (29).
The relaxosome is a complex of several DNA processing proteins (relaxosomal proteins) bound to a specific target sequence of the conjugative plasmid, the origin of transfer (*oriT*). The catalytic key component of the relaxosome is the relaxase. The relaxase produces a single-stranded scission at the *nic* cleavage site within *oriT* and becomes transiently attached to the 5' end of the DNA single strand (71). Additionally, the relaxase is coupled to the TraG-like protein, which is a membrane component that possibly forms a transport pore through the inner membrane. After cleavage at the *nic* site ("nicking"), the conjugative plasmid becomes relaxed and the cleaved DNA strand is transferred into the recipient cell, probably through a rolling-circle replication mechanism. DNA transfer proceeds with 5' to 3' polarity. During transfer, the 5'-attached relaxase possibly remains associated within the DNA transport channel, scanning the incoming DNA for the reconstituted *nic* site (111). When this reconstituted *nic* site passes the relaxase, a second strand transfer reaction takes place, recircularizing the transferred DNA single strand (110). Complementary strand synthesis in the donor and recipient cells, supercoiling of the covalently closed plasmids, and active dissociation of the mating partners complete the conjugative process.

1.2 Virulence-associated type IV secretion systems

Several pathogens deliver effector molecules to eukaryotic cells by means of secretion systems whose subunits are evolutionarily related to the Mpf complex and the TraG-like protein of bacterial conjugation systems. These secretion systems were named type IV secretion systems (T4SS), as originally proposed by Salmond (79). A few examples of T4SS are described in the following, starting with the *A. tumefaciens* T-DNA transfer system encoded by the Ti plasmid. This secretion system is the first characterized T4SS mediating transport from bacterial donors into eukaryotic hosts. It has been the reference for the T4SS studied later.

The *Agrobacterium tumefaciens* T-DNA transfer system

A. tumefaciens is a phytopathogen that induces tumorous growth (crown gall tumor) of infected plant tissues (92). It uses a T4SS to transfer oncogenic T-DNA and several effector proteins to the nuclei of plant cells. The genes encoding the

T4SS of *A. tumefaciens* map on the tumor-inducing plasmid (Ti plasmid). The known effector proteins translocated during plant infection are VirD2, VirE2, VirE3 and VirF. The function of VirF is uncertain. It may be involved in the targeted proteolysis of specific host proteins in early stages of the transformation process (86). VirD2 is a relaxase that, similar to relaxases of bacterial conjugation systems, acts in cleavage of the plasmid DNA. VirD2 specifically recognizes and cleaves the border repeat sequences flanking the T-DNA on the Ti plasmid. VirE2 is a single-stranded DNA-binding protein (SSB). The role of the virulence factor VirE3 is unknown. After entering the plant cell, VirE2 binds cooperatively to the VirD2-T-DNA conjugate. By virtue of nuclear localization sequences, VirE2 and VirD2 then guide the T-DNA to the nucleus of the plant cell. Here, the T-DNA becomes integrated into the plant genome by illegitimate recombination. The T-DNA contains plant oncogenes and genes encoding substrates collectively known as opines, which can be specifically consumed by the infecting *A. tumefaciens* cells. The oncogenes code for enzymes that catalyze the production of plant growth hormones, which in turn cause the formation of non-self-limiting tumors. The discovery of the *A. tumefaciens* T-DNA transfer system has led to the development of effective tools for the genetic transformation of plants [(42, 115) for review].

The *Helicobacter pylori* T4SS encoded by the *cag* pathogenicity island

H. pylori is a human pathogen causing gastric diseases such as gastritis, peptic ulcer and gastric cancer. Pathogenicity of *H. pylori* was found to be associated with the presence of the cancer-associated genes (*cag*) that localize on the pathogenicity island (PAI) (16). The *cag*-PAI consists of 27 putative genes, six of which show homologies to the *virB*/*virD* genes of *A. tumefaciens* or other genes encoding type IV secretion components. The T4SS encoded by *cag*-PAI directs the transfer of the 145 kDa CagA protein into gastric epithelial cells. After transfer, CagA undergoes tyrosine phosphorylation and induces changes in the tyrosine phosphorylation state of distinct cellular proteins (68, 89, 94). Independently of CagA secretion, *H. pylori* strains carrying the *cag*-PAI furthermore induce the production and secretion of chemokines such as interleukin-8 (IL-8) in gastric epithelial cells. Apart from *cag*-PAI, *H. pylori* encodes two other T4SS: the *comB* system for DNA-uptake (38) and the *tra*

region, which possibly enables conjugative transfer of chromosomal DNA between *H. pylori* cells (2). The different T4SS found in *H. pylori*, enabling both import and export of DNA or protein substrates, are a remarkable example for the diversity and the range of application of these secretion systems.

The *dot/icm* system of *Legionella pneumophila*

L. pneumophila is an intracellular pathogen. It is the causative agent of a potentially fatal form of pneumonia known as Legionnaires' disease. In the environment, *L. pneumophila* survives and replicates within amoebae. When infecting mammalians, it can alter the endocytic pathway of macrophages and replicate in the lungs of the host. *L. pneumophila* encodes a T4SS, the *dot/icm* system (*d*efect in *o*rganelle *t*rafficking/ *i*ntra*c*ellular *m*ultiplication), that is postulated to contribute to intracellular survival of *L. pneumophila*. Following contact formation with a macrophage, the bacterium induces the formation of a pore and enters the vacuole of the host cell. Such formed phagosomes containing *L. pneumophila* become specialized organelles that escape endosome or lysosome fusion. Trafficking of the phagosomes is controlled by the *dot/icm* system, which is thought to export effector molecules that regulate the maturation of the phagosomes [(91) for review]. *L. pneumophila* encodes a second T4SS (the *lvh* system), which possesses the ability for bacterial conjugation. It was however shown that this T4SS, although interfering with the *dot/icm* system, is not required for infection of macrophages and is thus not related to pathogenicity of *L. pneumophila* (90).

The T4SS of *Bartonella* spp.

Bartonellae are pathogenic bacteria adapted to cause intraerythrocytic infection in their mammalian hosts. The three major human pathogens are *Bartonella bacilliformis*, leading to Oroya fever and verruga peruana, *Bartonella quintana*, causing trench fever, and *Bartonella henselae*, the cause of cat-scratch disease, bacillary angiomatosis and peliosis, endocarditis, bacteraemia and neuroretinitis. As yet, sequence analysis of pathogenic *Bartonella* strains has revealed the presence of a T4SS in *B. henselae* and of two distinct T4SS in *Bartonella tribocorum*. The first one is most similar to the *A. tumefaciens* VirB/VirD system, whereas the other shares highest similarity with the conjugation system (Trw system) of plasmid R388 (6). The VirB/VirD-like T4SS of *B. tribocorum* was shown to be a factor mediating pathogenicity (88). The effector molecules

secreted by this secretion system and the mechanism by which intracellular survival of the bacteria is made possible remain to be determined. Equally, the role of the second, Trw-like T4SS is yet unknown, but investigations are underway (C. Dehio, personal communication).

The T4SS of *Brucella* spp.

Like *Legionella* and *Bartonella*, *Brucella* is an intracellular parasite of macrophages. Brucellae cause a chronic infectious disease in mammals, known as brucellosis or Malta fever. After invasion through wounds or mucosa, *Brucella* initially infects professional phagocytes. Subsequent infection occurs preliminarily in nonprofessional phagocytic cells in the tissue of reproductive organs, the placenta or the fetus. Therefore, typical manifestations of brucellosis are abortion in pregnant females (which is the origin of the name of *Brucella abortus*) and sterility in males. Intracellular survival and replication of *Brucella* are mediated by a T4SS that initiates a process involving the maturation of the pathogen-containing vacuole [(91) for review]. Similar to the *dot/icm* system of *L. pneumophila*, the T4SS of *Brucella* spp. alters the endocytic pathway of the infected macrophage and blocks the fusion of the phagosome with the lysosome. The bacteria proliferate in this phagosome, which forms a compartment bound by the endoplasmatic reticulum. The effector molecules secreted by the T4SS of *Brucella* are yet unknown.

The Ptl system of *Bordetella pertussis*

B. pertussis is the causative agent of whooping cough. The pathogenicity determinant of this bacterium is a T4SS called the Pertussis toxin liberation (Ptl) system (12, 21), which transports its substrate, the Pertussis toxin (PT), to the extracellular milieu. PT is a multisubunit toxin of the A/B family and is composed of five subunits forming a hexamer structure. The individual subunits are secreted into the periplasm by means of the general secretion pathway (GSP) and assemble into the holotoxin, which is then secreted into the exocytoplasm by the Ptl system. After secretion, the B domain of PT interacts with glycoprotein receptors of the mammalian host and targets the A domain of the toxin into the cytoplasm. The A domain then ADP-ribosylates protein components of the host cell, thus interfering with receptor-mediated activation and associated signalling pathways (12). The T4SS of *B. pertussis* is exceptional in the way that it delivers its substrate from the bacterial periplasm into the exocytoplasm instead of

directly transporting it from the cytoplasm into the host cell. This unique mechanistic exception comes along with the lack of an otherwise essential conserved T4SS component, the VirD4/TraG-like protein, giving a strong indication for the role of this component in the other T4SS (see below).

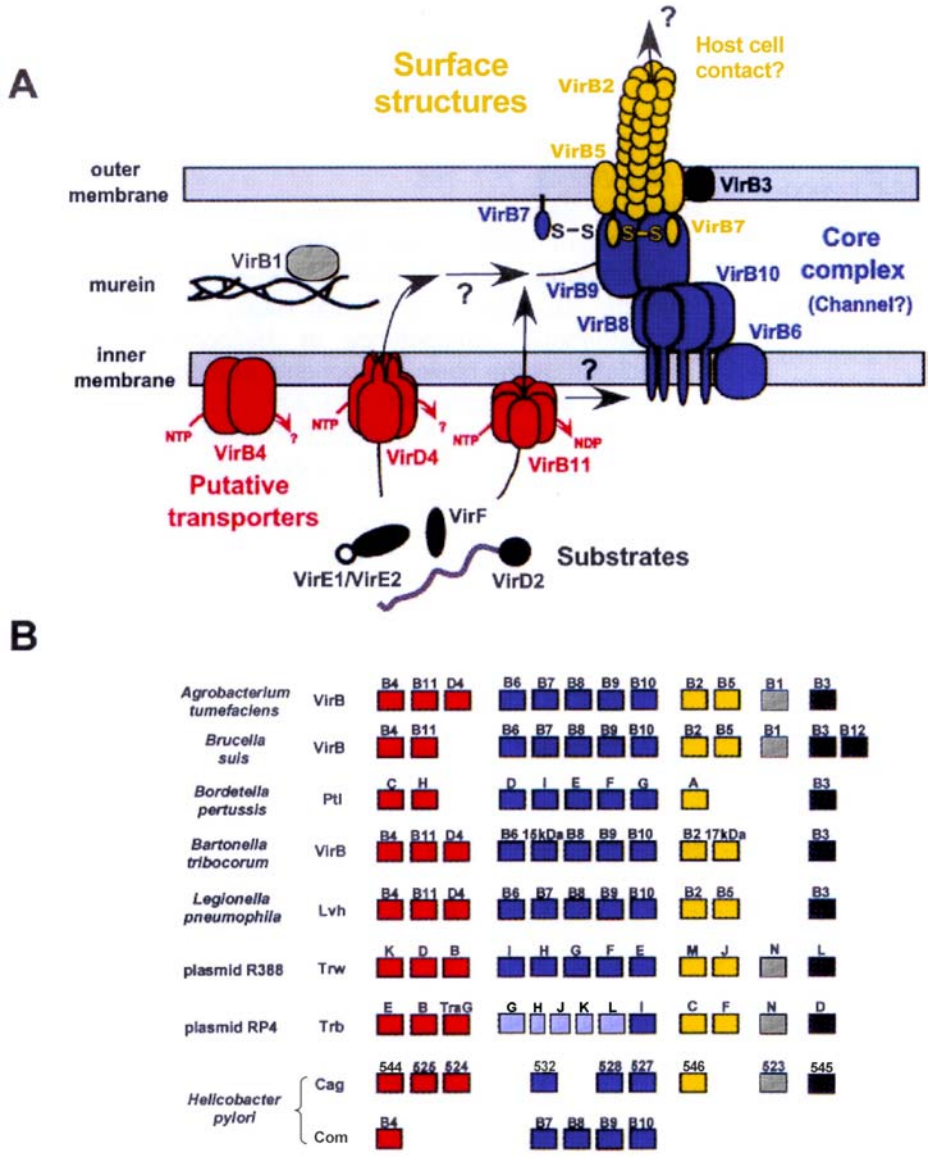

Figure 1.2. Conserved components of T4SS and their known or putative functions in the type IV secretion machinery. A. Model of the type IV secretion apparatus. The protein components are named according to the nomenclature of the *A. tumefaciens* VirB/VirD system. The cellular localization, oligomeric state and function of each component is indicated, provided that it has been determined for at least one representative of a T4SS. The association of energy-generating transporters needs to be defined, as well as the passage of substrates from the cytoplasm through the core complex or the translocation *via* the periplasm. Surface structures putatively initiate contact with host cells, but the substrate translocation pathway is unknown. The transglycosylase VirB1 and homologues may locally lyse the murein and thereby facilitate complex assembly. VirB7 may play a dual role stabilizing the complex by both association with VirB9 and association with the pilus. **B.** Alignment of conserved VirB/VirD components of different T4SS. The proteins were grouped according to their putative function(s): transporters (red), core components (blue), surface-associated proteins (yellow), lytic tranglycosylases (grey) and proteins without an assigned function (black). Light blue filling for some of the RP4 Trb proteins indicates that these proteins belong to the core complex but do not share significant sequence similarity with the VirB components of the other T4SS. [This figure originates from a review article by C. Baron *et al.* (6) and was slightly modified.]

1.3 The VirB/VirD4 transport machinery of type IV secretion

A set of membrane-associated and transmembrane proteins mediating the secretion of protein and/or DNA substrates is conserved in T4SS. In bacterial conjugation systems, these T4SS components were named the Mpf-complex, consisting of 11 proteins, and the TraG-like protein (named for the protein of IncP plasmid RP4, ref. 55). The related components of the *A. tumefaciens* T4SS and of other virulence-associated T4SS are the virulence (Vir) factors VirB (VirB1-VirB11) and VirD4, respectively.

The function of the VirB/VirD-like components and their localization in the transport apparatus has been partially characterized (Figure 1.2). The energy necessary for trafficking of the substrate molecules is probably delivered by nucleoside triphosphate (NTP) hydrolysis. Three putative NTPases (VirB4-, VirB11- and VirD4-like proteins) were identified among the T4SS components. However, NTP hydrolysis was demonstrated only for VirB11-like proteins (44, 45, 77). The pilus is formed by the VirB2-like protein (pilin) as the major structural component. Pilin is a cyclic peptide in case of VirB2 (*A. tumefaciens*) and TrbC (RP4) (26). Attached to the pilus are the VirB components VirB5 and VirB7. VirB7 functions as an interface between pilus and core complex by forming a heterodimer with the core component VirB9.

The core complex consists of a minimal set of three membrane proteins, VirB8-VirB10, that interact with one another. VirB3, VirB4 and VirB6 are probably also components of the core complex, since they each contribute to the stability of other core complex components or are themselves stabilized by such components. The putative transporter protein of the VirB11-type is a hexamer that probably interacts with the core complex at the cytoplasmic side of the inner membrane [(19) for review]. The VirD4-like putative transporter is equally located at the cytoplasmic side of the membrane, but it additionally contains a membrane anchor with a small periplasmic domain. It is unclear, whether VirD4-like proteins interact with any component of the VirB core complex. Possibly, these proteins exert a separate, VirB-independent, transport-related activity. A detailed description of VirD4/TraG-like proteins is given in the following.

1.4 The role of VirD4/TraG-like proteins in T4SS

TraG-like proteins are essential components of conjugative DNA transfer systems and the related T4SS (19), with exception of the Ptl system of *B. pertussis* (see above). In bacterial conjugation systems, the TraG-like protein is the only RP4 component needed in addition to the Mpf system for mobilization of the non self-transmissible IncQ plasmid RSF1010 (7, 13, 54). TraG proteins have been named 'coupling proteins' since they were proposed to mediate specific interaction between the Dtr and the Mpf functions (13, 14, 32, 54). Physical interactions between relaxosomal proteins and TraG-like proteins are suggested by genetic experiments (13, 14, 32, 54) and biochemical data (25, 43, 72, 97). In contrast, no direct evidence for interactions between a TraG-like protein and an Mpf/VirB component has been reported so far. TraG-like proteins share two common motifs resembling the Walker A and B motifs of NTPases and ABC transporters (55, 56, 85). These motifs were shown to be essential for transfer activity (4, 64), indicating that the postulated NTPase activity of TraG-like proteins plays an important role in type IV secretion.

The biochemical properties of two TraG-like proteins, TraD (F) and TrwB (R388), have been partially analyzed in the past. TraD was characterized as an inner membrane protein with potential DNA binding ability and ATPase activity (69). However, no ATPase activity has been detected for a soluble fragment of TrwB, lacking the transmembrane part [TrwBΔN70, ref. (64)], and the ATPase activity present in TraD was assigned to an impurity (this study). TraD was shown to interact with TraM of F (25), a relaxosomal component binding to three different sites within the *oriT* region of F (24). Membrane topology analysis of TraD revealed that TraD has a short cytoplasmic N-terminus, followed by a periplasmic domain of about 60 residues and a long cytoplasmic C-terminal tail (53). A similar topology was determined for the related VirD4 protein of *A. tumefaciens* (22).

The cytoplasmic domain of the TraG-like protein of plasmid R388, TrwBΔN70, has a hexameric pore-like structure that probably extends into the membrane (31, 39), indicating that TraG-like proteins may serve as transporters for gating of substrates through the inner membrane. Although TrwBΔN70 failed to hydrolyze nucleotides, it was found to bind to ATP (64). The nucleotide binding site of

TrwB was mapped in crystals of the nucleotide-bound form of the protein (30). Apart from binding to ATP, TrwBΔN70 binds to double-stranded (ds) and single-stranded (ss) DNA (64). Furthermore, TrwBΔN70 was reported to enhance relaxase activity of R388 TrwC and topoisomerase activity of *E. coli* topoisomerase I (64).

The whole genome sequence of *H. pylori* (100) revealed a TraG-like protein, HP0524, which is encoded by a gene of the *cag* PAI (see 1.2). HP0524 is essential for CagA export (94), but is dispensable for interleukin-8 (IL-8) induction in gastric epithelial cells (27).

1.5 The aim of this work

In the present work, it is envisaged to gain insights into the function of TraG/VirD4-like proteins of type IV secretion systems (T4SS). Different representatives of TraG-like proteins from conjugative plasmids and other T4SS are to be purified and characterized biochemically. Preparatory work in this field has been carried out with TraG (RP4) and TraD (F) in the laboratory of Dr. Lanka (43). The study of TraG-like proteins will be continued and extended to TrwB (R388) and HP0524 of the pathogenicity-related T4SS of *H. pylori*. Deletion- and point mutation derivatives of these TraG-like proteins will be constructed and purified. The ability to hydrolyze or bind nucleotides, to bind DNA and to interact with the relaxase TraI will be studied in detail. The results will help to dissect the different activities of TraG-like proteins and to assign functions to different domains or sequence signatures. A study for determination of the membrane topology of TraG has been underway since 1999 in cooperation with the work group of Dr. Traxler (University of Washington, Seattle). This study will be completed and extended in order to determine the membrane topology of TraG and to identify functionally essential domains. The comparative analysis of different members of the family of VirD4/TraG-like proteins will help to generalize the results and to make the conclusions applicable to other T4SS, with the objective to unravel the key role played by VirD4/TraG-like proteins in the mechanism of type IV secretion.