*Microbiology* (2010), **156**, 1963–1974 DOI 10.1099/mic.0.039248-0

# Membrane topology of conserved components of the type III secretion system from the plant pathogen *Xanthomonas campestris* pv. vesicatoria

Carolin Berger,[1]† Guillaume P. Robin,[2] Ulla Bonas[1] and Ralf Koebnik[1,2]

Correspondence
Ralf Koebnik
koebnik@gmx.de

[1]Institute of Biology, Department of Genetics, Martin-Luther-University, 06099 Halle, Germany

[2]Laboratoire Génome et Développement des Plantes, Université de Perpignan via Domitia–CNRS–IRD, UMR 5096, IRD Montpellier, France

Type III secretion (T3S) systems play key roles in the assembly of flagella and the translocation of bacterial effector proteins into eukaryotic host cells. Eleven proteins which are conserved among Gram-negative plant and animal pathogenic bacteria have been proposed to build up the basal structure of the T3S system, which spans both inner and outer bacterial membranes. We studied six conserved proteins, termed Hrc, predicted to reside in the inner membrane of the plant pathogen *Xanthomonas campestris* pv. vesicatoria. The membrane topology of HrcD, HrcR, HrcS, HrcT, HrcU and HrcV was studied by translational fusions to a dual alkaline phosphatase–$\beta$-galactosidase reporter protein. Two proteins, HrcU and HrcV, were found to have the same membrane topology as the *Yersinia* homologues YscU and YscV. For HrcR, the membrane topology differed from the model for the homologue from *Yersinia*, YscR. For our data on three other protein families, exemplified by HrcD, HrcS and HrcT, we derived the first topology models. Our results provide what is believed to be the first complete model of the inner membrane topology of any bacterial T3S system and will aid in elucidating the architecture of T3S systems by ultrastructural analysis.

Received 1 March 2010
Accepted 2 April 2010

## INTRODUCTION

Type III secretion (T3S) systems have evolved in Gram-negative bacteria to assemble flagella, secrete extracellular proteins, and deliver so-called effector proteins into the cytoplasm of eukaryotic cells. Nonflagellar T3S systems are important pathogenicity determinants and essential for the interaction of most Gram-negative bacterial pathogens with their eukaryotic hosts (Cornelis, 2006; Pallen *et al.*, 2005). T3S systems that transport proteins in a one-step process across the two bacterial membranes and the host plasma membrane directly into the host cell's cytoplasm have been identified in animal pathogens, e.g. *Yersinia* spp., *Salmonella* spp., *Shigella* spp. and *Escherichia coli* (Coburn *et al.*, 2007), and in most plant pathogens, e.g. *Pseudomonas syringae*, *Ralstonia solanacearum*, *Erwinia* spp. and *Xanthomonas* spp. (Büttner & Bonas, 2006; McCann & Guttman, 2008). Moreover, several symbiotic bacteria also express T3S systems (Preston, 2007).

Nonflagellar T3S systems are encoded by 20–25 genes (Cornelis, 2006). Many structural components are conserved among pathogenic and symbiotic bacteria, and are strictly required to build up a functional secretion system. Nine components of the apparatus are conserved between nonflagellar T3S systems and the flagellar export apparatus (He *et al.*, 2004; Nguyen *et al.*, 2000). The broadly conserved proteins are believed to localize in the cytoplasm or constitute the inner ring structure of the secretion apparatus. Among them are six predicted integral inner membrane proteins forming the inner core of the T3S system, which probably acts as a secretion pore (Aizawa, 2001; Pühler *et al.*, 2004; Tampakaki *et al.*, 2004). The flagellar counterparts of five of them, FlhA, FlhB, FliP, FliQ and FliR, have been shown to constitute the flagellar export apparatus (Fan *et al.*, 1997; Minamino & Macnab, 2000b). The related nonflagellar T3S proteins also show manifold interactions with each other, thus suggesting similar multiprotein complexes for the flagellar and nonflagellar T3S systems (Creasey *et al.*, 2003). This conclusion is supported by low-resolution electron microscopy, which demonstrates that the morphology of the nonflagellar T3S apparatus is remarkably similar to that of the flagellar

†Present address: Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstraße 3, 06466 Gatersleben, Germany.

Abbreviations: Rose-Gal, 6-chloro-3-indolyl-$\beta$-D-galactopyranoside; T3S, type III secretion; TMS, transmembrane segment; X-phos, 5-bromo-4-chloro-3-indolyl phosphate, disodium salt.

A supplementary table, showing the oligonucleotides used in this study, and three supplementary figures, showing amino acid sequences, the PhoA activity of HrcR reporter fusions, and a multiple sequence alignment of HrcV homologues, are available with the online version of this paper.

hook–basal body complex (Kubori *et al.*, 1998; Sekiya *et al.*, 2001; Tamano *et al.*, 2000). The main structural differences are found in the extracellular appendages. While non-flagellar T3S systems of animal pathogens are associated with a needle-like structure, plant pathogens possess a pilus-like structure on the surface, the Hrp pilus (Cornelis, 2006; He & Jin, 2003).

Nonflagellar T3S has been studied best in *Yersinia* species, and the corresponding components are termed the *lcr* (low calcium response), *ysc* (Yop secretion) and *yop* (yersinial outer protein) genes (Cornelis, 2006). Nevertheless, structural information is still very limited, which holds particularly true for the six inner membrane components YscV, YscU, YscT, YscS, YscR and YscD. The YscV (formerly LcrD) protein influences type III effector secretion and is related in sequence to the *Salmonella* inner membrane protein FlhA. Both YscV and FlhA are predicted to have seven transmembrane helices, with a cytoplasmic N terminus and a periplasmic C terminus (Melen *et al.*, 2003). These predictions are in contrast to experimental data that support a cytoplasmic location of the C-terminal domain of FlhA and YscV (Minamino & Macnab, 2000b; Plano *et al.*, 1991). FlhA associates with FlhB (Zhu *et al.*, 2002), which is related to the YscU protein. YscU, like FlhB, is thought to have four transmembrane helices followed by a cytoplasmic domain (Allaoui *et al.*, 1994; Minamino *et al.*, 1994). FlhB controls the substrate specificity and switches to the export of late structural subunits of the flagellum upon completion of early flagellar structures (Ferris & Minamino, 2006; Fraser *et al.*, 2003). This transition is accompanied by a proteolytic cleavage event in the C-terminal cytoplasmic domain of FlhB (Minamino & Macnab, 2000a). YscU plays a similar role in the nonflagellar T3S system (Edqvist *et al.*, 2003; Lavander *et al.*, 2002). Proteolytic cleavage in the C-terminal cytoplasmic domain of YscU has been found to be necessary for Yop translocator secretion (Sorg *et al.*, 2007). YscR, YscS and YscT are predicted to reside in the inner membrane with multiple transmembrane helices (Fields *et al.*, 1994; Ghosh, 2004), and the flagellar counterparts, FliP, FliQ and FliR, have been experimentally identified in the inner membrane (Ohnishi *et al.*, 1997). YscD has also been found in the inner membrane (Plano & Straley, 1995), and shares a low level of sequence conservation with its homologues PrgH, MxiG and HrcD (formerly HrpQ in *P. syringae*) (Ghosh, 2004; Pühler *et al.*, 2004). The YscD/HrpQ protein family (InterPro: IPR012843; http://www.ebi.ac.uk/interpro/IEntry?ac=IPR012843) is thought to form a multimeric ring, with each subunit being anchored in the inner membrane by one transmembrane helix (Moraes *et al.*, 2008).

To date, only a few high-resolution structures of T3S components have been solved, mainly restricted to soluble proteins or extramembraneous domains of membrane proteins (Deane *et al.*, 2008; Fadouloglou *et al.*, 2004; Moraes *et al.*, 2008; Zarivach *et al.*, 2008). However, there is no high-resolution structural information available for any of the membrane-embedded protein domains, and knowledge about how they assemble into a multiprotein complex in the bacterial membrane is very limited. Because of the difficulties of deriving high-resolution models of membrane proteins, alternative strategies have been developed to obtain structural information at lower resolution, such as the position of transmembrane helices within the polypeptide chain and their orientation in the membrane (van Geest & Lolkema, 2000).

In this study, we applied genetic tools to determine the membrane topology of the conserved inner membrane proteins of a T3S system, using the well-characterized plant pathogen *Xanthomonas campestris* pv. vesicatoria as a model (Gürlebeck *et al.*, 2006).

In our laboratory, we study the causal agent of bacterial spot disease on pepper and tomato, *X. campestris* pv. vesicatoria, which causes important economic losses in growing regions with a warm and humid climate. It has therefore being recommended for regulation as a quarantine pest by the European and Mediterranean Plant Protection Organization (http://www.eppo.org/QUARANTINE/listA2.htm). The interaction between *X. campestris* pv. vesicatoria and its host plants has been established as a model system for the molecular and genetic analysis of pathogenicity and plant resistance (Büttner & Bonas, 2006; Gürlebeck *et al.*, 2006). Basic pathogenicity is determined by a plant-inducible 23 kb *hrp* (hypersensitive reaction and pathogenicity) gene cluster which encodes a T3S system (Bonas *et al.*, 1991; Weber *et al.*, 2007). Most genes of the *hrp* gene cluster are absolutely required for T3S *in vitro* as well as translocation of effector proteins into plant cells (Huguet *et al.*, 1998; Rossier *et al.*, 1999, 2000). Eleven *hrp* genes which encode components of the T3S apparatus are also conserved in animal pathogens, and have been renamed *hrc* genes (*hrp* conserved), using the same letter code as the archetypical *ysc* homologue (Bogdanove *et al.*, 1996; Pühler *et al.*, 2004; Weber *et al.*, 2005). The expression of the *hrp* gene cluster is regulated by environmental signals and involves two regulatory proteins, HrpG and HrpX (Koebnik *et al.*, 2006; Wengelnik & Bonas, 1996; Wengelnik *et al.*, 1996b). Extracellular key components of the *X. campestris* pv. vesicatoria T3S system are the Hrp pilus and the translocon, which together serve as a conduit for protein translocation from the bacterium into the plant host cell (Büttner *et al.*, 2002; Weber *et al.*, 2005). Recently, type III-specific chaperones and substrate specificity control proteins have been identified in our laboratory (Lorenz *et al.*, 2008a, b). To gain novel insights into the molecular functionality of the T3S apparatus, we aimed at determining the membrane topology of all known inner membrane proteins. For this purpose, we performed a transposon mutagenesis leading to dual alkaline phosphatase–$\beta$-galactosidase (PhoA–LacZ$\alpha$) reporter fusions (Alexeyev & Winkler, 2002). Enzymic and molecular characterization of the fusion proteins allowed us to derive at high resolution models of the membrane topology of HrcD, HrcR, HrcS, HrcT, HrcU and HrcV.

## METHODS

**Bacterial strains and growth conditions.** *E. coli* cells were cultivated at 37 °C in lysogenic broth (LB) medium (Bertani, 1951). *X. campestris* pv. vesicatoria cells were cultivated at 28 °C in PSA medium (1 % peptone, 1 % sucrose, 0.1 % glutamic acid). Since the wild-type *X. campestris* pv. vesicatoria strain 85-10 produces large amounts of exopolysaccharides, the exopolysaccharide-negative strain 85E was used (Wengelnik *et al.*, 1996a). Plasmids were introduced into *E. coli* TOP10 cells (Invitrogen) and into *E. coli* S17-1 (Simon *et al.*, 1983) by electroporation. pBBR1-MCS5 constructs were transferred into *X. campestris* pv. vesicatoria strain 85E by biparental conjugation with *E. coli* S17-1 as donor.

Dual indicator plates for *E. coli* consisted of LB medium, 1.5 % (w/v) Bacto agar, 80 μg $ml^{-1}$ 5-bromo-4-chloro-3-indolyl phosphate, di-sodium salt (X-phos; Carl Roth), 100 μg $ml^{-1}$ 6-chloro-3-indolyl-β-D-galactopyranoside (Rose-Gal; AppliChem), 80 mM $K_2HPO_4$, pH 7.0, and 0.2 % L-arabinose (w/v). Indicator plates for *Xanthomonas* consisted of PSA medium, 1.5 % Bacto agar (w/v), 80 μg X-phos $ml^{-1}$ and 80 mM $K_2HPO_4$, pH 7.0. Antibiotics were added to the media at the following final concentrations: 100 μg ampicillin $ml^{-1}$; 25 μg kanamycin $ml^{-1}$; 25 μg gentamicin $ml^{-1}$.

***In vitro* transposon mutagenesis.** Coding sequences for HrcD, HrcR, HrcS, HrcT, HrcU and HrcV were amplified by PCR from genomic DNA of *X. campestris* pv. vesicatoria strain 85-10 (Bonas *et al.*, 1989; Thieme *et al.*, 2005) using *Pfu* DNA polymerase (Stratagene), and cloned into pCR 2.1-TOPO (Invitrogen). For primer sequences, see Supplementary Table S1. Inserts with the correct DNA sequence were subcloned into pBAD24 (Guzman *et al.*, 1995), which contains the $P_{BAD}$ promoter, the activity of which can be induced with 0.2 % L-arabinose. For transposon mutagenesis, plasmid pMA814, which carries the mini-Tn*5* transposon derivative Tn*pholac*1 (Alexeyev & Winkler, 2002), was used. Tn*pholac*1 contains the dual *phoA–lacZα* reporter gene and a kanamycin-resistance gene, flanked by 19 bp inverted repeats. When fused to an ORF, the PhoA–LacZα reporter will be preceded by 13 amino acids (LSLIHISWPMGPG) in front of $Pro_6$ of the mature PhoA protein.

*In vitro* transposition into pBAD derivatives containing *hrc* genes was achieved using EZ::TN transposase (Epicentre Technologies). Equimolar amounts of the transposon donor plasmid (pMA814) and the target plasmid (pBAD24 derivative with the respective *hrc* gene) were mixed with 1 μl transposase in EZ::TN reaction buffer (final volume 10 μl). After 2 h at 37 °C, the reaction was stopped and the mixture transformed into *E. coli* TOP10 cells. Transformants were plated on dual indicator plates and grown for 24 h at 37 °C.

**Mapping of Tn*pholac*1 insertions.** Transposon insertion sites were mapped by colony PCR (Supplementary Table S1). Based on fragment size, plasmid DNA was isolated from representative clones using the QIAprep Spin Miniprep kit (Qiagen). The transposon insertion sites were sequenced using the BigDye terminator sequencing kit on an ABI Prism 310 Genetic Analyzer (Applied Biosystems).

**Construction of site-specific reporter fusions.** To obtain site-specific reporter fusions, plasmid pDT was constructed. The *Pst*I Tn*pholac*1 fragment from pMA814 was cloned downstream of the $P_{BAD}$ promoter of pBAD24, and the resulting plasmid was digested with *Not*I and religated, thus removing the kanamycin-resistance gene and yielding plasmid pDT. The *hrc* gene fragments of interest were amplified by PCR using gene-specific primers (Supplementary Table S1) and cloned in-frame in front of the *phoA–lacZα* reporter using *Nhe*I and *Xba*I. Expression resulted in chimeric proteins consisting of the Hrc fragment, a 20 amino acid linker (SRVDLQPLSLIHISWPMGPG), and the PhoA–LacZα reporter starting at position $Pro_6$ of the mature PhoA protein.

**PhoA reporter assays in *Xanthomonas*.** To perform reporter assays in *X. campestris* pv. vesicatoria, *hrcR* reporter fusions were subcloned into a broad-host-range plasmid derived from pBBR1-MCS5 (Kovach *et al.*, 1995). *Bam*HI–*Sac*I fragments from pBAD reporter constructs were introduced into the pBBR1-MCS5 derivative downstream of a modified *hrpX* gene (A. Krüger and R. Koebnik, unpublished results), thus forming an operon fusion. This plasmid led to expression of a double-tagged HrpX protein from *X. campestris* pv. vesicatoria strain 85-10, with an N-terminal hexahistidine tag and a C-terminal *Strep*-tag II (IBA BioTAGnology). Upon conjugation into *X. campestris* pv. vesicatoria strain 85E, clones were selected on kanamycin/gentamicin plates and subsequently streaked on indicator plates. After 5 days of cultivation, bacteria were resuspended in 5 ml sterile water and assayed for coloration.

**Bioinformatics predictions.** Common algorithms were used for membrane topology predictions, using default parameters: TopPred II (http://bioweb.pasteur.fr/seqanal/interfaces/toppred.html) (von Heijne, 1992), TMHMM version 2.0 (http:/www.cbs.dtu.dk/services/TMHMM-2.0/) (Krogh *et al.*, 2001), HMMTOP version 2.0 (http:/www.enzim.hu/hmmtop/) (Tusnady & Simon, 2001), MEMSAT version 3 (http://bioinf.cs.ucl.ac.uk/psipred/) (Jones, 2007), PHDhtm (http://www.predictprotein.org/) (Rost *et al.*, 2004) and Phobius (http://phobius.cgb.ki.se/) (Käll *et al.*, 2004).

## RESULTS AND DISCUSSION

### Bioinformatics prediction of the membrane topology of six Hrc proteins

Several algorithms are available to predict the topology of membrane proteins (Elofsson & von Heijne, 2007). Since combined predictions are more reliable than single predictions (Ikeda *et al.*, 2002; Nilsson *et al.*, 2000), we applied six internet-based algorithms (TopPred, TMHMM, HMMTOP, MEMSAT, PHDhtm and Phobius) (Fig. 1) to predict the membrane topology of the predicted inner membrane proteins HrcD, HrcR, HrcS, HrcT, HrcU and HrcV from the *X. campestris* pv. vesicatoria T3S system. In no case did the six algorithms predict the same number of transmembrane segments (TMSs). Similarly, for most extramembraneous regions, different locations, i.e. in the cytoplasm or periplasm, were predicted. Also, the hidden Markov model-based algorithms TMHMM, HMMTOP and Phobius differed in their predictions. TMHMM was less sensitive: it did not predict any TMS in HrcD, missed TMS I of HrcU and the predicted TMS II of HrcT, and failed to predict two TMSs instead of only one at about amino acid residues 280–320 of HcrV. Surprisingly, TMHMM and HMMTOP often predicted opposite orientations of the protein within the membrane (HrcD, HrcS, most of HrcT and HrcU, and the C-terminal soluble domain of HrcV). Phobius, which has been designed to differentiate between TMSs and signal sequences of the general secretion pathway, predicted a signal sequence instead of a TMS at the N-terminal regions of HrcD and HrcR. Hence, from predictions, it remained unclear whether or not the N-terminal region is processed by the leader peptidase. PHDhtm, a two-layer neural network, tended to predict shorter TMSs than the other algorithms,

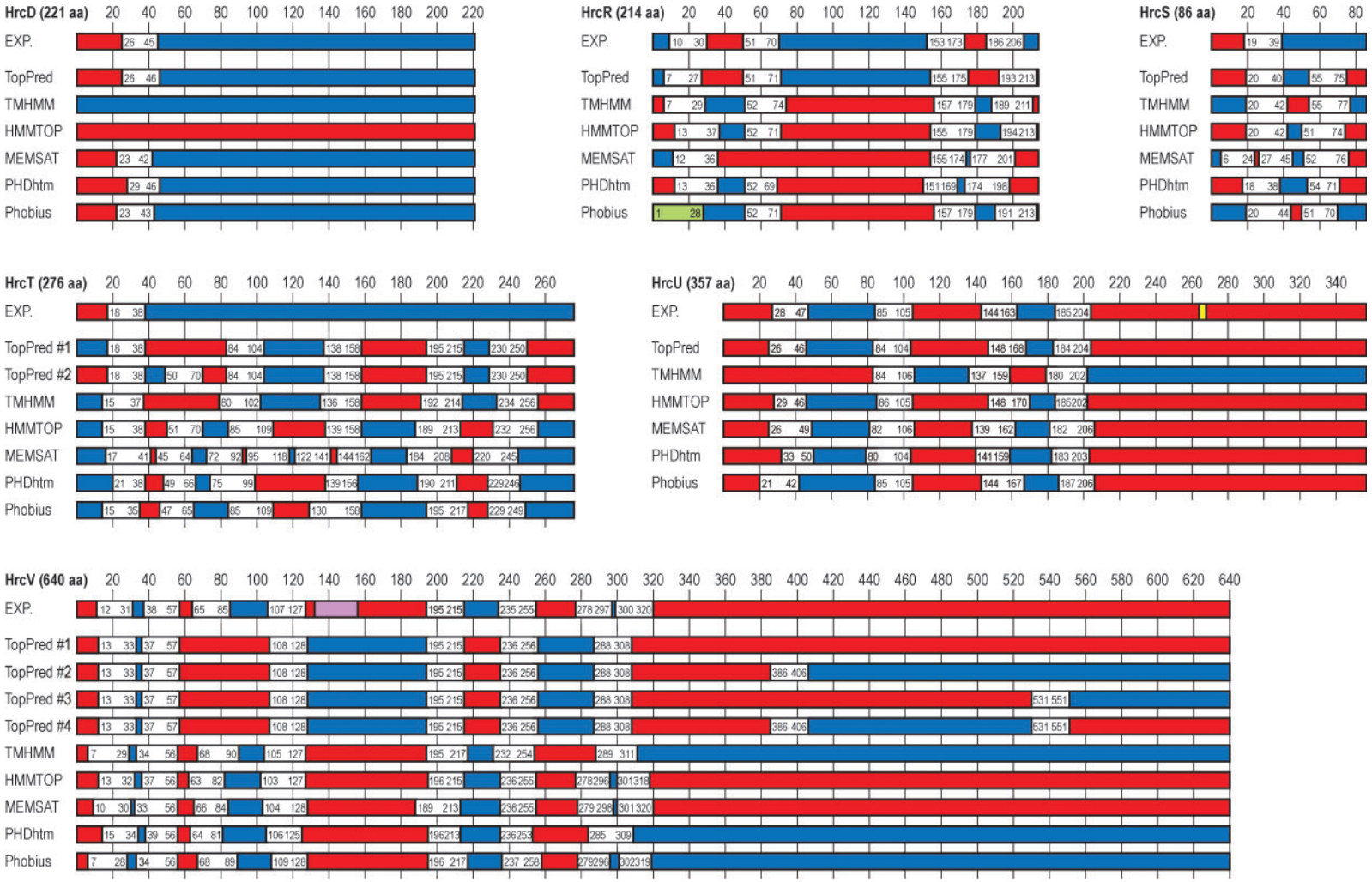


**Fig. 1.** Bioinformatics predictions and experimentally derived membrane topology of conserved inner membrane proteins of the *X. campestris* pv. vesicatoria T3S system. Six algorithms were applied to predict the membrane topology of HrcD, HrcR, HrcS, HrcT, HrcU and HrcV. For HrcT and HrcV, TopPred predicted two and four different models, respectively. For comparison, the experimentally derived topology models (EXP.) are shown at the top. Transmembrane segments are shown in white, periplasmic regions in blue and cytoplasmic regions in red. Predicted signal sequences are indicated in green. The cytoplasmic NPTH motif of HrcU, corresponding to the strictly conserved cleavage site within the YscU/FlhB protein family, is indicated in yellow. The cytoplasmic FHIPEP motif of HrcV is indicated in purple.

probably due to the second layer which aims at correcting for too-long TMS lengths. MEMSAT, which uses multiple sequence alignments to recognize topology models by expectation maximization, tended to predict two short TMSs instead of one long TMS for extended regions of high hydrophobicity (two TMSs for TMS I of HrcS and two extra TMSs for HrcT). In summary, no unambiguous topology model was predicted for any of the six selected Hrc proteins.

## Isolation of reporter fusions to inner membrane Hrc proteins

Because of the difficulties of obtaining reliable topology predictions we studied the membrane topology of six Hrc proteins by reporter fusions, using Tn*pholac*1-based mutagenesis in *E. coli* (Alexeyev & Winkler, 2002). The simultaneous use of two reporter enzymes, alkaline phosphatase (PhoA) and *β*-galactosidase (LacZ), allows enzymic readouts for both periplasmic and cytoplasmic localization at each fusion point.

Coding sequences for the six predicted inner membrane proteins HrcD, HrcR, HrcS, HrcT, HrcU and HrcV from *X. campestris* pv. vesicatoria were cloned under the control of an arabinose-inducible $P_{BAD}$ promoter and mutagenized by Tn*pholac*1 *in vitro*. After selection on plates containing the chromogenic substrates X-Phos and Rose-Gal, four types of colonies were obtained. Blue colonies, indicating PhoA activity, correspond to hybrid proteins with a fusion site in or close to a periplasmic region of the Hrc protein. Red colonies were due to LacZ activity and suggested a fusion in or close to a cytoplasmic domain. White colonies, which were obtained predominantly, were likely due to out-of-frame fusions of the reporter genes with the *hrc* gene, to have an insertion of the reporters in the opposite direction to the *hrc* gene, or to harbour the reporters somewhere else in the mutagenized plasmid. Occasionally, purple colonies were observed. In these cases, the fusion point might be located within a TMS, thus allowing only some but not all reporter moieties to be translocated to the periplasm (Alexeyev & Winkler, 2002). Preliminary studies had shown that out-of-frame fusions tended to display an erratic behaviour, probably due to translation reinitiation in the vicinity of *lacZα*. Therefore, we only analysed colonies with intense coloration.

Generally, many different, randomly distributed fusion points were obtained for the six Hrc proteins, thus allowing us to build precise topology models. In all, we selected nine different translationally in-frame Tn*pholac*1 fusions in *hrcD*, 26 in *hrcR*, nine in *hrcS*, 25 in *hrcT*, 21 in *hrcU* and 32 in *hrcV* (Fig. 2), and labelled them with the number of the amino acid residue after which the fusion occurred. To close gaps in regions where no fusions were obtained, site-specific fusions were constructed, i.e. after positions $S_{50}$, $L_{80}$ and $P_{86}$ in HrcS, after positions $Q_{119}$, $S_{223}$ and $I_{229}$ in HrcT, and after position $S_{134}$ in HrcU (see Methods).

## Model building

For model building, we took advantage of the bioinformatics prediction of candidate TMSs (Fig. 1). We considered all potential TMSs which were predicted by at least one algorithm. However, predictions of pairs of TMSs which corresponded to only one experimentally supported TMS (TMS I and TMS II of HrcS, and TMS III–VI of HrcT, as predicted by MEMSAT) were not considered. When several algorithms predicted a TMS, we defined those amino acid residues as the core of the TMS that were predicted by all algorithms (Supplementary Fig. S1). Since the core sizes varied between five and 20 amino acids, we normalized all TMSs to a reasonable length of 20 or 21 residues necessary to span the lipid bilayer by adding equal numbers of amino acids to both sides of the core. TMSs are counted from the N to the C terminus of the protein by roman numbers: TMS I, TMS II, etc.

The theoretical topology models were then compared with our experimental data which localized the hydrophilic stretches of the polypeptide chain either in the cytoplasm or in the periplasm. Predicted TMSs were not considered when the enzymic activities of fusions at both sides of a predicted TMS did not support their existence. Often, fusions with LacZ or PhoA activity were found too close to each other to be separated by a complete TMS, suggesting that the fusion sites were located within the TMS. This finding was not surprising, since it has been demonstrated before that an incomplete TMS in $N_{in}$–$C_{out}$ orientation can be sufficient to transport a downstream protein segment across the membrane (Boyd *et al.*, 1987; Calamia & Manoil, 1990; Pourcher *et al.*, 1996; Ujwal *et al.*, 1995). This is probably especially true in our experimental system, in which the reporter fusion adds a short stretch of amino acids with elevated hydrophobicity (LSLIHI) due to the inverted repeat of the mini transposon (see Methods). Hence, fusions within the N-terminal third of an $N_{in}$–$C_{out}$ TMS are expected to have LacZ activity, while fusions in the C-terminal two-thirds can lead to PhoA activity. Analogously, fusions within or at the C-terminal end of an $N_{out}$–$C_{in}$ TMS often have PhoA activity. This is probably due to missing downstream topogenic information, which holds the reporter moiety back in the cytoplasm (van Geest & Lolkema, 2000). In these cases, we therefore scrutinized the sequence context for the presence of topogenic signals, i.e. positively charged amino acids (von Heijne, 1986).

## HrcD contains a single N-terminal transmembrane helix followed by a periplasmic protein domain

We first studied the 221 amino acid HrcD protein (formerly HrpD5) (Fig. 2 and Supplementary Fig. S1), which originally was predicted to have 312 amino acids when we started the project (Huguet *et al.*, 1998). Later we found that the formerly predicted start codon was located upstream of the transcriptional start site, as determined by rapid amplification of 5′ complementary DNA ends (5′-RACE) experiments (Weber *et al.*, 2007). As a con-

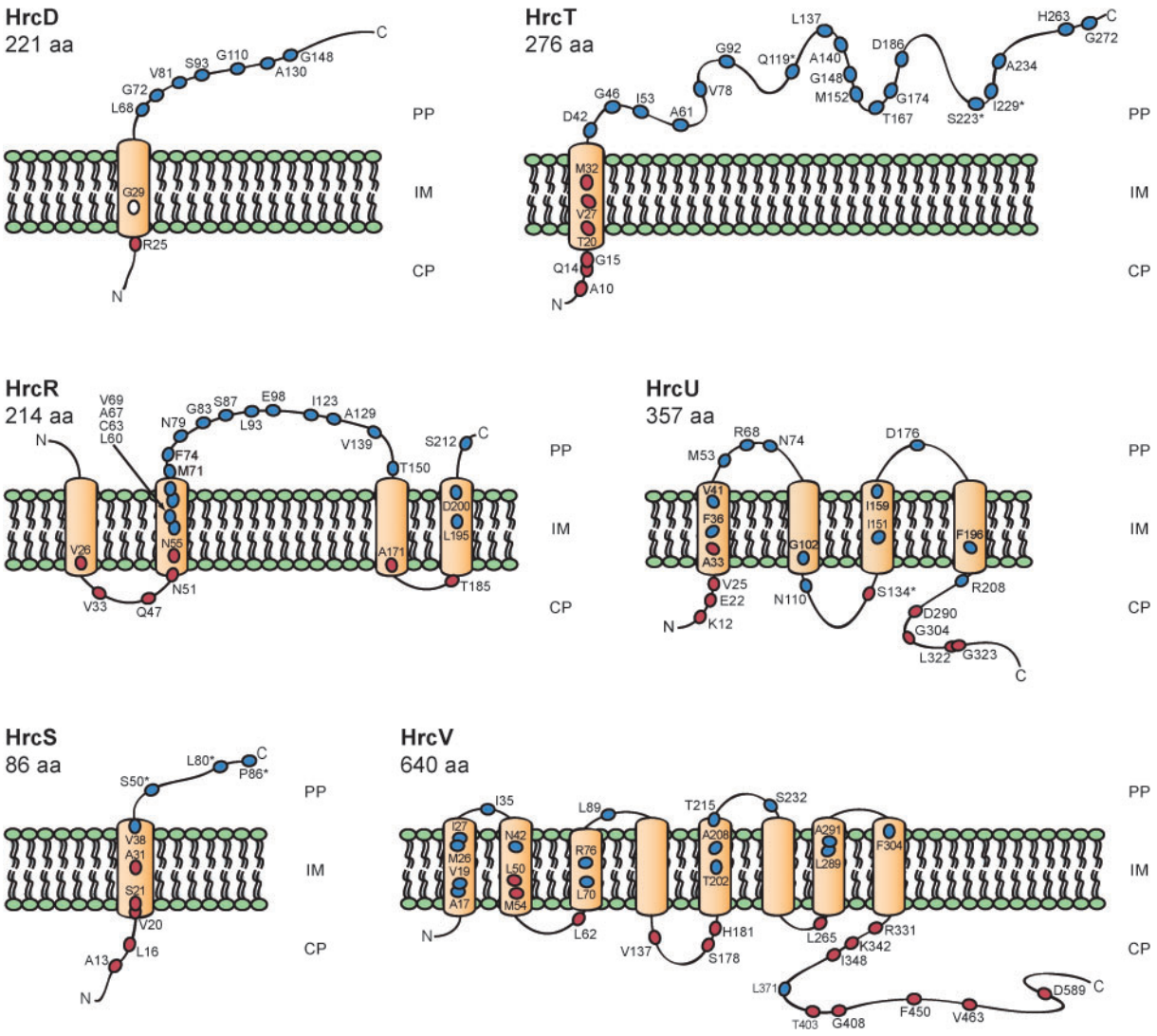


**Fig. 2.** Models of the membrane topology of HrcD, HrcR, HrcS, HrcT, HrcU and HrcV. Phenotypes of *phoA–lacZ*α reporter fusions are indicated by blue (PhoA activity) and red (LacZ activity) ellipses, along with the amino acid position at which the reporters were fused. For HrcD, a purple colony was observed, indicating the activity of both enzymes (white ellipse; see text). PP, periplasm; IM, inner membrane; CP, cytoplasm. Orange cylinders represent transmembrane segments. Asterisks indicate fusions in HrcS ($S_{50}$, $L_{80}$, $L_{86}$), HrcT ($Q_{119}$, $S_{223}$, $I_{229}$) and HrcU ($S_{134}$) which were obtained by site-directed mutagenesis.

sequence, a new translational start codon was defined for *hrcD* that is located 91 codons downstream of the previously predicted start site. The new start codon is preceded by a possible Shine–Dalgarno sequence and would result in a 24 kDa protein. Support for this prediction comes from the expression of a C-terminally hexahistidine-tagged HrcD in *X. campestris* pv. vesicatoria, which allows the purification of a 25 kDa protein (Weber *et al.*, 2007). However, probably due to N-terminal modification of the protein, the N-terminal amino acid sequence could not be determined (C. Berger & R. Koebnik, unpublished results). Based on the first prediction, we have used the 312 codon sequence of *hrcD* for plasmid construction and transposon mutagenesis. Since in this case the first 91 codons do not contain any stretch of hydrophobic amino acids sufficient to span the lipid bilayer, all conclusions from the longer sequence are also valid for the shorter one. Indeed, we had obtained six fusions within the first 91 codons which all led to red colonies, indicating the absence of a putative TMS in this region (data not shown). Therefore, we will refer to the refined numbering throughout this work.

We selected nine different in-frame *hrcD*::Tn*phoA*lac1 mutants, one red, one purple and seven blue colonies (Fig. 2). The red colony had a fusion after position $R_{25}$, indicating that the N terminus of HrcD is located in the cytoplasm. For the purple colony, the fusion site was found

after position $G_{29}$, which overlaps with a hydrophobic segment centred at position $A_{35}$ (Supplementary Fig. S1), indicative of a location within a TMS. All fusions downstream of amino acid $L_{68}$ resulted in blue colonies, suggesting the absence of an additional TMS in this region and a periplasmic location of the C-terminal protein domain. Our results are supported by four predictions, TopPred, MEMSAT, PHDhtm and Phobius (Fig. 1). Only two hidden Markov model-based algorithms, TMHMM and HMMTOP, predicted a soluble protein without any TMS in the 221 amino acid protein.

HrcD is similar (30 % sequence identity) to HrpW of *R. solanacearum* (323 amino acids; accession no. CAD18008), and weakly related to HrpQ of *P. syringae* (330 amino acids; AAO54923), YscD of *Yersinia* (419 amino acids; AAC62551), PrgH of *Salmonella* (392 amino acids; AAB60188) and MxiG of *Shigella* (371 amino acids; CAC05811) (Ghosh, 2004; Huguet *et al.*, 1998; Pühler *et al.*, 2004; Van Gijsegem *et al.*, 2002). Since PrgH is a component of the needle complex, the cell-surface appendage of the T3S system of animal pathogens (Galan & Wolf-Watz, 2006; Kubori *et al.*, 1998), HrcD might play a role in the assembly of the Hrp pilus, the appendage of plant pathogens. Such a role is supported by the co-transcription of *hrcD* with two other genes devoted to Hrp pilus assembly, *hrpD6* and *hrpE*, the latter encoding the major pilus subunit (Weber *et al.*, 2005, 2007). Some members of the YscD/HrpQ protein family contain a (cytoplasmic) FHA (forkhead-associated) domain at the N terminus. The FHA domain is a phosphopeptide recognition domain found in many regulatory proteins, e.g. kinases and phosphatases (Durocher & Jackson, 2002). HrcD is significantly shorter at the N terminus and lacks the FHA domain. This finding is not surprising because the FHA domains of YscD-like proteins might just represent a molecular fossil which is undergoing evolutionary erosion (Pallen *et al.*, 2005).

## HrcR is a four-transmembrane-helix protein with a central periplasmic domain

For *hrcR*, we selected seven red and 18 blue Tn*pholac*1 mutants (Fig. 2 and Supplementary Fig. S1). All fusions upstream of position $G_{56}$ resulted in red colonies, suggesting a cytoplasmic location of this region. Since all six algorithms predict a hydrophobic, membrane-interacting segment within the first 37 amino acid residues (Fig. 1), we are confident about the presence of an $N_{out}$–$C_{in}$ TMS (TMS I) in this region, for which, however, experimental proof cannot be obtained by our approach. The Phobius algorithm suggested that the candidate TMS I is a signal sequence which is removed by the leader peptidase. If this is the case, one would expect that fusions immediately downstream of the processing site would lead to PhoA activity. However, fusions after $V_{26}$ and $V_{33}$ resulted in red colonies. Moreover, bioinformatics analyses (Phobius) of 20 randomly chosen HrcR homologues with less than 85 % pairwise identity, including YscR from *Yersinia pestis*, did not predict a classical signal sequence for any of them (data not shown). Thus, the N-terminal hydrophobic segment likely corresponds to a TMS and not to a signal sequence.

The most N-terminal fusion leading to a blue colony was obtained after position $L_{60}$, in agreement with a model placing this fusion site in TMS II ($N_{in}$–$C_{out}$). A periplasmic loop of approximately 80 amino acid residues between TMS II and TMS III, from $M_{71}$ to $A_{152}$, is supported by more than 10 insertions resulting in blue colonies. Two fusions at $A_{171}$ and $T_{185}$, leading to red colonies, support the existence of TMS III in the C-terminal half of the protein, and three fusions with a blue phenotype after positions $L_{195}$, $D_{200}$ and $S_{212}$ clearly localize the C-terminal region in the periplasmic space. Hence, the polypeptide chain of HrcR traverses the inner membrane by four transmembrane helices.

Five of six algorithms predicted an opposite orientation of the protein in the membrane and only TopPred supports our model of HrcR. This situation is reminiscent of topology studies with the HrcR homologue from *Y. pestis*, YscR. Based on six Tn*phoA* fusions, of which only one fusion at amino acid residue 45 gave rise to high PhoA activity, Fields *et al.* (1994) proposed four TMSs for YscR, with both protein termini residing in the cytoplasm. Thus, there are two alternative models for members of the YscR/HrcR protein family with opposite orientation in the membrane, a situation that is not without precedent (von Heijne, 2006). Still, this finding is surprising, since HrcR and YscR share 50 % sequence identity. Because of the high density of transposon insertions in *hrcR* and the double enzymic readout for all the fusion proteins we are confident that our HrcR model is correct. Our model is also supported by two PhoA fusions in the central loop of HrcR (formerly HrpT) from *R. solanacearum* (Van Gijsegem *et al.*, 1995), corresponding to positions $S_{84}$ and $V_{90}$ of HrcR from *X. campestris* pv. vesicatoria.

To exclude any artefacts that might originate from the use of *E. coli*, we transferred six representative HrcR fusions into *X. campestris* pv. vesicatoria. Reporter fusions were cloned into a broad-host-range plasmid downstream of a constitutively expressed *hrpX* derivative. This plasmid complemented an *hrpX* mutant and activated the expression of *hrpX*-dependent *hrp* operons (A. Krüger, G. P. Robin and R. Koebnik, unpublished results). On indicator plates for PhoA activity, we observed bluish colonies with fusions after positions $N_{79}$, $E_{98}$ and $T_{150}$ (Supplementary Fig. S2). In contrast, fusions after positions $N_{51}$, $A_{171}$ and $T_{185}$ led to the same greenish colonies as the strain with the 'empty' *hrpX* plasmid, most likely due to low background activity of the endogenous *phoA* gene and the yellow pigment xanthomonadin (Supplementary Fig. S2). Since xanthomonads expressing the LacZ Ω fragment in a strain without endogenous β-galactosidase are not available, we could not determine the LacZ activity of the reporter fusions. In summary, we observed the same topology in *E.*

*coli* and in *X. campestris* pv. vesicatoria regardless of the presence or absence of other Hrc proteins, including the HrcJ/HrcD ring as a molecular platform for T3S apparatus assembly (Yip *et al.*, 2005; Spreter *et al.*, 2009).

## HrcS has at least one transmembrane helix and a periplasmic C terminus

HrcS, with only 86 amino acids, is the smallest inner membrane Hrc protein. The topology model for HrcS derived from nine transposon insertions shows only one TMS (Fig. 2 and Supplementary Fig. S1). The insertions $A_{13}$, $L_{16}$, $V_{20}$ and $S_{21}$ had LacZ activity, thus being compatible with a cytoplasmic location of the N terminus. Although insertions $A_{31}$ and $V_{38}$ lie close together, they led to LacZ and PhoA activity, respectively. The two amino acid residues are separated by a hydrophobic stretch (GLLIAF), and are thus probably located in the N- and C-terminal halves of a TMS. Three additional fusions after positions $S_{50}$, $L_{80}$ and $P_{86}$ in HrcS were constructed in pDT and resulted in high PhoA activity. The same phenotype has been observed with PhoA fusions after position $S_{50}$ of HrcS (formerly HrpU) of *R. solanacearum* (Van Gijsegem *et al.*, 1995), corresponding to position $S_{50}$ of HrcS from *X. campestris* pv. vesicatoria. In contrast to the experimental data, all algorithms predicted two or three TMSs with different orientations in the membrane (Fig. 1). Only the N-terminally located predicted TMS fits well to our data because experimental support for the existence of a second (or third) TMS in the C-terminal part of the protein is missing. However, we cannot entirely rule out the possibility that the N-terminal TMS (in $N_{in}$–$C_{out}$ orientation) exports the reporter in the $V_{38}$ and $S_{50}$ fusions, leading to PhoA activity. But when more amino acid residues of HrcS are present in the fusion protein, the N-terminal TMS might adopt the opposite orientation ($N_{out}$–$C_{in}$) and a second TMS ($N_{in}$–$C_{out}$) might be formed. This scenario would be compatible with the predictions of TMHMM and Phobius. In any case, the C terminus of HrcS is located on the periplasmic side, as suggested by the two fusions after positions $L_{80}$ and $P_{86}$.

## HrcT contains an N-terminal transmembrane anchor followed by a large periplasmic domain

Our experimental data suggest that the 276 amino acid HrcT protein is an integral membrane protein with only one TMS close to the N terminus which serves as a membrane anchor (Fig. 2 and Supplementary Fig. S1). This conclusion is based on 25 PhoA–LacZ$\alpha$ fusions. All six fusions before amino acid $F_{33}$ resulted in high LacZ activity, while all 19 fusions behind amino acid $D_{42}$ resulted in high PhoA activity. Because we did not isolate any Tn*phoA*lac1-based fusion between positions $G_{92}$ and $L_{137}$ and between positions $D_{186}$ and $A_{234}$, we constructed three additional site-specific fusions after positions $Q_{119}$, $S_{223}$ and $I_{229}$. Since these constructs also resulted in high PhoA activity we conclude that HrcT contains a large periplasmic domain, starting at approximately amino acid $T_{39}$ (Supplementary Fig. S1).

Surprisingly, this model is in sharp contrast to bioinformatics predictions, which consistently predicted five to eight TMSs for HrcT (Fig. 1). The N terminus is predicted, except for one of the two TopPred models, to be in the periplasm. Since we considered all the rules of van Geest and Lolkema, i.e. a minimum of one fusion in each predicted extra-membraneous domain and at least one fusion for every 30 amino acid residues in regions of intermediate hydrophobicity (van Geest & Lolkema, 2000), and obtained a positive readout (i.e. high PhoA activity) for all fusions downstream of position $D_{42}$, we believe that HrcT contains a large C-terminal periplasmic domain.

## HrcU has an N-terminal membrane-embedded domain and a C-terminal cytoplasmic domain

For the 357 amino acid HrcU protein, we isolated 21 different PhoA–LacZ$\alpha$ fusions (Fig. 2 and Supplementary Fig. S1). Fusions with PhoA activity were restricted to the N-terminal 210 amino acid residues, indicative of a membrane-embedded domain in the N-terminal protein region. Our data are in agreement with four TMSs in this region, as predicted by all algorithms except TMHMM, which did not predict the first TMS. The insertions $A_{33}$ and $F_{36}$, resulting in red and blue colonies, respectively, are probably both located in the first TMS. Insertions $N_{110}$ and $R_{208}$, although leading to PhoA activity, are probably located on the cytoplasmic side of the membrane close to the C-terminal end of an $N_{out}$–$C_{in}$ TMS (Fig. 1). Scrutiny of the amino acid sequence context revealed that both residues are followed by two or three basic amino acids within the next 10 residues, which might have a topogenic function (Supplementary Fig. S1) (von Heijne, 1986). From this we conclude that missing downstream topogenic information is responsible for this atypical, but not unexpected, behaviour (van Geest & Lolkema, 2000). Because Tn*phoA*lac1 mutagenesis did not lead to fusions supportive of TMS II and TMS III, we constructed a site-specific fusion after position $S_{134}$. As expected, this fusion resulted in high LacZ activity, thus confirming the four-TMS model for the N-terminal 205 amino acid residues.

All four reporter fusions downstream of amino acid $R_{208}$ gave rise to high PhoA activity, indicative of a large cytoplasmic protein domain (Fig. 2 and Supplementary Fig. S1). Our model of HrcU is corroborated by a model of YscU from *Yersinia enterocolitica* which is based on 10 PhoA fusions (Allaoui *et al.*, 1994). Thus, both proteins share a large C-terminal cytoplasmic domain with a conserved proteolytic cleavage site which plays a role in substrate specificity switching (Lavander *et al.*, 2002; Minamino & Macnab, 2000a). HrcU cleavage has been observed by immunoblot analysis, suggesting a similar but distinct mechanism of substrate specificity control (C. Berger & R. Koebnik, unpublished results; Lorenz *et al.*, 2008b).

## HrcV is a two-domain membrane protein with a large cytoplasmic domain in the C-terminal region

For *hrcV*, we selected 15 red and 17 blue Tn*pholac*1 mutants (Fig. 2 and Supplementary Fig. S1). Within the N-terminal half of the 640 amino acid HrcV protein we observed clusters of reporter fusions with either LacZ or PhoA activity along the polypeptide chain, indicative of a large membrane-embedded domain. These data are in agreement with the presence of eight TMSs (TMS I to TMS VIII), as predicted by most algorithms (Fig. 1 and Supplementary Fig. S1). According to the positive-inside rule and the presence of two positively charged amino acids ($R_2$ and $R_5$), the N terminus was expected to reside in the cytoplasm. This was also predicted by all six algorithms (Fig. 1). Not unexpectedly, several fusions within the N-terminal half of an $N_{in}$–$C_{out}$ TMS (i.e. after positions $A_{17}$, $V_{19}$, $L_{70}$ and $T_{202}$) showed PhoA activity. Our model places the diagnostic FHIPEP motif (PROSITE accession no. PS00994) into the second cytoplasmic loop between TMS IV and TMS V, as proposed for FlhA (Fig. 1) (McMurry *et al.*, 2004).

Except for one fusion ($L_{371}$, see below), all fusions downstream of amino acid $R_{331}$ in HrcV led to LacZ activity, suggesting that HrcV possesses a large C-terminal cytoplasmic domain of approximately 320 amino acids. The localization of this domain varied in the predictions (Fig. 1). A C-terminal TMS at amino acid positions 531–551, which was predicted only by TopPred, is unlikely because the downstream reporter fusion after position $D_{589}$ led to LacZ activity. Moreover, the corresponding region of the flagellar homologue, FlhA, has been purified and crystallized without detergent, supporting a soluble protein domain (Saijo-Hamano *et al.*, 2004, 2005). A cytoplasmic location of the C-terminal domain of the YscV/FlhA protein family is also supported by the finding that the C-terminal domain of FlhA interacts with the cytoplasmic proteins FliJ, FliH and FliI (Minamino & Macnab, 2000b). Moreover, FlhA has been shown to interact with the C-terminal domain of FlhB, the HrcU homologue (Zhu *et al.*, 2002). Similar findings have been obtained in our laboratory using a bacterial two-hybrid system (Dmitrova *et al.*, 1998), showing that the soluble domain of HrcV ($HrcV_{322\text{-}645}$) interacts with itself and with the soluble domain of HrcU ($HrcU_{265\text{-}357}$) (C. Berger & R. Koebnik, unpublished results). Thus, the large C-terminal domain of HrcV is localized in the cytoplasm.

The reporter fusion after position $L_{371}$, which led to PhoA activity, is of special interest. To include this particular fusion into the model, an upstream $N_{in}$–$C_{out}$ TMS (TMS IX) would have to be postulated. We believe that the fairly hydrophobic sequence $L_{359}$SMRLSPQLAALL$_{371}$, which is just upstream of the $L_{371}$ fusion point, has the potential to form a TMS when fused to a few additional hydrophobic amino acids (LSLIHI) derived from the PhoA–LacZα reporter (Supplementary Fig. S1). Hence, an artificial TMS might be formed upon fusion after $L_{371}$, leading to the export of the reporter, and the presumed TMS IX is likely absent in the wild-type protein. In support of this, a comparison of HrcV homologues from 14 different genera shows that a fairly hydrophobic sequence is not conserved in this region of the polypeptide chain (Supplementary Fig. S3). Thus, the fusion after position $L_{371}$ most likely led to an experimental artefact and illustrates the importance of careful analysis of all the fusion constructs.

The HrcV homologue from *Y. pestis*, YscV (formerly LcrD), has been studied by Tn*phoA* mutagenesis, leading to a model of YscV with eight TMSs and a large cytoplasmic C-terminal domain (Plano *et al.*, 1991). The last TMS of YscV, TMS VIII, is predicted between amino acids 330 and 347; however, this region is fairly hydrophilic and not likely to form a TMS. We therefore suggest that TMS VIII resides between positions $F_{301}$ and $L_{320}$ in YscV, which corresponds to positions $F_{299}$ and $I_{318}$ in HrcV (TMS VIII). This model would explain the PhoA activity of YscV fusions $L_{297}$ and $V_{304}$ (Plano *et al.*, 1991). The PhoA activity of two fusions at the C-terminal end of TMS VIII ($G_{317}$ and $L_{320}$ of YscV) could be explained by the lack of downstream topogenic information, i.e. the conserved Arg-Lys motif at amino acid residues 322 and 323 (van Geest & Lolkema, 2000).

## Validity of the PhoA–LacZ reporter fusion approach to study the membrane topology of inner membrane proteins

Our experimental approach relied on the assumption that inner membrane proteins can be studied individually upon heterologous expression in *E. coli* although they are normally part of a multiprotein complex, i.e. the T3S apparatus. The validity of the gene fusion approach for a heterologous protein of a multiprotein complex has been demonstrated for the L-subunit of the photosynthetic reaction centre from *Rhodobacter sphaeroides* (Yun *et al.*, 1991). Gene fusions with high PhoA activity were shown to be in regions of the polypeptide known to be at or near the periplasmic surface, as defined by the high-resolution X-ray structure. Later, several proteins of heterologous multiprotein complexes have been studied by reporter fusions expressed in *E. coli*, e.g. the pro-sigma(K) processing complex of *Bacillus subtilis* and components of the type IV secretion system of *Helicobacter pylori* (Green & Cutting, 2000; Hofreuter *et al.*, 2003). As an example of special interest, we studied the membrane topology of HrcR in *E. coli* and in *X. campestris* pv. vesicatoria. All six reporter fusions adopted the same topology in both species, thus validating the approach of heterologous expression in *E. coli*.

The reporter fusion approach also relies on the assumption that C-terminal truncations of the protein under study do not affect its native topology, i.e. the formation and orientation of TMSs. This assumption appears to be justified because dozens of membrane proteins have been studied this way and many topology models were later confirmed by other biochemical approaches or high-resolution structures (van Geest & Lolkema, 2000). However, a few studies indicate

that fusion proteins can occasionally display an anomalous behaviour with respect to membrane topology (for a detailed discussion, see van Geest & Lolkema, 2000). For instance, it is possible that more C-terminal transmembrane helices can influence the topology of more N-terminal transmembrane helices (Ota *et al.*, 1998; van Geest & Lolkema, 1996). Such pitfalls guided the formulation of rules to optimize the reporter approach, as summarized by van Geest & Lolkema (2000). It has been suggested that experimental support for a topological model requires a minimum of one fusion in each extra-membraneous domain. If the protein contains amino acid stretches of intermediate hydrophobicity that cannot unambiguously be predicted as membrane spanning, fusions should be made approximately every 30 residues. The use of PhoA as a reporter protein is advantageous because a positive result (i.e. enzymic activity) requires the export of the mature reporter enzyme moiety into the periplasm, while the use of LacZ as a cytoplasmic reporter may also give an enzymic activity as a result of artefacts, such as jamming of the export machinery. Throughout our study, we followed the rules of van Geest & Lolkema (2000). We are thus confident that our topology models are correct.

### Concluding remarks

Electron microscopy experiments have unravelled the morphology of some nonflagellar T3S systems (Blocker *et al.*, 2001; Kubori *et al.*, 1998; Marlovits *et al.*, 2004), but a detailed understanding of the structural characteristics, organization and precise molecular organization of T3S systems is still lacking. Here, we present what is believed to be the first complete model of the inner membrane topology of a T3S system, which is based on a bioinformatics consensus prediction and experimentally derived constraints, including 129 different enzymically active reporter fusions. Our models validate and revise older models of T3S membrane proteins and will be instrumental for further model building, using X-rays, NMR and high-resolution microscopy.

## ACKNOWLEDGEMENTS

We thank Hannelore Espenhahn for excellent technical assistance, Antje Krüger (Martin-Luther-University, Halle) for providing the *hrpX* pBBR1-MCS5 derivative, Philipp Franken for comments on the manuscript, and Ernst Weber for fruitful discussions. We are very grateful to Mikhail F. Alexeyev and Herbert H. Winkler (University of South Alabama) for the Tn*pholac* mutagenesis system. This work was funded in part by grant KO 1686/3-2 from the Deutsche Forschungsgemeinschaft to R. K.

## REFERENCES

**Aizawa, S. I. (2001).** Bacterial flagella and type III secretion systems. *FEMS Microbiol Lett* **202**, 157–164.

**Alexeyev, M. F. & Winkler, H. H. (2002).** Transposable dual reporters for studying the structure–function relationships in membrane proteins: permissive sites in *R. prowazekii* ATP/ADP translocase. *Biochemistry* **41**, 406–414.

**Allaoui, A., Woestyn, S., Sluiters, C. & Cornelis, G. R. (1994).** YscU, a *Yersinia enterocolitica* inner membrane protein involved in Yop secretion. *J Bacteriol* **176**, 4534–4542.

**Bertani, G. (1951).** Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* **62**, 293–300.

**Blocker, A., Jouihri, N., Larquet, E., Gounon, P., Ebel, F., Parsot, C., Sansonetti, P. & Allaoui, A. (2001).** Structure and composition of the *Shigella flexneri* 'needle complex', a part of its type III secreton. *Mol Microbiol* **39**, 652–663.

**Bogdanove, A. J., Beer, S. V., Bonas, U., Boucher, C. A., Collmer, A., Coplin, D. L., Cornelis, G. R., Huang, H. C., Hutcheson, S. W. & other authors (1996).** Unified nomenclature for broadly conserved *hrp* genes of phytopathogenic bacteria. *Mol Microbiol* **20**, 681–683.

**Bonas, U., Stall, R. E. & Staskawicz, B. (1989).** Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet* **218**, 127–136.

**Bonas, U., Schulte, R., Fenselau, S., Minsavage, G. V., Staskawicz, B. J. & Stall, R. E. (1991).** Isolation of a gene cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. *Mol Plant Microbe Interact* **4**, 81–88.

**Boyd, D., Manoil, C. & Beckwith, J. (1987).** Determinants of membrane protein topology. *Proc Natl Acad Sci U S A* **84**, 8525–8529.

**Büttner, D. & Bonas, U. (2006).** Who comes first? How plant pathogenic bacteria orchestrate type III secretion. *Curr Opin Microbiol* **9**, 193–200.

**Büttner, D., Nennstiel, D., Klüsener, B. & Bonas, U. (2002).** Functional analysis of HrpF, a putative type III translocon protein from *Xanthomonas campestris* pv. vesicatoria. *J Bacteriol* **184**, 2389–2398.

**Calamia, J. & Manoil, C. (1990).** *lac* permease of *Escherichia coli*: topology and sequence elements promoting membrane insertion. *Proc Natl Acad Sci U S A* **87**, 4937–4941.

**Coburn, B., Sekirov, I. & Finlay, B. B. (2007).** Type III secretion systems and disease. *Clin Microbiol Rev* **20**, 535–549.

**Cornelis, G. R. (2006).** The type III secretion injectisome. *Nat Rev Microbiol* **4**, 811–825.

**Creasey, E. A., Delahay, R. M., Daniell, S. J. & Frankel, G. (2003).** Yeast two-hybrid system survey of interactions between LEE-encoded proteins of enteropathogenic *Escherichia coli*. *Microbiology* **149**, 2093–2106.

**Deane, J. E., Graham, S. C., Mitchell, E. P., Flot, D., Johnson, S. & Lea, S. M. (2008).** Crystal structure of Spa40, the specificity switch for the *Shigella flexneri* type III secretion system. *Mol Microbiol* **69**, 267–276.

**Dmitrova, M., Younes-Cauet, G., Oertel-Buchheit, P., Porte, D., Schnarr, M. & Granger-Schnarr, M. (1998).** A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in *Escherichia coli*. *Mol Gen Genet* **257**, 205–212.

**Durocher, D. & Jackson, S. P. (2002).** The FHA domain. *FEBS Lett* **513**, 58–66.

**Edqvist, P. J., Olsson, J., Lavander, M., Sundberg, L., Forsberg, A., Wolf-Watz, H. & Lloyd, S. A. (2003).** YscP and YscU regulate substrate specificity of the *Yersinia* type III secretion system. *J Bacteriol* **185**, 2259–2266.

**Elofsson, A. & von Heijne, G. (2007).** Membrane protein structure: prediction versus reality. *Annu Rev Biochem* **76**, 125–140.

**Fadouloglou, V. E., Tampakaki, A. P., Glykos, N. M., Bastaki, M. N., Hadden, J. M., Phillips, S. E., Panopoulos, N. J. & Kokkinidis, M. (2004).** Structure of $HrcQ_B$-C, a conserved component of the bacterial type III secretion systems. *Proc Natl Acad Sci U S A* **101**, 70–75.

**Fan, F., Ohnishi, K., Francis, N. R. & Macnab, R. M. (1997).** The FliP and FliR proteins of *Salmonella typhimurium*, putative components of the type III flagellar export apparatus, are located in the flagellar basal body. *Mol Microbiol* **26**, 1035–1046.

**Ferris, H. U. & Minamino, T. (2006).** Flipping the switch: bringing order to flagellar assembly. *Trends Microbiol* **14**, 519–526.

**Fields, K. A., Plano, G. V. & Straley, S. C. (1994).** A low-$Ca^{2+}$ response (LCR) secretion (*ysc*) locus lies within the *lcrB* region of the LCR plasmid in *Yersinia pestis*. *J Bacteriol* **176**, 569–579.

**Fraser, G. M., Hirano, T., Ferris, H. U., Devgan, L. L., Kihara, M. & Macnab, R. M. (2003).** Substrate specificity of type III flagellar protein export in *Salmonella* is controlled by subdomain interactions in FlhB. *Mol Microbiol* **48**, 1043–1057.

**Galan, J. E. & Wolf-Watz, H. (2006).** Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**, 567–573.

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

**Green, D. H. & Cutting, S. M. (2000).** Membrane topology of the *Bacillus subtilis* pro-$\sigma^K$ processing complex. *J Bacteriol* **182**, 278–285.

**Gürlebeck, D., Thieme, F. & Bonas, U. (2006).** Type III effector proteins from the plant pathogen *Xanthomonas* and their role in the interaction with the host plant. *J Plant Physiol* **163**, 233–255.

**Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. (1995).** Tight regulation, modulation, and high-level expression by vectors containing the arabinose $P_{BAD}$ promoter. *J Bacteriol* **177**, 4121–4130.

**He, S. Y. & Jin, Q. (2003).** The Hrp pilus: learning from flagella. *Curr Opin Microbiol* **6**, 15–19.

**He, S. Y., Nomura, K. & Whittam, T. S. (2004).** Type III protein secretion mechanism in mammalian and plant pathogens. *Biochim Biophys Acta* **1694**, 181–206.

**Hofreuter, D., Karnholz, A. & Haas, R. (2003).** Topology and membrane interaction of *Helicobacter pylori* ComB proteins involved in natural transformation competence. *Int J Med Microbiol* **293**, 153–165.

**Huguet, E., Hahn, K., Wengelnik, K. & Bonas, U. (1998).** *hpaA* mutants of *Xanthomonas campestris* pv. *vesicatoria* are affected in pathogenicity but retain the ability to induce host-specific hypersensitive reaction. *Mol Microbiol* **29**, 1379–1390.

**Ikeda, M., Arai, M., Lao, D. M. & Shimizu, T. (2002).** Transmembrane topology prediction methods: a re-assessment and improvement by a consensus method using a dataset of experimentally-characterized transmembrane topologies. *In Silico Biol* **2**, 19–33.

**Jones, D. T. (2007).** Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics* **23**, 538–544.

**Käll, L., Krogh, A. & Sonnhammer, E. L. (2004).** A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**, 1027–1036.

**Koebnik, R., Krüger, A., Thieme, F., Urban, A. & Bonas, U. (2006).** Specific binding of the *Xanthomonas campestris* pv. vesicatoria AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. *J Bacteriol* **188**, 7652–7660.

**Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., II & Peterson, K. M. (1995).** Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**, 175–176.

**Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. (2001).** Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.

**Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. E. & Aizawa, S. I. (1998).** Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. *Science* **280**, 602–605.

**Lavander, M., Sundberg, L., Edqvist, P. J., Lloyd, S. A., Wolf-Watz, H. & Forsberg, A. (2002).** Proteolytic cleavage of the FlhB homologue YscU of *Yersinia pseudotuberculosis* is essential for bacterial survival but not for type III secretion. *J Bacteriol* **184**, 4500–4509.

**Lorenz, C., Kirchner, O., Egler, M., Stuttmann, J., Bonas, U. & Büttner, D. (2008a).** HpaA from *Xanthomonas* is a regulator of type III secretion. *Mol Microbiol* **69**, 344–360.

**Lorenz, C., Schulz, S., Wolsch, T., Rossier, O., Bonas, U. & Büttner, D. (2008b).** HpaC controls substrate specificity of the *Xanthomonas* type III secretion system. *PLoS Pathog* **4**, e1000094.

**Marlovits, T. C., Kubori, T., Sukhan, A., Thomas, D. R., Galan, J. E. & Unger, V. M. (2004).** Structural insights into the assembly of the type III secretion needle complex. *Science* **306**, 1040–1042.

**McCann, H. C. & Guttman, D. S. (2008).** Evolution of the type III secretion system and its effectors in plant–microbe interactions. *New Phytol* **177**, 33–47.

**McMurry, J. L., Van Arnam, J. S., Kihara, M. & Macnab, R. M. (2004).** Analysis of the cytoplasmic domains of *Salmonella* FlhA and interactions with components of the flagellar export machinery. *J Bacteriol* **186**, 7586–7592.

**Melen, K., Krogh, A. & von Heijne, G. (2003).** Reliability measures for membrane protein topology prediction algorithms. *J Mol Biol* **327**, 735–744.

**Minamino, T. & Macnab, R. M. (2000a).** Domain structure of *Salmonella* FlhB, a flagellar export component responsible for substrate specificity switching. *J Bacteriol* **182**, 4906–4914.

**Minamino, T. & Macnab, R. M. (2000b).** Interactions among components of the *Salmonella* flagellar export apparatus and its substrates. *Mol Microbiol* **35**, 1052–1064.

**Minamino, T., Iino, T. & Kutuskake, K. (1994).** Molecular characterization of the *Salmonella typhimurium flhB* operon and its protein products. *J Bacteriol* **176**, 7630–7637.

**Moraes, T. F., Spreter, T. & Strynadka, N. C. (2008).** Piecing together the type III injectisome of bacterial pathogens. *Curr Opin Struct Biol* **18**, 258–266.

**Nguyen, L., Paulsen, I. T., Tchieu, J., Hueck, C. J. & Saier, M. H., Jr (2000).** Phylogenetic analyses of the constituents of type III protein secretion systems. *J Mol Microbiol Biotechnol* **2**, 125–144.

**Nilsson, J., Persson, B. & von Heijne, G. (2000).** Consensus predictions of membrane protein topology. *FEBS Lett* **486**, 267–269.

**Ohnishi, K., Fan, F., Schoenhals, G. J., Kihara, M. & Macnab, R. M. (1997).** The FliO, FliP, FliQ, and FliR proteins of *Salmonella typhimurium*: putative components for flagellar assembly. *J Bacteriol* **179**, 6092–6099.

**Ota, K., Sakaguchi, M., Hamasaki, N. & Mihara, K. (1998).** Assessment of topogenic functions of anticipated transmembrane segments of human band 3. *J Biol Chem* **273**, 28286–28291.

**Pallen, M. J., Beatson, S. A. & Bailey, C. M. (2005).** Bioinformatics, genomics and evolution of non-flagellar type-III secretion systems: a Darwinian perspective. *FEMS Microbiol Rev* **29**, 201–229.

**Plano, G. V. & Straley, S. C. (1995).** Mutations in *yscC*, *yscD*, and *yscG* prevent high-level expression and secretion of V antigen and Yops in *Yersinia pestis*. *J Bacteriol* **177**, 3843–3854.

**Plano, G. V., Barve, S. S. & Straley, S. C. (1991).** LcrD, a membrane-bound regulator of the *Yersinia pestis* low-calcium response. *J Bacteriol* **173**, 7293–7303.

**Pourcher, T., Bibi, E., Kaback, H. R. & Leblanc, G. (1996).** Membrane topology of the melibiose permease of *Escherichia coli* studied by *melB–phoA* fusion analysis. *Biochemistry* **35**, 4161–4168.

**Preston, G. M. (2007).** Metropolitan microbes: type III secretion in multihost symbionts. *Cell Host Microbe* **2**, 291–294.

**Pühler, A., Arlat, M., Becker, A., Göttfert, M., Morrissey, J. P. & O'Gara, F. (2004).** What can bacterial genome research teach us about bacteria–plant interactions? *Curr Opin Plant Biol* **7**, 137–147.

**Rossier, O., Wengelnik, K., Hahn, K. & Bonas, U. (1999).** The *Xanthomonas* Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. *Proc Natl Acad Sci U S A* **96**, 9368–9373.

**Rossier, O., Van den Ackerveken, G. & Bonas, U. (2000).** HrpB2 and HrpF from *Xanthomonas* are type III-secreted proteins and essential for pathogenicity and recognition by the host plant. *Mol Microbiol* **38**, 828–838.

**Rost, B., Yachdav, G. & Liu, J. (2004).** The PredictProtein server. *Nucleic Acids Res* **32**, W321–W326.

**Saijo-Hamano, Y., Minamino, T., Macnab, R. M. & Namba, K. (2004).** Structural and functional analysis of the C-terminal cytoplasmic domain of FlhA, an integral membrane component of the type III flagellar protein export apparatus in *Salmonella*. *J Mol Biol* **343**, 457–466.

**Saijo-Hamano, Y., Imada, K., Minamino, T., Kihara, M., Macnab, R. M. & Namba, K. (2005).** Crystallization and preliminary X-ray analysis of the C-terminal cytoplasmic domain of FlhA, a membrane-protein subunit of the bacterial flagellar type III protein-export apparatus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**, 599–602.

**Sekiya, K., Ohishi, M., Ogino, T., Tamano, K., Sasakawa, C. & Abe, A. (2001).** Supermolecular structure of the enteropathogenic *Escherichia coli* type III secretion system and its direct interaction with the EspA-sheath-like structure. *Proc Natl Acad Sci U S A* **98**, 11638–11643.

**Simon, R., Priefer, U. & Pühler, A. (1983).** A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. *Nat Biotechnol* **1**, 784–791.

**Sorg, I., Wagner, S., Amstutz, M., Müller, S. A., Broz, P., Lussi, Y., Engel, A. & Cornelis, G. R. (2007).** YscU recognizes translocators as export substrates of the *Yersinia* injectisome. *EMBO J* **26**, 3015–3024.

**Spreter, T., Yip, C. K., Sanowar, S., André, I., Kimbrough, T. G., Vuckovic, M., Pfuetzner, R. A., Deng, W., Yu, A. C. & other authors (2009).** A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nat Struct Mol Biol* **16**, 468–476.

**Tamano, K., Aizawa, S., Katayama, E., Nonaka, T., Imajoh-Ohmi, S., Kuwae, A., Nagai, S. & Sasakawa, C. (2000).** Supramolecular structure of the *Shigella* type III secretion machinery: the needle part is changeable in length and essential for delivery of effectors. *EMBO J* **19**, 3876–3887.

**Tampakaki, A. P., Fadouloglou, V. E., Gazi, A. D., Panopoulos, N. J. & Kokkinidis, M. (2004).** Conserved features of type III secretion. *Cell Microbiol* **6**, 805–816.

**Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., Büttner, D., Caldana, C., Gaigalat, L., Goesmann, A. & other authors (2005).** Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. *J Bacteriol* **187**, 7254–7266.

**Tusnady, G. E. & Simon, I. (2001).** The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849–850.

**Ujwal, M. L., Jung, H., Bibi, E., Manoil, C., Altenbach, C., Hubbell, W. L. & Kaback, H. R. (1995).** Membrane topology of helices VII and XI in the lactose permease of *Escherichia coli* studied by *lacY–phoA* fusion analysis and site-directed spectroscopy. *Biochemistry* **34**, 14909–14917.

**van Geest, M. & Lolkema, J. S. (1996).** Membrane topology of the sodium ion-dependent citrate carrier of *Klebsiella pneumoniae*. Evidence for a new structural class of secondary transporters. *J Biol Chem* **271**, 25582–25589.

**van Geest, M. & Lolkema, J. S. (2000).** Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol Mol Biol Rev* **64**, 13–33.

**Van Gijsegem, F., Gough, C., Zischek, C., Niqueux, E., Arlat, M., Genin, S., Barberis, P., German, S., Castello, P. & Boucher, C. (1995).** The *hrp* gene locus of *Pseudomonas solanacearum*, which controls the production of a type III secretion system, encodes eight proteins related to components of the bacterial flagellar biogenesis complex. *Mol Microbiol* **15**, 1095–1114.

**Van Gijsegem, F., Vasse, J., De Rycke, R., Castello, P. & Boucher, C. (2002).** Genetic dissection of *Ralstonia solanacearum hrp* gene cluster reveals that the HrpV and HrpX proteins are required for Hrp pilus assembly. *Mol Microbiol* **44**, 935–946.

**von Heijne, G. (1986).** The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J* **5**, 3021–3027.

**von Heijne, G. (1992).** Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225**, 487–494.

**von Heijne, G. (2006).** Membrane-protein topology. *Nat Rev Mol Cell Biol* **7**, 909–918.

**Weber, E., Ojanen-Reuhs, T., Huguet, E., Hause, G., Romantschuk, M., Korhonen, T. K., Bonas, U. & Koebnik, R. (2005).** The type III-dependent Hrp pilus is required for productive interaction of *Xanthomonas campestris* pv. vesicatoria with pepper host plants. *J Bacteriol* **187**, 2458–2468.

**Weber, E., Berger, C., Bonas, U. & Koebnik, R. (2007).** Refinement of the *Xanthomonas campestris* pv. *vesicatoria hrpD* and *hrpE* operon structure. *Mol Plant Microbe Interact* **20**, 559–567.

**Wengelnik, K. & Bonas, U. (1996).** HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. vesicatoria. *J Bacteriol* **178**, 3462–3469.

**Wengelnik, K., Marie, C., Russel, M. & Bonas, U. (1996a).** Expression and localization of HrpA1, a protein of *Xanthomonas campestris* pv. vesicatoria essential for pathogenicity and induction of the hypersensitive reaction. *J Bacteriol* **178**, 1061–1069.

**Wengelnik, K., Van den Ackerveken, G. & Bonas, U. (1996b).** HrpG, a key *hrp* regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. *Mol Plant Microbe Interact* **9**, 704–712.

**Yip, C. K., Kimbrough, T. G., Felise, H. B., Vuckovic, M., Thomas, N. A., Pfuetzner, R. A., Frey, E. A., Finlay, B. B., Miller, S. I. & Strynadka, N. C. (2005).** Structural characterization of the molecular platform for type III secretion system assembly. *Nature* **435**, 702–707.

**Yun, C. H., Van Doren, S. R., Crofts, A. R. & Gennis, R. B. (1991).** The use of gene fusions to examine the membrane topology of the L-subunit of the photosynthetic reaction center and of the cytochrome *b* subunit of the $bc_1$ complex from *Rhodobacter sphaeroides*. *J Biol Chem* **266**, 10967–10973.

**Zarivach, R., Deng, W., Vuckovic, M., Felise, H. B., Nguyen, H. V., Miller, S. I., Finlay, B. B. & Strynadka, N. C. (2008).** Structural analysis of the essential self-cleaving type III secretion proteins EscU and SpaS. *Nature* **453**, 124–127.

**Zhu, K., Gonzalez-Pedrajo, B. & Macnab, R. M. (2002).** Interactions among membrane and soluble components of the flagellar export apparatus of *Salmonella*. *Biochemistry* **41**, 9516–9524.

Edited by: C. A. Boucher