CHAPTER 1

LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

Pseudomonas aeruginosa, an ubiquitous Gram-negative bacterium, is an important opportunistic pathogen of humans, causing serious infections in immunocompromised patients such as those with cancer or AIDS, as well as patients suffering from cystic fibrosis and severe burns (Van Delden and Iglewski, 1998). The pathogenesis of this bacterium is attributed to the combined effect of extracellular virulence determinants, including lipases and phospholipases, proteases, exopolysaccharides, alkaline phosphatases, together with properties such as adherence, biofilm formation and resistance to antibiotics (Liu, 1974; Lazdunski *et al.*, 1990; Van Delden and Iglewski, 1998; Elkins *et al.*, 1999; Davey and O'Toole, 2000; Watnick and Kolter, 2000; Donlan, 2002). Despite general agreement that biofilms are the basis for persistent or chronic infection, the understanding of the molecular mechanisms implicated in the biofilm process in still growing (Donlan, 2002).

Biofilms are currently defined as structured bacterial communities enclosed in a self-produced exopolysaccharide matrix and adherent to abiotic or biological surfaces (Costerton *et al.*, 1995). Adherence is often mediated by proteinaceous appendages (flagella, pili, fimbriae) protruding from the cell envelope. Among the best-characterized surface appendages are the type 1 and P pili of *Escherichia coli* (Orndorf and Bloch, 1990; Saulino *et al.*, 1998), type IV pili of *P. aeruginosa* (Strom and Lory, 1993; Hahn, 1997), and curli of *E. coli* and *Salmonella enteritidis* (Olsen *et al.*, 1989; Romling *et al.*, 1998). Recently, a potentially novel class of pili was identified in *Actinobacillus actinomycetemcomitans* (Kachlany *et al.*, 2000) and *Caulobacter crescentus* (Skerker and Shapiro, 2000), which in the case of *A. actinomycetemcomitans*, are associated with the ability of the bacterium to bind nonspecifically to inert surfaces.

The above-mentioned pili or fimbriae are all filamentous multimeric macromolecules and are synthesized through the ordered polymerization of pilin subunits. In general, the bacterial pilus is composed of a repeating polypeptide packed into a helical assembly of which the tip may display a protein adhesin that binds to host cells (Wizemann *et al.*, 1999; Sauer *et al.*, 2000). Their biogenesis involves many genes, including those that encode the major subunit, minor components, proteins required for biogenesis and assembly, and regulatory proteins (Soto and Hultgren, 1999; Sauer *et al.*, 2000). In Gram-negative bacteria, most of the pilus components have to be secreted through the inner membrane, the periplasm and the outer

membrane before reaching their final destination. The general secretory pathway (GSP), which is widespread among Gram-negative bacteria (Thanassi and Hultgren, 2000), permits these proteins to cross first the cytoplasmic membrane, via the Sec system, and then the outer membrane, via specific terminal branches, depending on the structure considered (Soto and Hultgren, 1999).

As the role of pili in the biofilm process is closely related to the aims of this investigation, information pertinent to the role of these bacterial surface appendages during the early stages of biofilm development, as well as the molecular events in the biogenesis of pili found in Gram-negative bacteria will be discussed in greater detail in this review of the literature.

1.2 BIOFILM DEVELOPMENT

Over the past few years, much progress has been made towards understanding the development of bacterial biofilms. This progress has been largely due to the recent focus of analyzing biofilms using genetic (O'Toole and Kolter, 1998a; 1998b; Whiteley *et al.*, 2001), proteomic (Sauer and Camper, 2001; Steyn *et al.*, 2001) and molecular biological (Tolker-Nielsen *et al.*, 2000; Heydorn *et al.*, 2002) approaches. In addition, the results obtained by various biophysical, structural and chemical studies have led to a basic model for biofilm structure (Costerton *et al.*, 1995). In this model, bacteria form microcolonies surrounded by copious amounts of exopolysaccharide (EPS). Interspersed between the microcolonies are water-filled channels that may serve to promote the influx of nutrients and the efflux of waste products (Costerton *et al.*, 1995; 1999). Despite much having been learned about the structure and characteristics of bacterial biofilms, the gene products required for biofilm formation have remained elusive and consequently, the pathways leading to biofilm formation and dissolution have remained poorly understood.

1.2.1 Steps in biofilm development

The formation of a well-developed biofilm (Fig. 1.1) is believed to occur in a sequential process of transport of microorganisms to a surface; initial microbial attachment; formation of microcolonies and finally, the formation of well-developed biofilms (Marshall, 1985; Van Loosdrecht *et al.*, 1990). Prior to surface colonization, a conditioning film, composed of

proteins, glycoproteins and organic nutrients, is believed to form on the attachment surface upon its immersion in liquid (Marshall *et al.*, 1971). Once a surface has been conditioned, its properties are altered so that the affinity of an organism for a native or a conditioned surface can be quite different (Boland *et al.*, 2000). Planktonic bacteria may be brought into close approximation of the conditioned surface by either a random (*e.g.* sedimentation and liquid flow) or in a directed fashion (*e.g.* chemotaxis and active motility) (Quirynen *et al.*, 2000). Initial attachment of the bacteria to the conditioned surface is then facilitated by van der Waals forces, electrostatic and hydrophilic interactions and specific interactions, or by a combination of these, depending on the proximity of the organism to the attachment surface (Carpentier and Cerf, 1993; Zottola and Sasahara, 1994; An *et al.*, 2000).

The initial attachment is followed by a phase during which production of bacterial EPS results in more stable attachment by forming organic bridges between the cells and substratum and/or receptor-specific ligands located on pili, fimbriae and fibrillae or both (Jacob-Dubuisson *et al.*, 1993; Jones *et al.*, 1995; Rudel *et al.*, 1995; Pratt and Kolter, 1998). Once the bacteria have irreversibly attached to a surface, the process of biofilm maturation begins. During this process, the growth and multiplication of firmly attached primary colonizing bacteria lead to the formation of microcolonies, which may subsequently develop into mushroom- or pillar-like structures interspersed with fluid-filled channels (Costerton *et al.*, 1995; Kurchma and O'Toole, 2000). Once fully developed, a biofilm generates altered patterns of bacterial growth, physiological cooperation and metabolic efficiency (Costerton *et al.*, 1995; 1999).

The growth potential of the bacterial biofilm is ultimately limited by the availability of nutrients in the immediate environment, the expression of quorum-sensing molecules released in response to nutrient limitation, accumulation of toxic by-products and other factors, including pH, oxygen perfusion, carbon source availability and osmolarity (La Tourette Prosser *et al.*, 1987; Carpentier and Cerf, 1993; Allison *et al.*, 1998; Davies *et al.*, 1998). At some point, the biofilm reaches critical mass and a dynamic equilibrium is reached at which the cells farthest from the surface may consequently detach and together with progeny of other biofilm cells may colonize other surfaces (Korber *et al.*, 1989; Heydorn *et al.*, 2002).

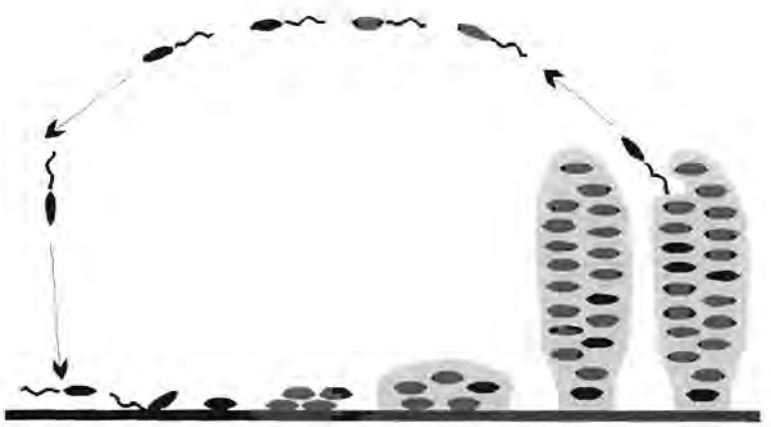

Fig. 1.1 Model of biofilm development. In response to environmental cues, planktonic cells initiate cell-to-surface and cell-to-cell contacts resulting in the formation of microcolonies. In response to developmental signals, microcolonies undergo differentiation to form a well-developed biofilm characterized by pillar- or mushroom-like structures surrounded by bacterial exopolysaccharides (EPS) and interspersed with fluid-filled channels. Once the biofilm has reached critical mass, some of the biofilm cells may detach to colonize other surfaces (Modified from O'Toole *et al.*, 2000).

1.3 BACTERIAL COMPONENTS REQUIRED FOR INITIAL ATTACHMENT

Of the processes leading to the formation of well-developed biofilms, bacterial structural components required for initial attachment have been best characterized, primarily through mutation analysis. The rate and extent of attachment of bacterial cells to a surface is influenced by cell surface hydrophobicity, presence of flagella, pili and adhesins, outer membrane proteins and production of EPS (O'Toole and Kolter, 1998a; 1998b; DeFlaun *et al.*, 1999; Genevaux *et al.*, 1999; Espinosa-Urgel *et al.*, 2000). In addition, recent evidence suggests that the primary development of a biofilm might be regulated at the level of population density-dependent gene expression controlled by cell-to-cell signaling molecules such as acylated homoserine lactones (McLean *et al.*, 1997; Allison *et al.*, 1998; Davies *et al.*, 1998). In the following section, advances made towards revealing the role of flagella and pili in bacterial adhesion to surfaces will be specifically addressed.

1.3.1 Importance of flagella, pili and adhesins in bacterial attachment to surfaces

The requirement for flagella has emerged as a common theme in biofilm formation in several Gram-negative bacteria subjected to genetic analysis, e.g. *E. coli* (Pratt and Kolter, 1998), *P. aeruginosa* and *P. fluorescence* (O'Toole and Kolter, 1998a; 1998b), and *Vibrio cholerae* (Watnick and Kolter, 1999). For each of these bacteria, mutations in genes involved in flagellar-mediated motility hinder biofilm formation under quiescent conditions, *e.g.* microtitre plate wells.

Several non-motile mutant strains of *P. aeruginosa* PA14 have been isolated in screens defective for biofilm formation (O'Toole and Kolter, 1998a). The *P. aeruginosa* mutants, designated *sad* for surface attachment defective, could be divided into two groups. One group of strains were found to harbor mutations in genes with homology to flagellar genes of other organisms, while a second group of *sad* mutants were defective in the biogenesis of type IV pili, which are known to be involved in surface-associated movement referred to as twitching motility. Microscopic analysis of wild-type *P. aeruginosa*, non-motile *P. aeruginosa* and twitch-negative *P. aeruginosa* revealed that flagellar-mediated motility is important in establishing cell-surface contacts, whereas the twitching motility appears to play a role in the formation of microcolonies within the biofilm (O'Toole and Kolter, 1998a).

E. coli has also been reported to require flagella and pili to initiate the early attachment process (Genevaux *et al.*, 1996; Pratt and Kolter, 1998). However, the biofilm phenotype of *E. coli* flagellum mutants is different from that of *P. aeruginosa* mutants and the roles that flagella play in the formation of biofilms of *E. coli* and *P. aeruginosa* appear to be different (Pratt and Kolter, 1998). Attachment is not completely eliminated in *E. coli* 2K1056 flagellum mutants, although it is severely impaired, and the biofilm that forms consists of isolated microcolonies (Pratt and Kolter, 1998). In *E. coli*, flagellar-mediated motility may be required for movement parallel to the surface, in addition to bringing the bacteria into proximity to the surface (Pratt and Kolter, 1998). In non-motile strains of *E. coli*, cell surface adhesins, known as curli, have been reported to play a role during early attachment events in biofilm formation. In a study performed by Vidal *et al.* (1998), a non-motile *E. coli* K-12 mutant strain was used to select for mutants that gained the ability to attach to polyvinylchloride (PVC). A gain-of-function allele in *ompR* was isolated and shown to increase production of curli, which, in turn, was shown to be required for biofilm formation in the non-motile strain (Vidal *et al.*, 1998). Thus, it may be possible that under certain conditions, a different pathway is utilized that bypasses the requirement for flagellar-mediated motility, and this distinct pathway uses the curli surface adhesin (Vidal *et al.*, 1998). Attachment is also reduced by mutations in the type 1 pili biosynthetic gene *fimH*, which encodes a mannose-specific adhesin (Pratt and Kolter, 1998). Type 1 pili, however, do not appear to play a role in moving the bacteria across the surface.

The role of surface structures in the ability of *V. cholerae* El Tor to form biofilms appears to be similar to what has been observed for *E. coli* (Watnick and Kolter, 1999; Watnick *et al.*, 1999). Although motility is important for *V. cholerae* biofilm formation in Luria-Bertani broth, biofilms do eventually form in *V. cholerae* flagellum mutants, albeit at a slower rate than the wild-type (Watnick and Kolter, 1999). The flagella are thought to be important for bringing bacteria in close proximity of a surface and for bacterial spread across the surface. Depending on the surface to which *V. cholerae* attaches, the bacterium appears to utilize different pathways for initial attachment. For example, *in vivo* the toxin-coregulated pilus (Tcp) is required for colonization of the intestine (Herrington *et al.*, 1988), whereas the type IV mannose-sensitive hemagglutinin pilus (MshA) is required for attachment to abiotic surfaces and does not play a role in host colonization (Thelin and Taylor, 1996). The MshA pilus also appears to speed the attachment of bacteria to a surface. The analysis of mature biofilms formed by *V. cholerae* flagellum and *mshA* mutant strains, using confocal scanning

laser microscopy (CSLM), revealed that although they are slightly delayed in biofilm formation, the mature biofilm formed by mutants lacking these surface structures is indistinguishable from that formed by the wild-type strain (Watnick *et al.*, 1999).

1.4 FIMBRIAL EXPRESSION AND ASSEMBLY

From the preceding section, it follows that adhesion of bacteria to a surface is an essential first step in the initiation of biofilm development. Consequently, several adhesion factors that may play a role in this process have been studied. Whereas some attachments are achieved by non-fimbrial structural adhesins that are present as monomers or oligomers on the outer membrane, other attachment is mediated by surface organelles such as fimbriae or pili (Abraham *et al.*, 1998; O'Toole and Kolter, 1998a; Pratt and Kolter, 1998). The fimbriae have been classified based on morphological, serological, biochemical or functional criteria (Sauer *et al.*, 2000) and details on the molecular mechanisms of fimbrial biosynthesis have been worked out to varying degrees in different systems. Until now, at least five fundamentally different secretion systems have been described in Gram-negative bacteria, some of which are sometimes associated with surface appendages (Lory, 1998; Soto and Hultgren, 1999; Thanassi and Hultgren, 2000). Whereas type IV pili in *P. aeruginosa* are assembled via the type II secretion system (Nunn, 1999), type 1 and P pili in *E. coli* are translocated to the cell surface by a chaperone-usher pathway (Sauer *et al.*, 2000) and curli fimbriae, in *E. coli* and *Salmonella* spp., are assembled by an extracellular nucleation-precipitation pathway (Romling *et al.*, 1998). These pili and their biogenesis will subsequently be discussed in greater detail.

1.4.1 Type 1 and P pili

The type 1 and P pili are important virulence factors expressed in uropathogenic *E. coli* (UPEC) that promote colonization of the urinary tract by mediating binding to mucosal epithelial cells (Roberts *et al.*, 1994; Connell *et al.*, 1996). Whereas type 1 pili mediate binding to mannose-oligosaccharides (Krogfeldt *et al.*, 1990), the P pili mediate binding to glycolipid receptors on the uroepithelial cells (Zhang and Normark, 1996). Binding of these fimbriae to the host cell, however, signals the host and consequently triggers the host cell-

signaling pathways to respond to the bacterial attachment by eliciting the release of pro-immunoinflammatory cytokines in epithelial cells (Svanborg *et al.*, 1996; Mulvey *et al.*, 1998). Thus, these fimbriae are not only involved in bacteria-host interaction, but may also be involved in host-bacteria signaling. Genetic, biochemical and structural studies have revealed that type 1 and P pili resemble each other in their gene order, organization, assembly and regulation (Kuehn *et al.*, 1992; Saulino *et al.*, 1998). The type 1 and P pili gene clusters are depicted in Fig. 1.2.

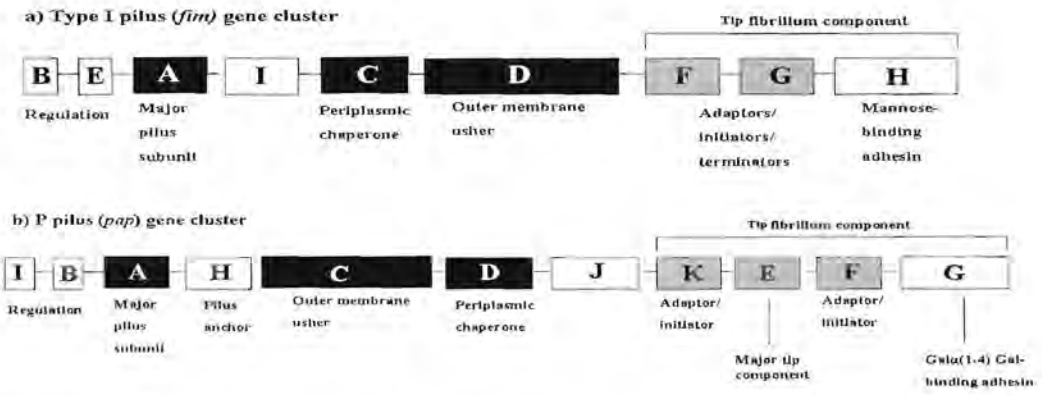

Fig. 1.2 Operons encoding type 1 (a) and P pili (b) in uropathogenic *E. coli* strains. The operons contain the genes encoding the structural subunits of the pilus shaft and tip fibrillum, including the adhesin, as well as their respective periplasmic chaperones and outer membrane ushers. Notably, these two operons display a similar gene organization and the gene arrangement resembles the structural organization of the pili (after Schilling *et al.*, 2001).

1.4.1.1 General characteristics

The expression and assembly of type 1 pili requires at least nine genes (Fig. 1.2a), which are present in the type 1 *fim* gene cluster (Hull *et al.*, 1981; Hultgren *et al.*, 1991). The type 1 pili are composite structures consisting of a long rod and a thin tip. Whereas the long rod is arranged in a right-handed helical conformation by FimA subunits, the short tip fibrillar structure contains FimG, the mannose-sensitive FimH adhesin and possibly FimF (Maurer and Orndorff, 1985; Jones *et al.*, 1995). The type 1 subunits are arranged in a helix with an external diameter of 6 to 7 nm and an axial hole of 20 to 25 Å, with a pitch distance of 23.1 Å

and 3.125 subunits per turn (Brinton, 1965; Kuehn *et al.*, 1994; Saulino *et al.*, 1998). The two minor components of type 1 pili, FimF and FimG, are involved in the initiation and termination of pili assembly, respectively (Russell and Orndorff, 1992; Jones *et al.*, 1995). Two proteins, FimC and FimD, play a role in fimbrial assembly. Whereas the FimC protein is a chaperone-like molecule (Jones *et al.*, 1993), FimD is an outer membrane usher protein that controls translocation of fimbrial subunits across the outer membrane and stabilizes the chaperone-adhesin complex to initiate fimbrial assembly (Klemm and Christiansen, 1990).

By contrast to type 1 pili, eleven genes in the *pap* gene cluster (Fig. 1.2b) are required for the expression and assembly of the P pili (Hull *et al.*, 1981; Hultgren *et al.*, 1991; Marklund *et al.*, 1992). The P pili are also composite fibers consisting of flexible fibrillae joined end-to-end to pilus rods (Kuehn *et al.*, 1992). The rod is composed of repeating PapA subunits packed into a right-handed helical assembly, with an external diameter of 68 Å, an axial hole of 15 Å, and a pitch distance of 24.9 Å, with 3.28 subunits per turn of the helical cylinder (Bullitt and Makowski, 1995; Gong and Makowski, 1992). The rod is terminated by PapH, which may serve to anchor the pilus in the membrane (Baga *et al.*, 1987). The tip fibrillum is comprised mostly of PapE subunits, as well as several other minor subunits (Kuehn *et al.*, 1992; Bullitt and Makowski, 1995). The adhesin of P pili, PapG, is located at the distal end of the tip and is joined to the PapE fibrillum via an adaptor protein, PapF, and another adaptor protein, PapK, joins the adhesion-containing tip to the PapA rod (Kuehn *et al.*, 1992; Jacob-Dubuisson *et al.*, 1993). Two assembly proteins, a periplasmic chaperone PapD and an outer membrane usher PapC, orchestrate P pilus biogenesis (Norgen *et al.*, 1987; Kuehn *et al.*, 1991; Dodson *et al.*, 1993).

1.4.1.2 Model for biogenesis of type 1 pili and P pili of *E. coli*

The assembly of type 1 and P pili proceeds by the highly conserved chaperone-usher pathway (Kuehn *et al.*, 1994), which participates in the biogenesis of at least 30 adhesive organelles in a wide range of Gram-negative pathogenic bacteria (Hung and Hultgren, 1998; Sauer *et al.*, 2000). The assembly machinery is comprised of two specialized classes of proteins, a periplasmic immunoglobulin-like chaperone and an outer membrane usher.

During biogenesis of type 1 and P pili (Fig. 1.3), the respective pilus subunits are translocated across the cytoplasmic membrane in an unfolded state via the general secretion pathway

(Sec). A periplasmic chaperone, FimC or PapD, interacts with newly translocated pilus subunits and facilitates the proper folding of subunits (Jacob-Dubuisson *et al.*, 1994; Jones *et al.*, 1997). The pilus subunits are targeted to the periplasmic chaperone through motifs in the amino-terminal and carboxy-terminal regions of the subunits (Kuehn *et al.*, 1993; Soto *et al.*, 1998). Folding of subunits on the chaperone may occur concomitantly with their release from the membrane, with the chaperone providing a template that nucleates the folding process. However, the periplasmic disulfide isomerase DsbA also plays a role in subunit folding and may be part of a two-step process in which a newly secreted subunit first interacts with DsbA and is then passed to the chaperone (Jacob-Dubuisson *et al.*, 1994). In the absence of chaperone, free subunits undergo misfolding and aggregation. Such protein aggregation in the periplasm is sensed by at least two signal transduction systems, namely the CpxA-CpxR two-component system in which CpxA is the membrane-bound sensor/kinase and CpxR is the DNA-binding response/regulator, and the sigma E (σ^E) modulatory pathway in which RseA and RseB act to negatively regulate the σ^E (RpoE) transcription factor (Danese and Silhavy, 1997; Delas *et al.*, 1997; Pogliano *et al.*, 1997). Both these pathways activate transcription of the *degP* gene, which encodes the DegP periplasmic protease normally responsible for breaking down subunit aggregates in the periplasm (Jones *et al.*, 1997).

The crystal structure of PapD (Holmgren and Brändén, 1989) and the crystal structures of the PapD-PapK chaperone-subunit complex (Sauer *et al.*, 1999), as well as the FimC-FimH chaperone-adhesin complex (Choudhury *et al.*, 1999) have all been solved. The chaperone (PapD or FimC) consists of two immunoglobulin (Ig)-like domains, forming L-shaped molecules. The pilus subunits (PapK or FimH) have incomplete Ig-like structure, lacking the seventh C-terminal β-strand present in canonical Ig folds. The absence of this strand generates a deep groove along the surface of the pilin domain and exposes its hydrophobic core. In the presence of the chaperone, the chaperone contributes its G1 β-strand to the incomplete folds in the subunits to fill the groove by running parallel to the subunit carboxy-terminal F strand, and thereby stabilizes the interaction. This process is termed donor strand complementation and thus leads to the capping of one of the subunit's interactive surfaces and prevents premature pilus formation in the periplasm (Choudhury *et al.*, 1999; Sauer *et al.*, 1999). Subunits assembled by the chaperone-usher pathway have an amino-terminal extension that does not contribute to the Ig fold of the subunit, but rather projects away from the rest of the pilin domain where it would be free to interact with another subunit (Sauer *et al.*, 1999). During pilus biogenesis, the highly conserved amino-terminal domain of one

subunit may therefore insert anti-parallel to the F strand of the neighboring subunit and thereby displaces the chaperone G1 β-strand from its preceding subunit in a mechanism termed donor-strand exchange. The mature pilus would thus always consist of a complete array of canonical Ig domains, each of which donates a strand to the fold of the preceding subunit to form a highly stable organelle (Barnhardt *et al.*, 2000).

Following the proper folding of the pilus subunits and their release from the cytoplasmic membrane, the chaperone-subunit complexes are targeted to the usher PapC (or FimD) in the outer membrane (Dodson *et al.*, 1993). The usher is needed for translocation of subunits across the outer membrane. Both the purified PapC and FimD ushers have been shown to form pores when reconstituted into liposomes (Thanassi *et al.*, 2002). The P pilus usher, PapC, has been shown to assemble into 15-nm-diameter ring-shaped complexes containing central pores of 2-3 nm in diameter, and the PapC complexes consist of at least six and possibly up to 12 subunits (Thanassi *et al.*, 1997). PapC is furthermore predicted to have a largely β-sheet secondary structure, typical of bacterial outer membrane pore-forming proteins, and they probably present large regions to the periplasm for interaction with chaperone-subunit complexes (Valent *et al.*, 1995). To facilitate pilus assembly, the usher must be able to translocate pilus subunits across the outer membrane. The 2-nm-wide linear tip fibrillum would be able to pass through the 2- to 3-nm-diameter usher channel, but the 6.8-nm-wide helical pilus rod would not be able to fit through the usher. It has subsequently been shown that the P pilus rods, as well as type 1 pilus rods, can be unraveled into linear fibers measuring 2 nm in diameter, which would therefore be narrow enough to pass through the usher pore (Abraham *et al.*, 1992; Bullitt and Makowski, 1995; Thanassi *et al.*, 1998). The linear pilus fiber is thus proposed to be translocated across the outer membrane and adopts its final helical conformation upon reaching the cell surface.

The chaperone-usher pathway does not appear to require input energy of external energy for assembly and secretion of pili across the outer membrane. Winding of the PapA (or FimA) fiber into a helix on the external surface of the cell may provide the driving force for the translocation of the pilus across the outer membrane (Jacob-Dubuisson *et al.*, 1994; Saulino *et al.*, 1998). This, combined with the targeting affinities of the chaperone-subunit complexes for the usher and the binding specificities of the subunits for each other, may provide all the energy and specificity necessary for the ordered assembly and translocation of pili across the outer membrane.

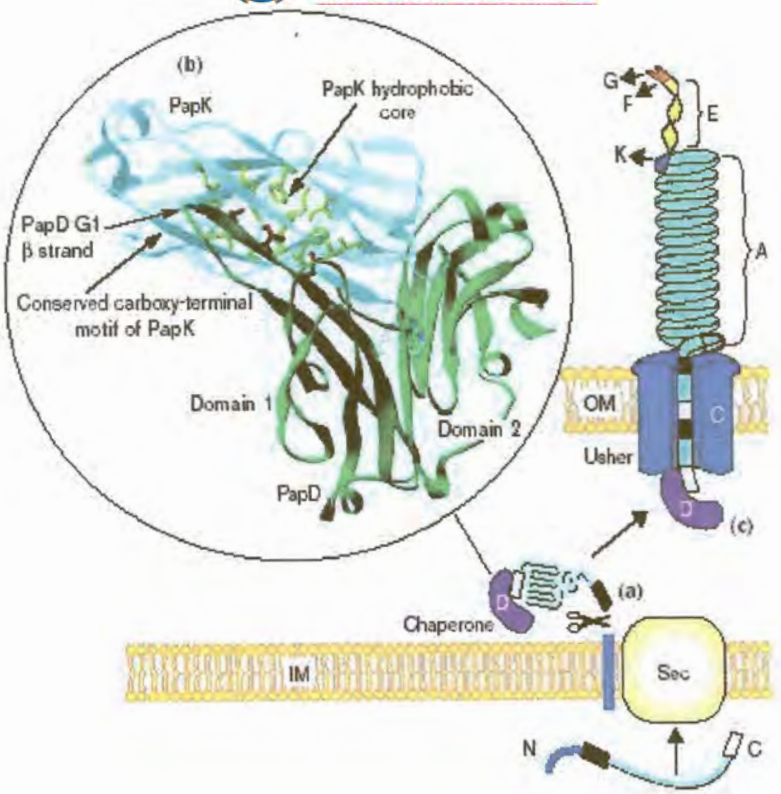

Fig. 1.3 Model for biogenesis of P pili by the chaperone-usher pathway. (a) Pilus subunits cross the inner membrane (IM) via the Sec system, followed by cleavage of their amino-terminal signal sequence. The periplasmic chaperone PapD binds to each subunit via a conserved carboxy-terminal subunit motif (white box), allowing proper subunit folding and preventing premature subunit-subunit interactions. (b) Crystal structure of the PapD-PapK chaperone-subunit complex. The chaperone (green) consists of two Ig folds. The subunit (blue) consists of a single Ig fold that lacks the usual seventh β strand, resulting in exposure of its hydrophobic core. The G1 β strand of PapD binds to the conserved carboxy-terminal motif of PapK, donating its hydrophobic residues to complete the structure of the subunit in a mechanism termed donor strand complementation. (c) Chaperone-subunit complexes are targeted to the outer membrane (OM) usher for assembly into pili and secretion across the outer membrane. Subunit-subunit interactions are thought to take place by interaction of conserved amino-terminal (black box) and carboxy-terminal (white box) motifs. The amino-terminal motif of one subunit may complete the structure of the preceding subunit in a mechanism termed donor strand exchange to build the pilus fiber. The usher channel is only able to allow passage of a linear fiber of folded subunits, forcing the pilus rod to adopt its final helical conformation at the cell surface. The location of Pap subunits in the pilus is indicated. (Reprinted from Current Opinions in Cell Biology, Vol. 12, D.G. Thanassi and S.J. Hultgren, Multiple pathways allow protein secretion across the bacterial outer membrane, pp. 420-430, 2000, from Elsevier).

1.4.2 Curli

Several *E. coli* and *Salmonella* strains produce a class of long, thin, irregular, flexible and highly aggregated surface structures known as curli (Collinson *et al.*, 1996; Olsen *et al.*, 1998; Romling *et al.*, 1998; Prigent-Combaret *et al.*, 1999). These surface organelles are distinct from other types of pili in terms of both their morphology and mechanism of assembly (Soto and Hultgren, 1999). Curli bind to several matrix and plasma proteins, including fibronectin, plasminogen, laminin and surfaces such as agar and plastic to initiate adherence and colonization (Arnquist *et al.*, 1992; Olsen *et al.*, 1998; Bian *et al.*, 2000). Bacterial cells expressing curli on their surface tend to autoaggregate, indicating that curli may mediate adhesion to other bacterial cells, a property that may be important for curli-mediated biofilm formation (Vidal *et al.*, 1998; Prigent-Combaret *et al.*, 1999).

1.4.2.1 General characteristics

The cluster of genes coding for curli expression and biosynthesis is termed the *csg* operon in *E. coli* and the *agf* operon in *Salmonella* (Romling *et al.*, 1998). The two divergently transcribed curli operons, *agfBAC* and *agfDEFG*, of *Samonella* exhibit the same gene organization as in *E. coli* (Collinson *et al.*, 1996; Romling *et al.*, 1998), and the predicted proteins encoded by these operons share a significant level of primary sequence similarity (86-99%) with those of *E. coli* (Romling *et al.*, 1998). In *E. coli*, the two divergently transcribed operons, *csgBA* and *csgDEFG*, are located on one chromosomal region in *E. coli* (Hammar *et al.*, 1995). Whereas the *csgBA* operon encodes the components of the curli, the *csgDEFG* operon encodes a transcriptional activator for curli production and three putative assembly factors.

The major component of *E. coli* curli is the CsgA protein, which is secreted into the extracellular milieu as a soluble protein. CsgB is a minor component and may be found associated with the outer membrane or distributed along the length of the curli fiber where it has been suggested to be able to initiate branching of the fibrillar structure (Bian and Normark, 1997). The *csgD* gene encodes a transcription regulator, CsgD, belonging to the LuxR/UhpA family of transcriptional regulators, which controls the transcription of the *csgBA* operon (Hammar *et al.*, 1995). The roles of CsgE and CsgF are not well understood. Whereas inactivation of the *csgE* gene does not significantly affect formation of curli fibers

(Provence *et al.*, 1992), mutation of *csgF* eliminates curli, without disrupting CsgA secretion into the extracellular environment (Romling *et al.*, 1998). CsgG, a lipoprotein located in the outer membrane, is required for maintaining the stability of the CsgA-CsgB complex, thereby preventing them from premature proteolysis (Loferer *et al.*, 1997). Since the CsgE, CsgF and CsgG proteins do not form part of the final curli structure, it is thought that they may form part of the assembly apparatus required for assembly of the curli (Hammar *et al.*, 1995).

Expression of curli is highly regulated by a range of different environmental signals such as temperature, osmolarity and growth conditions. Curli are expressed optimally at temperatures below 30°C, under low nutrients and low medium osmolarity, and during the stationary phase (Olsen *et al.*, 1989; 1993). Transcriptional activation of the *csgD* promoter in both *E. coli* and *Salmonella typhimurium* is dependent on RpoS, while OmpR, an osmolarity-sensing transcription regulator, is needed for activation of both *csgBA* and *csgDEFG* promoters in *E. coli* and *S. typhimurium* (Hammar *et al.*, 1995; Romling *et al.*, 1998; Vidal *et al.*, 1998).

1.4.2.2 Model for biosynthesis of curli fimbriae

It has been proposed that the assembly of the curli fiber of *E. coli* occurs extracellulary by a nucleation-precipitation pathway (Hammar *et al.*, 1996). Unlike P and type IV pili that undergo assembly from the base, the formation of curli fimbriae seems to occur from both the tops and the cell-associated bases of the fibers. However, how this assembly effectively proceeds extracellulary and what mechanism(s) is involved in regulation of this process remain unknown.

In *E. coli*, protein products encoded by both the *csgBA* and *csgDEFG* operons are required for biogenesis of curli. The two components of curli, CsgA and CsgB, have significant sequence homology and structural similarity. Both proteins contain *sec*-like amino-terminal signal peptides, and their mature domains have 4 to 5 repeats of a consensus motif with a β-strand-turn-β-strand-turn structure (Hammar *et al.*, 1996). Only the mature forms of CsgA and CsgB (13-kDa polypeptides) can be detected in the curli fibers. As with CsgA and CsgB, the CsgG lipoprotein contains a signal peptide that is removed during secretion. It has been reported that CsgG is located on the periplasmic side of the outer membrane, and mutants lacking CsgG accumulate CsgA and CsgB in the periplasmic space (Loferer *et al.*, 1997). Consequently, CsgG is needed for export of the two components of curli across the outer

membrane. However, it remains to be determined whether a multimeric form of CsgG itself may function as a Csg-specific channel within the outer membrane or whether CsgG might function as a chaperone that prevents misfolding and proteolysis of the curli subunits in the periplasm. Both CsgE and CsgF are also candidate chaperones of the curli secretion system. In the absence of CsgB, CsgA does not polymerize on the cell surface, but is released into the extracellular milieu. It has thus been proposed that CsgB functions either as a nucleator that triggers polymerization of CsgA or as a platform for the assembly of curli fibers (Bian and Normark, 1997).

1.4.3 Type IV pili

The pili from a broad spectrum of Gram-negative bacteria are grouped as type IV on the basis of amino acid sequence similarities among their major pilin subunit. The homology between different type IV pilins is highest at their amino-terminus, but also extends to some areas of the carboxy-terminus, where other conserved features, like a pair of cysteine residues that form a disulfide bridge in the mature protein, can be identified (Strom and Lory, 1993). The type IV pili are regarded as important virulence factors and occur in many different pathogens, including *P. aeruginosa*, enteropathogenic *E. coli* (EPEC), *Moraxella bovis*, pathogenic *Neisseria* and *V. cholerae* (Strom and Lory, 1993). These pili have been implicated in a variety of functions, including adhesion to biotic and abiotic surfaces (Hahn, 1997; Kang *et al.*, 1997; O'Toole and Kolter, 1998b), modulation of target cell specificity (Bieber *et al.*, 1998), bacteriophage adsorption (Rehmat and Shapiro, 1983; Roncero *et al.*, 1990; Mattick *et al.*, 1996) and twitching motility (Darzins 1993; 1994; Wall and Kaiser, 1999).

1.4.3.1 General characteristics

The major subunit of the type IV pilus is type IV pilin, which possesses a number of unique properties. It contains a short (6 to 7 amino acids) basic amino-terminus leader peptide, a modified amino acid (N-methylophenylalanine) at the amino terminus of mature pilin, a highly hydrophobic amino-terminal domain, and a disulfide-bonded carboxy-terminal domain (Hobbs and Mattick, 1993; Pugsley, 1993; Alm and Mattick, 1997). A specialized/specific inner membrane signal peptidase is required to remove the leader peptide and for maturation

of the pilin subunit (Nunn and Lory, 1991; Strom and Lory, 1992). Based on immunologic and crystallographic data, the type IV pili structures have a diameter of 60 Å and are typically up to 4 000 nm long, with a pitch distance of approximately 40 Å and about 5 subunits per turn (Parge *et al.*, 1995). Although most type IV pili are flexible rod-like structures and arranged in a helical manner, the toxin-coregulated pili (Tcp) of *V. cholerae* (Taylor *et al.*, 1987) and the bundle-forming pili (Bfp) of enteropathogenic *E. coli* (Girón *et al.*, 1991) have some unique characteristics among type IV pili. The TcpA and BfpA pilins assemble into straight fibers of variable length that have a strong tendency to aggregate laterally (Bieber *et al.*, 1998). In addition, whereas most type IV prepilin proteins are characterized by a short basic leader sequence and a conserved phenylalanine residue that follows the cleavage site (referred to as type IVA pili), the latter prepilins are characterized by larger leader peptides and the absence of phenylalanine in the position immediately after the conserved leader peptide cleavage site (referred to as type IVB pili).

At present, only two tip adhesins have been identified among type IV pili. These are PilC of *N. gonorrhoeae* (Rudel *et al.*, 1995) and *N. meningitides* (Ryll *et al.*, 1997), respectively. There is a homologue of *Neisseria pilC* in *P. aeruginosa*, which is termed *pilY* (Alm *et al.*, 1996). However, the role of PilY is unclear since the main adhesion determinant of *P. aeruginosa* type IV pili is found in their major pilin subunit (Farinha *et al.*, 1994; Lee *et al.*, 1994). Thus, although tip adhesins may exist in type IV pili, it would appear that the major pilins can participate directly in adhesion to surfaces.

1.4.3.2 Model for type IV pilus biogenesis

Although the genes responsible for biogenesis of type IV pili are typically located in various regions in the bacterial chromosome (Strom and Lory, 1993), exceptions do occur. Whereas the *tcp* genes of *V. cholerae* are clustered in a single region of the chromosome (Kaufman *et al.*, 1993), the *bfp* genes of enteropathogenic *E. coli* (EPEC) are present in a 80-kb virulence plasmid (Girón *et al.*, 1991; Stone *et al.*, 1996). In recent years, *P. aeruginosa* has been used as the primary model for studying fimbrial biology. Characterization of *P. aeruginosa* mutants which lack twitching motility has led to the identification of a large number of genes, located in six chromosomal clusters, involved in fimbrial biogenesis and function. Many of these genes have homology to other gene/protein sets involved in protein secretion and DNA uptake in various bacteria (Hobbs and Mattick, 1993; Alm and Mattick, 1997). These genes

can be divided into four groups: (i) four transcriptional regulators (*pilS*, *pilR*, *fimS*, *algR*); (ii) eight *che*-like genes which control flagellar-based swimming and type IV pilus-based twitching (Darzins and Russell, 1997) (*pilG*, *H*, *I*, *J*, *K*, *L*, *chpA* and *chpB*); (iii) 19 type IV pilus biogenesis genes (*pilA*, *B*, *C*, *D*, *E*, *F*, *M*, *N*, *O*, *P*, *Q*, *V*, *W*, *X*, *Y1*, *Y2*, *Z*, *fimT* and *fimU*); and (iv) two pilus function genes (*pilT* and *pilU*) (Alm and Mattick, 1997; Wall and Kaiser, 1999).

Several of the chromosomal gene clusters contain genes directly involved in fimbrial assembly. The *pilABCD* locus encodes the major pilin (PilA) and three ancilliary proteins, namely: PilB, a cytoplasmic protein possessing an ATP-binding motif; PilC, an inner membrane protein; and PilD, a prepilin peptidase (Nunn *et al.*, 1990; Nunn and Lory, 1991; Nunn and Lory, 1992). Inactivation of the *pilB* gene or mutagenesis of the ATP-binding motif results in the inability to assemble fimbriae on the cell surface (Turner *et al.*, 1993). It is thus likely that PilB plays a role in providing energy for the assembly and function of the type IV pilus export apparatus. Interestingly, both the PilT and PilU proteins, like the PilB protein, contain nucleotide-binding motifs, suggesting that they also have NTPase activity (Whitchurch *et al.*, 1991; Whitchurch and Mattick, 1994). However, by contrast to *pilB* mutants, both *pilT* and *pilU* mutants express structurally normal pili, but lack motility (Whitchurch *et al.*, 1991; Whitchurch and Mattick, 1994; Wu *et al.*, 1997). Thus, although these proteins are not thought to be part of the pilus fiber, they are essential for twitching motility. Located 25 kb from *pilABCD*, the genes *fimT*, *fimU*, *pilV* and *pilE* encode products of which the function(s) is not yet known, but they contain prepilin-like amino-terminal hydrophobic domains (Russell and Darzins, 1994; Alm and Mattick, 1995; 1996). Whereas *pilV*, *pilE* and *fimU* mutants are unable to produce extracellular assembled fimbriae and accumulate PilA within the membrane fraction, a *fimT* mutant resembles the wild-type phenotype (Alm and Mattick, 1996).

The operon *pilMNOPQ* encodes components of pilus assembly located at the inner membrane and outer membrane, and mutants in these genes are non-fimbriated (Martin *et al.*, 1995). PilQ appears to be the sole protein of the assembly system that is an integral outer membrane component (Martin *et al.*, 1993) and it is capable of oligomerizing to form a gated channel with a central cavity of *ca.* 55 Å (Tonjum and Koomey, 1997; Bitter *et al.*, 1998). The size of the central pore formed by PilQ oligomers is in agreement with the calculated outer diameter (52 Å) of the type IV pili (Folkhard *et al.*, 1981). These properties make PilQ the ideal

the ideal candidate for assembly of the channel through which the pilus rod is exported to the extracellular milieu. Apart from PilQ, the specific components of type IV pilus secretion systems are inner membrane or cytoplasmic proteins. The only periplasmic protein recognized to play a role in these secretion systems is the oxidoreductase DsbA (Raina and Missiakas, 1997). DsbA is required for pilin stability and assembly into bundle forming pili in enteropathogenic *E. coli* (Zhang and Donnenberg, 1996), while in *V. cholerae*, a gene homologue to *dsbA* (*tcpG*) was found to be necessary for secretion of cholera toxin and the assembly of toxin-coregulated pili (Yu *et al.*, 1992).

With reference to the other proteins involved in fimbrial biogenesis, there are a number of other genes whose products are involved in the export of the fimbrial subunit, as mutants in these genes are non-fimbriated. These include the *pilW* and *pilX* genes, which encode prepilin-like proteins, the *pilY1* gene, which encodes a part homologue of the gonococcal PilC adhesin and the *pilY2* gene, which encodes a novel small protein of unknown function (Alm *et al.*, 1996). In addition, both *pilF* and *pilZ* mutants are blocked in the export/assembly of fimbrial biogenesis, with the processed subunit accumulating in the membrane fraction (Alm *et al.*, 1997; Watson *et al.*, 1996a; 1996b).

Although knowledge regarding type IV pilus biogenesis remains incomplete, the following working model for type IV pilus assembly of *P. aeruginosa* (Fig. 1.4) has been proposed by Alm and Mattick (1997) and Hahn (1997). Following translocation of the pre-PilA precursor subunits into the periplasmic compartment by the general secretory pathway, these molecules are transitorily anchored into the inner membrane, at the periplasmic side, by a conserved hydrophobic domain located at the amino-terminus, immediately after the signal peptide, while the hydrophilic C-terminal domains are oriented towards the periplasm (Pugsley, 1996; Kaufman *et al.*, 1991). The hydrophobic amino-terminus is subsequently cleaved at the cytoplasmic side of the inner membrane by prepilin peptidase, PilD, to generate mature PilA (Nunn and Lory, 1991). The mature pilins pack into a pilus fiber in a helical structure with its hypervariable regions exposed. The core of the pilus forms a continuous hydrophobic layer with the inner membrane. This hydrophobic continuum facilitates the low energy requiring polymerization and depolymerization of the pili that is under control of the inner membrane-associated protein complex. PilT may function as an ATPase or kinase and transduce energy for depolymerization. However, PilB, another nucleotide-binding motif containing protein, may also transduce energy for pilus polymerization (Wall and Kaiser, 1999). The assembled

pilin is thought to be translocated across the outer membrane via a gated channel formed by a multimeric complex of PilQ. PilP, a probable lipoprotein (Martin *et al.*, 1995), may function in stabilizing the PilQ complex. In the case of *Neisseria* spp., PilC facilitates the translocation of the pili through the pore and thereby presents itself to the pilus tip, where it functions as a specific adhesin (Nassif *et al.*, 1994; Rudel *et al.*, 1995).

1.4.3.3 Relationship of the type IV pilus biogenesis system with the type II secretion system

All type IV pili use a similar secretion and assembly machinery, which is functionally and evolutionary related to components of the type II secretion system of Gram-negative bacteria (Pugsley, 1993; Russel, 1998) (Fig. 1.4). Consequently, the type IV pilus assembly machinery has been grouped as part of the type II secretion pathway. In *P. aeruginosa*, PilD is shared between the type II (Xcp) export system (Filloux *et al.*, 1998) and the type IV pilus biogenesis system (Strom *et al.*, 1991). PilD (XcpA) is a bifunctional enzyme carrying out both cleavage and N-methylation of the PilA pilin subunit, as well as four other proteins (XcpT-XcpW) (Nunn and Lory, 1992; Strom *et al.*, 1993) that are part of the type II export machinery. The XcpTUVW proteins are referred to as pseudopilins, *i.e.* proteins homologous to the major subunit of type IV pili in the amino-terminal moiety (Filloux *et al.*, 1998), and they are involved in type IV pilus biogenesis, probably as components of the type II secretion apparatus (Filloux *et al.*, 1998). This suggests that these pseudopilin subunits may form a pilus-like fiber similar to that of type IV pili.

In addition to the central requirement for PilD and the sequence similarities of the pilin and pseudopilin proteins, the similarity of the two systems also extends to the presence of at least three additional sets of proteins whose homologues can be seen in both: PilB and XcpR, PilC and XcpS, and PilQ and XcpQ. PilB and XcpR belong to a family of proteins characterized by their nucleotide-binding motif and peripheral membrane localization, and it is likely that these proteins play a role in providing energy for the assembly and function of the type IV pilus and type II export apparatus (Turner *et al.*, 1993; Possot and Pugsley, 1994). PilC and XcpS are integral cytoplasmic membrane proteins (Thomas *et al.*, 1997) and it has been postulated that this component interacts at the cytoplasmic face of the inner membrane with the corresponding nucleotide-binding PilB homologue (XcpR) (Bally *et al.*, 1992). Of all the components of the type II and type IV pilus biogenesis systems, only one protein in each

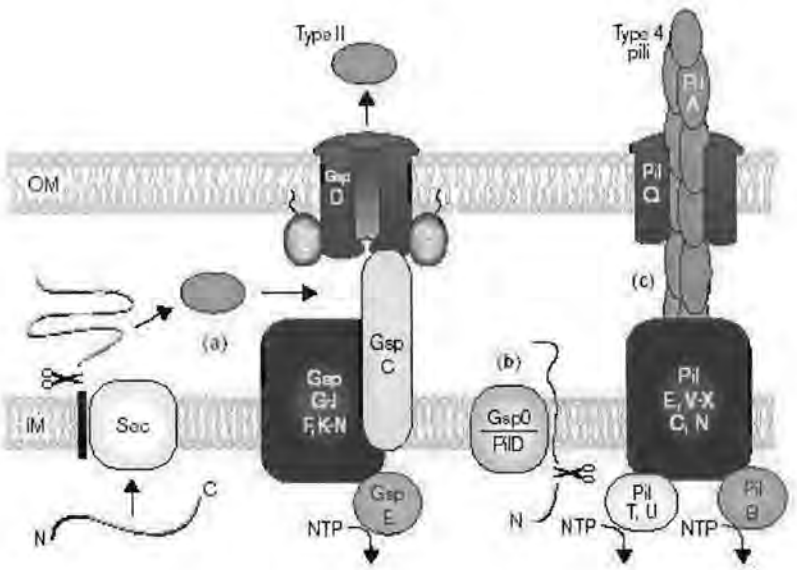

Fig. 1.4 Models for type II secretion and type IV pilus biogenesis. Components of the type II secretion are indicated using the general secretory pathway (Gsp) nomenclature, and type IV pilus proteins are labeled according to the *P. aeruginosa* Pil system. Similar shading and location indicates homologous components. (a) Type II substrates cross the inner membrane (IM) via the Sec system followed by signal-sequence cleavage and protein folding in the periplasm. The GspD secretin, indicated as a complex with the GspS lipoprotein, serves as a gated channel for secretion of substrates to the cell surface. GspC may transmit energy from the inner membrane, presumably generated by the cytoplasmic GspE nucleotide-binding protein, to the outer membrane (OM) complex. GspG-J exhibit homology to the pilin subunit PilA and are processed by the GspO prepilin peptidase. (b) GspO cleaves their amino-terminal leader sequence on the cytoplasmic face of the inner membrane. In *P. aeruginosa*, the type IV prepilin peptidase PilD is the same protein as GspO. (c) Type IV pilus biogenesis requires the outer membrane secretin PilQ. The type IV pilus system contains at least four pilin-like components (PilE and PilV-X), in addition to PilA, which are processed by PilD. Additional nucleotide-binding proteins (PilT, PilU) are present and are involved in a pilus-generated movement termed twitching motility. (Reprinted from Current Opinions in Cell Biology, Vol. 12, D.G. Thanassi and S.J. Hultgren, Multiple pathways allow protein secretion across the bacterial outer membrane, pp. 420-430, 2000, from Elsevier).

system is localized within the outer membrane. PilQ and its homologue in the type II export apparatus, XcpQ, form multimers of 12-14 subunits (Tonjum and Koomey, 1997; Bitter *et al.*, 1998) that suggest that the protein is able to form a large channel in the outer membrane.

1.5 NOVEL FIMBRIAL EXPRESSION AND ASSEMBLY SYSTEMS

1.5.1 The Flp fimbriae of *Actinobacillus actinomycetemcomitans*

Actinobacillus actinomycetemcomitans is a Gram-negative coccobacillus and has been implicated as a primary etiological agent in localized juvenile and severe adult periodontitis (Slots *et al.*, 1980; Slots and Ting, 1999), as well as other human diseases including infective endocarditic and brain abscesses (Das *et al.*, 1997; Fives-Taylor *et al.*, 2000). *A. actinomycetemcomitans* produces several potential virulence factors and toxins, including a leucotoxin, cytolethal distending toxin, trypsin-like protease and iron- and haemin-binding proteins (Fives-Taylor *et al.*, 2000). Fresh clinical isolates are able to adhere tightly to solid abiotic surfaces, *e.g.* glass, plastic and hydroxyapatite, and to form a tenacious biofilm (Fine *et al.*, 1999a; Kachlany *et al.*, 2000). Electron microscopy has revealed that these clinical isolates express long, thick fibrils (Holt *et al.*, 1980; Inouye *et al.*, 1990). Each fibril consists of a parallel array of individual pili of approximately 5 to 7 nm in diameter (Inouye *et al.*, 1990; Kachlany *et al.*, 2001), and the fibrils become interlocked by sharing individual pili. Fibrils are often several microns long and up to 100 nm thick (Kachlany *et al.*, 2001). On solid medium, colonies of fresh clinical isolates exhibit a rough colony morphology (Inouye *et al.*, 1990; Fine *et al.*, 1999a). The rough colonies can, however, convert to a smooth phenotype upon subculture and this conversion is accompanied by loss of fimbriae and adhesiveness (Inouye *et al.*, 1990; Fine *et al.*, 1999b).

Biochemical analysis has indicated that the pili of *A. actinomycetemcomitans* are composed of a 6.5-kDa polypeptide, which has been designated Flp (for fimbrial low molecular weight protein) (Inoue *et al.*, 1998). Subsequent amino acid sequence analysis of Flp has led to the identification of the *flp-1* gene, which predicts that the Flp-1 precursor has a signal peptide sequence at its amino-terminus that is likely to be cleaved upon export from the bacterial cell (Inoue *et al.*, 1998; Kachlany *et al.*, 2001). Immediately downstream from *flp-1* is another gene, *flp-2*, whose predicted product is 51% identical to Flp-1. Phylogenetic analysis of Flp-2

has indicated that not only is it a homologue of Flp-1 but both Flp-1 and Flp-2 belong to a distinct subfamily of the type IV pilin subunits (Kachlany *et al.*, 2001). Although insertion mutagenesis studies have indicated that *flp-1* mutants fail to adhere to surfaces and do not express pili, the functional significance of the *flp-2* gene is not known, as the gene does not appear to be expressed in *A. actinomycetemcomitans* (Kachlany *et al.*, 2001).

The *flp* genes of *A. actinomycetemcomitans* are located upstream of a cluster of seven novel genes, which have been termed *tadABCDEFG* (for tight adherence) (Kachlany *et al.*, 2000). Like *flp-1* mutants, mutations in any of the *tad* genes resulted in a defect in adherence and failure to produce fibrils (Kachlany *et al.*, 2000). The *flp-1-tadA* region also contains two genes, *rcpA* and *rcpB*, which have been reported by Haase *et al.* (1999) to encode outer membrane proteins that are expressed specifically in rough, adherent bacteria but not in smooth, non-adherent variants. In addition to these genes, it also contains an open reading frame (ORF) termed *orfB*, of which the predicted amino acid sequence is similar to prepilin peptidase (Haase *et al.*, 1999), which possibly removes the leader peptide sequence from prepilin for assembly into pili. Analysis of the amino acid sequences of the proteins encoded by the *tadABCDEFG* genes revealed that only TadA is similar to proteins of known function and it was subsequently reported that the TadA protein of *A. actinomycetemcomitans* is an ATPase required for fibril production and tenacious adherence (Bhattacharjee *et al.*, 2001). Moreover, phylogenetic analysis revealed that TadA is the first representative of a distinct subfamily of potential type IV secretion NTPase genes (Planet *et al.*, 2001).

Based on the phylogenies of the *tadA* and *flp* genes (Kachlany *et al.*, 2000; Planet *et al.*, 2001), it was suggested that they have experienced an evolutionary history that is independent of other secretion systems. Consequently, it was recently proposed that the *flp-rcp-tad* region specifies a novel system for the assembly and secretion of Flp pili that is distinct from the type II and type IV secretion systems (Bhattacharjee *et al.*, 2001). Interestingly, similar *tad* loci are present in the genomes of a wide variety of Gram-negative and Gram-positive bacteria and in Archaea (Kachlany *et al.*, 2000; Planet *et al.*, 2001), and it has been proposed that the *tad* loci may be important for microbial colonization in a variety of environmental niches.

1.6 AIMS OF THIS STUDY

From the review of the literature, it is apparent that microbial adhesion to biotic and abiotic surfaces is mediated by non-specific physical or specific ligand-receptor interactions whereafter growth and survival of the microbial community or biofilm depends on adaptation to a series of changing environmental milieus. Advances in genetic analysis of biofilm formation by microorganisms have resulted in substantial progress in the understanding of the molecular mechanisms involved in this process (Costerton *et al.*, 1995; Davey and O'Toole, 2000; O'Toole *et al.*, 2000). With reference to the early events in biofilm formation, several studies have suggested that flagella, fimbriae and other protein receptors are essential for bacterial attachment to surfaces. In all instances, the single or diverse components of these surface structures, however, must cross the double-layered envelope of Gram-negative bacteria before reaching their final destination outside the cell, and to this end, diverse secretory pathways have evolved. In *P. aeruginosa*, the involvement of two surface appendages has clearly been established (O'Toole and Kolter, 1998a; 1998b). These structures include the primary organelles of motility; the flagellum and type IV pili. However, the involvement of pili other than type IV pili in the attachment of *P. aeruginosa* to surfaces has not yet been considered for this organism. Recently, a novel pilus biogenesis/secretion system has been identified in *A. actinomycetemcomitans*, which has subsequently been reported to occur in the genomes of a wide variety of Gram-negative and Gram-positive bacteria. It has been proposed that this system may play a role in microbial colonization in a variety of environmental niches (Kachlany *et al.*, 2000; 2001).

Therefore, the aims of this investigation were the following:

- To analyse the available genome sequence of *P. aeruginosa* PAO1 in order to identify homologues of genes previously described to comprise a novel pilus biogenesis system in various different bacteria, amongst other *A. actinomycetemcomitans*.

- To generate *P. aeruginosa* mutant strains by directed insertional mutagenesis using allelic exchange vectors harboring inactivated copies of selected genes.

- To construct a complementation plasmid whereby the generated *P. aeruginosa* mutant strains can be complemented with wild-type copies of the insertionally inactivated genes.

- To characterize the cellular morphology and growth characteristics of the *P. aeruginosa* wild-type and mutant strains.

- To compare the capacity of *P. aeruginosa* wild-type and mutant strains to form biofilms using glass wool as attachment substratum.

- To analyse the extracellular protein profiles of *P. aeruginosa* wild-type and mutant strains by two-dimensional gel electrophoresis and amino-terminal amino acid sequence analysis of selected protein spots.

CHAPTER 2

IN SILICO IDENTIFICATION OF A NOVEL PUTATIVE PILUS BIOGENESIS SYSTEM IN *Pseudomonas aeruginosa*

2.1 INTRODUCTION

Although *Pseudomonas aeruginosa* is an ubiquitous environmental Gram-negative bacterium, it is also an opportunistic human pathogen that is responsible for serious damage to the respiratory tract of cystic fibrosis patients (May *et al.*, 1991; Govan and Deretic, 1996). The success of *P. aeruginosa* to grow and cause infections in diverse environmental niches is attributed to its broad metabolic diversity and its many cell-associated and secreted virulence factors (Lazdunski *et al.*, 1990; Van Delden and Iglewski, 1998). Furthermore, the ability of *P. aeruginosa* to attach to both biotic and abiotic surfaces with the subsequent development of biofilms can be considered a major virulence trait in a variety of infections (Watnick and Kolter, 2000; Donlan, 2002). The ability of bacteria to form biofilms may also contribute to their survival in environmental niches, since the cells have access to the nutrients absorbed by the surface and may be protected from exogenous antibiotics and competitive colonization by other species (Lawrence *et al.*, 1991; Elkins *et al.*, 1999). In *P. aeruginosa*, both flagellar motility and type IV pilus-based twitching motility have been proposed to play an important role during the initial attachment of the cells to surfaces (DeWeger *et al.*, 1987; DeFlaun *et al.*, 1994; O'Toole and Kolter, 1998a; 1998b).

Actinobacillus actinomycetemcomitans is a Gram-negative coccobacillus responsible for localized juvenile periodontitis and other systemic infections (Fives-Taylor *et al.*, 2000). Similar to *P. aeruginosa*, the bacterium produces several potential virulence factors (Fives-Taylor *et al.*, 1999; Graber *et al.*, 1998) and fresh clinical isolates are able to adhere tightly to a variety of abiotic surfaces to form a tenacious biofilm (Fine *et al.*, 1999b; Kachlany *et al.*, 2000). *A. actinomycetemcomitans* produces long fibrils of bundled pili that are required for adherence (Haase *et al.*, 1999; Kachlany *et al.*, 2000; 2001). Recently, Kachlany *et al.* (2000) identified a cluster of seven novel *tad* genes (*tadABCDEFG*) required for tight nonspecific adherence of *A. actinomycetemcomitans* to surfaces. Non-polar mutations in any of the *tad* genes resulted in a failure of the cells to adhere to surfaces, they were unable to auto-aggregate and showed no evidence of pili or fibrils (Kachlany *et al.*, 2000). Subsequent investigations have indicated that several genes and open reading frames (ORFs) that are located upstream from the *tad* genes also appear to form part of the *tad* operon (Haase *et al.*, 1999; Kachlany *et al.*, 2001). Cells containing mutations in a *flp-1* gene, which is located upstream from the *tad* genes and encodes the major subunit of the pili, fail to adhere to glass and do not express pili (Kachlany *et al.*, 2001). It was subsequently concluded that the entire

flp-rcp-tad region specifies a novel system for the assembly and secretion of Flp pili, which in turn, mediate tight adherence of *A. actinomycetemcomitans* to surfaces.

Although similar *flp-rcp-tad* gene clusters have been identified in Bacteria and Archaea (Planet *et al.*, 2001), little is known about the functions of the *tad* loci in other organisms. A related *pilA-cpa* locus in *Caulobacter crescentus* was found to be responsible for the production of pili of unknown function, although it appears that the pili are used as receptors by bacteriophage ΦCbK (Skerker and Shapiro, 2000). The *C. crescentus* pilus-encoding region contains a pilin gene, *pilA*, which is a member of the *flp* superfamily, as well as several homologues of genes present in the *flp-rcp-tad* locus of *A. actinomycetemcomitans* (Skerker and Shapiro, 2000). These results provide strong supporting evidence that the *tad* loci in other bacteria may be involved in the assembly and secretion of novel pili. It would appear that *tad* loci, homologous to those described above, might, in addition to playing a role in colonization, also be important for disease. Recently, a 15-gene cluster has been identified in *Haemophilus ducreyi* that encodes predicted protein products with significant homology to those encoded by the *A. actinomycetemcomitans flp-rcp-tad* locus (Nika *et al.*, 2002). *H. ducreyi* is the etiological agent of the sexually transmitted disease chancroid (Trees and Morse, 1995). Mutations within the *flp-1* and *flp-2* genes of *H. ducreyi* were found to significantly reduce the ability of the bacteria to attach and to form microcolonies when cultured *in vitro* with human foreskin fibroblasts. Furthermore, a mutant *H. ducreyi* strain with an inactivated *tadA* gene exhibited a decrease in virulence in a rabbit model for experimental chancroid (Nika *et al.*, 2002).

The aim of this part of the investigation was therefore to determine whether *P. aeruginosa* possesses a pilus biogenesis/secretion system similar to those described above. This was investigated by searching for homologues of the above-mentioned Tad- and Cpa-encoding genes in the *P. aeruginosa* genome sequence, followed by sequence analysis of the putative *P. aeruginosa* proteins to identify conserved features that may be related to their function.

2.2 MATERIALS AND METHODS

2.2.1 Computer analysis

Nucleotide sequences of *P. aeruginosa* potentially coding for proteins similar to the *flp*, *rcp* and *tad* gene products of *A. actinomycetemcomitans* or the *pilA* and *cpa* gene products of *C. crescentus* were identified using the BLAST alignment program to search the *P. aeruginosa* PAO1 genome database (http://www.pseudomonas.com). Deduced amino acid sequences of the putative *P. aeruginosa* proteins were then compared to the entries of both the GenBank and the Unfinished Genome Databases by making use of the BLASTP program (Altschul *et al.*, 1997) available on the National Centre for Biotechnology Information web page (http://www.ncbi.nlm.nih.gov/BLAST/). Default settings of the BLAST programmes were used and the names of previously identified sequences obtained from these searches were retained in this study. Pair-wise alignments were performed using LALIGN (Pearson *et al.*, 1997), while multiple alignments were carried out using CLUSTALW (Thompson *et al.*, 1994).

The deduced amino acid sequences were also subjected to an online PROSITE database search (at http://ca.expasy.org/prosite) and conserved motifs or domains were predicted using the SMART tool (at http://smart.embl-heidelberg.de). Signal peptide prediction was performed using SIGNALP (at http://www.cbs.dtu.dk/services/SignalP) and the cellular location of proteins was predicted using PSORT (at http://psort.ims.u-tokyo.ac.jp). Membrane-spanning regions were predicted using TMPRED (at http://www.ch.embnet.org/software/TMPRED).

The isoelectric point (pI) and molecular mass (Mw) of the different proteins were determined using the pI/Mw tool (at http://ca.expasy.org/tools/pi_tool.html). Searches for consensus promoter sequences were performed using the NEURAL NETWORK PROMOTER PREDICTION PROGRAM (at http://www-hgc-lbl.gov/projects/promoter.html) as well as the SEQSCAN program (at http://www.bmb.psu.edu/seqscan).

2.3 RESULTS

2.3.1 Sequence analysis

2.3.1.1 Identification of a gene cluster in *P. aeruginosa* homologous to the *flp-rcp-tad* and *pilA-cpa* gene clusters

Haase *et al.* (1999) and Kachlany *et al.* (2000; 2001) described the existence of a 14-gene cluster (*flp-rcp-tad*) in the genome of *A. actinomycetemcomitans* that encodes proteins involved in the ability of the organism to form surface fibrils and adhere tightly to glass. Regions showing sequence homologies with the genes described in *A. actinomycetemcomitans* have also been reported in the genomes of *C. crescentus* (*pilA-cpa*) and *H. ducreyi* (*flp-rcp-tad*). These genes have been shown to be required for pili formation (Skerker and Shapiro, 2000; Nika *et al.*, 2002). Towards understanding the molecular mechanisms underlying attachment of *P. aeruginosa* to surfaces, and since pili have been reported to play an important role during the initial stages of attachment (O'Toole and Kolter, 1998a), the first step was to identify whether genes similar to those described above are indeed present in the *P. aeruginosa* genome.

BLAST searches of the *P. aeruginosa* PAO1 genome (http://www.pseudomonas.com) with each of the Tad and Cpa proteins, as well as further sequence analysis, led to the discovery of a 9.172-kb region of the *P. aeruginosa* DNA, located at nucleotides 482 1381 - 483 0553, which contained nine open reading frames (ORFs) that encoded predicted proteins with homology to some of the Tad and/or Cpa proteins. These genes (PA4297 through PA4305) appear to be organized in a single transcribed operon. Furthermore, a protein encoded by the PA4306 gene, located immediately upstream from this gene cluster, was found to display significant sequence similarity to the Flp and PilA proteins of *A. actinomycetemcomitans* and *C. crescentus*, respectively (Fig. 2.1).

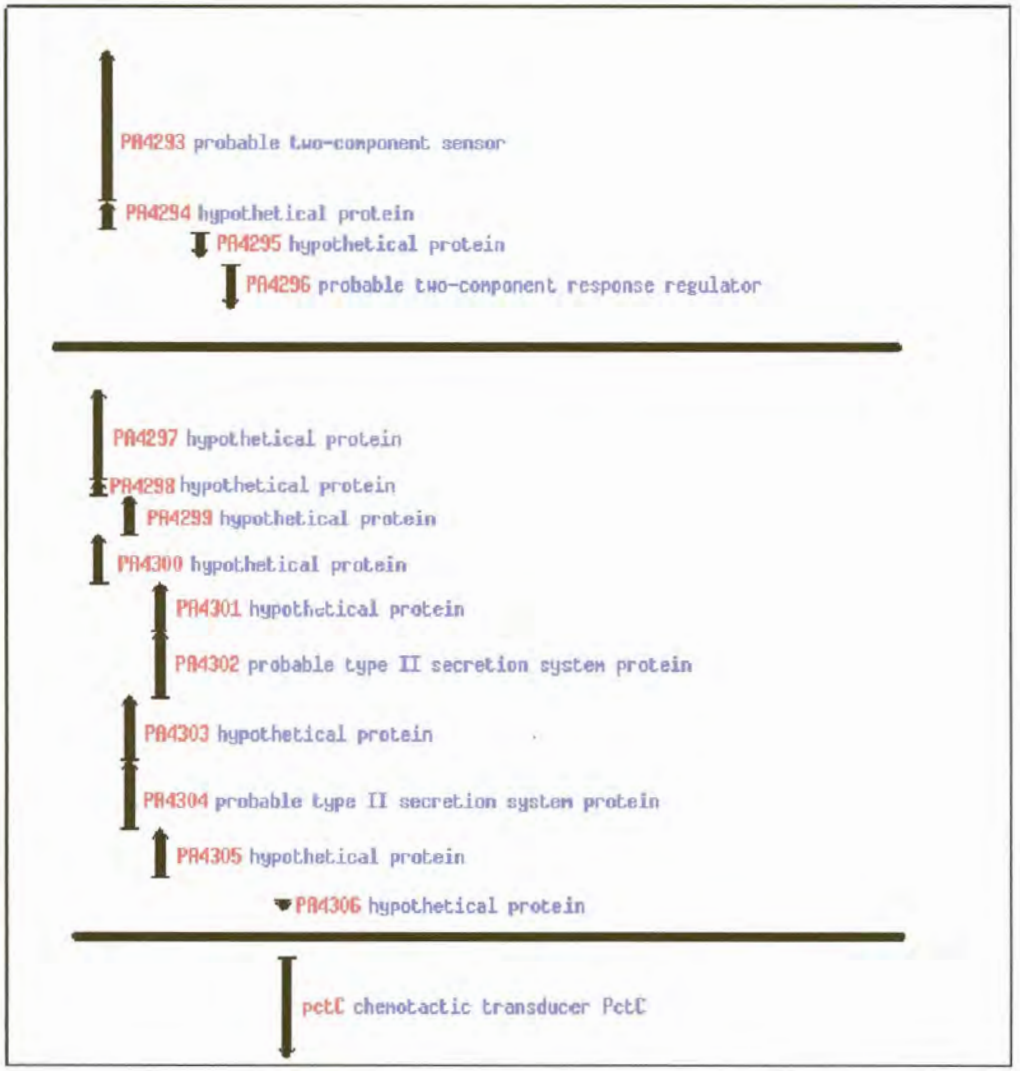

Fig. 2.1 Diagrammatic representation of the *P. aeruginosa* open reading frames (PA4306 - PA4297), which display homology to the pilus biogenesis systems of *A. actinomycetemcomitans* and *C. crescentus*, respectively. The orientation and approximate size of the different open reading frames are indicated by the direction and length of the arrows. (The diagram was obtained from the *Pseudomonas aeruginosa* genome database at http://www.pseudomonas.bit.uq.edu.au/).

2.3.1.2 Homology searches

To detect homologues of the predicted protein products encoded by the *P. aeruginosa* PA4297 through PA4305 genes, each of the putative *P. aeruginosa* proteins was compared to the protein sequences in the GenBank Database using the BLASTP alignment tool (Altschul *et al.*, 1997). The percentage identity between the closest matching sequences was calculated for full-length proteins using LALIGN (Pearson *et al.*, 1997). With the exception of the proteins encoded by the PA4297, PA4298 and PA4299 genes, each of the proteins displayed a significant degree of sequence relatedness to proteins of other putative and previously characterized adhesive and secretion systems of various bacteria, including: *Pasteurella multocida, Yersinia pestis, Agrobacterium tumefaciens, Mesorhizobium loti* and *Ralstonia solanacearum*. The results of this analysis, excluding hypothetical or putative proteins, are summarized in Table 2.1.

The closest matches to the protein encoded by the PA4300 gene are the TadC proteins of *P. multocida*, *A. actinomycetemcomitans*, *Chlorobium tepidum* and *H. ducreyi* (20-24% sequence identity). A similar level of homology was observed with the CtpI protein of *A. tumefaciens* (23% sequence identity), followed by the CpaA protein of *C. crescentus* (19% sequence identity). The protein encoded by the PA4301 gene displays a higher level of homology to the TadB proteins of *C. tepidum* and *C. crescentus* (26% sequence identity) compared to the TadB proteins of *H. ducreyi, P. multocida* and *A. actinomycetemcomitans* or the CtpH protein of *A. tumefaciens* (21-22% sequence identity). In the case of *A. actinomycetemcomitans*, both the TadB and TadC proteins have been implicated in the tight adherence of the bacterium to surfaces (Kachlany *et al.*, 2000).

The PA4302 gene encodes a protein which was found to display a significant degree of homology to the TadA proteins of *A. actinomycetemcomitans*, *P. multocida* and *H. ducreyi* (40% sequence identity), the CpaF protein of *C. crescentus* (37% sequence identity) and the CtpE protein of *A. tumefaciens* (35% sequence identity), as well as to various members of the TrbB/VirBII family of proteins. This family of proteins has been reported to be involved in DNA uptake, extracellular secretion and pilus assembly (Hobbs and Mattick, 1993; Christie, 1997).

The proteins encoded by the PA4303 and PA4305 genes displayed homology to a number of putative pilus assembly proteins. The closest match to the protein encoded by PA4303 is the CpaE protein of *C. crescentus* (20% sequence identity), followed by the CtpF protein of *A. tumefaciens* (18% sequence identity). The closest matches to the protein encoded by PA4305 are the CpaB protein (25% sequence identity) and the CtpC protein (23% sequence identity) of the same two organisms. Notably, both the CpaB and CpaE proteins have been reported to be involved in pili biogenesis of *C. crescentus* (Skerker and Shapiro, 2000).

The protein encoded by the PA4304 gene shares homology with the RcpA proteins of *A. actinomycetemcomitans*, *P. multocida*, *C. tepidum* and *H. ducreyi* (25-29% sequence identity), the CpaC protein of *C. crescentus* (27% sequence identity) and the CtpD protein of *A. tumefaciens* (25% sequence identity). Whereas the CpaC protein has been reported to function in pili formation (Skerker and Shapiro, 2000), a lack of expression of the RcpA protein in *A. actinomycetemcomitans* is associated with little or no expression of fimbriae (Haase *et al.*, 1999).

In addition, a protein encoded by the PA4306 gene displays significant sequence similarity to several pilin subunit proteins. The closest matches to the protein encoded by the PA4306 gene are the PilA protein of *C. crescentus*, the CtpA protein of *A. tumefaciens* and the Flp-1 protein of *P. multocida* (35-36% sequence identity), followed by the Flp-1 and Flp-2 proteins of *A. actinomycetemcomitans* (31% and 27% sequence identity, respectively). The *P. aeruginosa* protein also exhibits 21%, 22% and 28% sequence identity to the *H. ducreyi* Flp-1, Flp-2 and Flp-3 proteins, respectively.

2.3.2 *In silico* analysis of the *P. aeruginosa* PA4305 - PA4297 gene cluster

2.3.2.1 Identification of putative consensus promoter sequences

To identify consensus promoter sequences, the genomic sequence of *P. aeruginosa* encompassing the PA4305 to PA4297 genes, as well as the intergenic region of 410 bp between the PA4306 and PA4305 genes were analyzed using the NEURAL NETWORK PROMOTER PREDICTION and SEQSCAN programmes.

Whereas the arrangement and orientation of the *flp-rcp-tad* ORFs in *A. actinomycetemcomitans*, *H. ducreyi* and *P. multocida* are similar, they differ only in the number of *flp* alleles. By contrast to *A. actinomycetemcomitans* and *P. multocida*, which both possesses two *flp* genes, *H. ducreyi* possesses three *flp* genes. Although the orientation of the ORFs of the *pilA-cpa* system in *C. crescentus* and the *ctp* system of *A. tumefaciens* is similar to those of the *flp-rcp-tad* system, the ORFs are arranged differently and consist of fewer genes compared to the number of genes contained in the *flp-rcp-tad* gene clusters (at least 13). Despite the *P. aeruginosa* PAO1 gene cluster being similar to the *flp-rcp-tad* and *pilA-cpa* gene clusters, it has several unique features:

(i) The *P. aeruginosa* gene cluster lacks a significant portion of the *tad* gene cluster. Comparative sequence analysis of this region revealed that there are ORFs in *P. aeruginosa* PAO1 that encode proteins with homology to the Flp, TadA, TadB, TadC and RcpA proteins only. However, comparison of the *P. aeruginosa* gene cluster to the *pilA-cpa* gene cluster of *C. crescentus* led to the identification of homologues for the *pilA* and each of the *cpa* genes, except for *cpaD*. Similarly, homologues for each of the *A. tumefaciens ctp* genes, except *ctpE*, could be identified in the *P. aeruginosa* gene cluster.

(ii) The *P. aeruginosa* gene cluster contains three ORFs that encode proteins for which no similarity to any of the previously reported Tad or Cpa proteins could be demonstrated. However, BLASTP analysis indicated that the protein encoded by PA4299 displays homology to a probable lipoprotein of *Ralstonia solanacearum*, while the proteins encoded by PA4298 and PA4297, respectively, exhibit homology to a probable signal peptide protein and a probable transmembrane protein of the same organism (GenBank accession no. AL646082).

(iii) Although a homologue of the *flp* prepilin gene could be identified in the *P. aeruginosa* genomic sequence (PA4306), it does not appear to form part of the putative operon structure, but rather appears to be transcribed in the opposite direction compared to the PA4305 - PA4297 gene cluster. Moreover, by contrast to the systems of all the other bacteria, except *C. crescentus* and *A. tumefaciens*, *P. aeruginosa* contains only a single copy of the putative prepilin gene.

(iv) For pili to be assembled, a peptidase is required that process the signal peptide found in the prepilin (Christie, 1997; Soto and Hultgren, 1999). In the *flp-rcp-tad* system of *A. actinomycetemcomitans*, an ORF downstream of *flp-2*, *orfB*, encodes a protein that is similar to a prepilin peptidase (Haase *et al.*, 1999). Likewise, the CpaA protein may be a functional peptidase required for processing of the *C. crescentus* PilA subunit (Skerker and Shapiro, 2000). Using pair-wise alignments, the protein encoded by PA4300 of *P. aeruginosa* displays 19% sequence identity to the CpaA protein of *C. crescentus*, but a lower level of identity to OrfB is observed (16% sequence identity). Notably, downstream of the *P. aeruginosa* gene cluster is a gene, PA4295, that encodes a predicted protein that displays 19% sequence identity to CpaA of *C. crescentus*, and 23% sequence identity to the OrfB protein of *A. actinomycetemcomitans*. However, only the PA4295-encoded protein contains putative active sites of prepilin petidases (Skerker and Shapiro, 2000) (Fig. 2.4), but the PA4295 gene does not form part of the putative operon structure and appears to be divergently transcribed (Fig. 2.1). Further experimental analyses are required to determine which of the two proteins or whether both of these proteins function as a peptidase.

In conclusion, the sequence similarities and the order of the homologous genes in *P. aeruginosa*, with the exception of the orientation of the *flp* gene, more closely resembles that of the *pilA-cpa* system of *C. crescentus* and the *ctp* system of *A. tumefaciens* than the *flp-rcp-tad* systems of *A. actinomycetemcomitans*, *P. multocida* and *H. ducreyi*. Based on the above results, the *P. aeruginosa* gene cluster has been termed *htp* based on their **h**omology to **t**ype IV **p**ilus biogenesis proteins encoded by the *tad* and *cpa* genes of other bacteria. In the sections that follow, the PA4305 through PA4297 genes will be referred to as *htpABCDEFGHI*, and the upstream PA4306 gene, encoding a putative pilus monomer, will be referred to as *htpP* (**h**omologous to **t**ype IV **p**repillin protein).

2.3.3 Properties of the putative Htp proteins of *P. aeruginosa*

The predicted protein products encoded by the *P. aeruginosa htp* ORFs were subsequently analyzed, using various different programmes, to gain a better understanding regarding their possible function(s). The PROSITE (available at the ExPASy server) and SMART (available at the EMBL server) programmes were used to identify conserved motifs and functional

domains, while analysis of the proteins by TMPRED (Hofmann and Stoffel, 1993), SIGNALP (Nielsen *et al.*, 1997) and PSORT (Nakai and Kanehisa, 1991) were used to identify putative transmembrane-spanning regions, signal peptide sequences and the cellular location of the proteins, respectively. The results of these analyses are summarized in Table 2.1.

2.3.3.1 Proteins predicted to be localized to the outer membrane

The proteins encoded by genes PA4304 and PA4299 are both predicted to be localized to the outer membrane. Analysis of the amino acid sequence of the protein encoded by PA4304 (HtpB) indicated that the protein consists of 416 amino acid residues and has an estimated molecular mass of 44.2 kDa. The HtpB protein is predicted to have four transmembrane-spanning helices and the first 28 amino acids of the protein possess characteristics of a signal peptide, with the cleavage site predicted to be between Ala_{27} and Lys_{28} in the sequence $_{22}$LGVALA↓LPALAL$_{33}$. Cleavage at this site would yield a mature polypeptide with a theoretical molecular mass of 42.0 kDa. Notably, the protein possesses a secretion protein motif (GGX$_{12}$VP[L/F]LXXIPXIGXL[F/L]), located near the carboxy-terminus of the protein, resembling that of the bacterial type II secretion system protein D (GspD). Protein D is involved in the general (type II) secretion pathway within Gram-negative bacteria and, being located in the outer membrane, is thought to be involved in transporting exoproteins from the periplasm, across the outer membrane, to the extracellular environment (Hobbs and Mattick, 1993). The notion that protein HtpB may function as a secretin is furthermore supported by BLASTP analysis (Section 2.3.1.2), which indicated that the protein shares significant sequence homology to members of the pulD/pIV family of outer membrane proteins (Fig. 2.5). These proteins, also referred to as secretins, are involved in extracellular secretion and pilus biogenesis (Russel, 1998; Nouwen *et al.*, 1999). They are believed to function as a gated channel in the outer membrane through which protein substrates are secreted (Russel *et al.*, 1997; Guilvout *et al.*, 1999; Nouwen *et al.*, 1999).

Analysis of the amino acid sequence of the protein encoded by PA4299 (HtpG) indicated that the protein consists of 245 amino acid residues and has an estimated molecular mass of 26.5 kDa. Although the protein is being discussed here, it should be noted that the protein is also predicted to be localized to the inner membrane with an equal probability. The protein is predicted to have one transmembrane-spanning helix and appears to be a lipoprotein. Evidence for the lipoprotein nature of the protein is derived from it possessing a consensus

prokaryotic membrane lipoprotein lipid attachment site ($_{7}$IGLCAALLGGC$_{17}$). This site resembles the consensus ([LY]-[AST]-[GA]↓C) of the lipoprotein signal sequence cleavage site (Sutcliffe and Russell, 1995). In the case of HtpG, the cleavage site is predicted to be between Gly_{16} and Cys_{17} in the above sequence and would thus result in a mature polypeptide having a theoretical molecular mass of 25.1 kDa. In prokaryotes, membrane lipoproteins are cleaved by a specific lipoprotein signal peptidase (signal peptidase II) that recognizes the conserved sequence and cleaves upstream of the Cys residue to which a glyceride-fatty acid lipid is then attached (Hayashi and Wu, 1990; Paetzel *et al.*, 2000). Such lipoprotein modification provides a versatile mechanism by which proteins may be anchored within the cell envelope (Sutcliffe and Russell, 1995). The notion that HtpG is indeed a lipoprotein is also supported by the results of BLASTP analysis that indicated that the protein shows a homology to a probable transmembrane lipoprotein of *Ralstonia solanacearum* (27% sequence similarity).

2.3.3.2 Proteins predicted to be localized to the periplasmic space

Only two of the proteins, HtpA, encoded by the PA4305 gene, and HtpH, encoded by the PA4298 gene, are predicted to be localized to the periplasmic space. Analysis of the amino acid sequence indicated that the HtpA protein consists of 303 amino acid residues with a predicted molecular mass of 31.7 kDa. The protein is also predicted to possess two transmembrane-spanning helices, but has a low probability of occurring in the outer membrane. The first 37 amino acids in the sequence possess characteristics of a prokaryotic signal peptide, with the cleavage site predicted to be between Ala_{36} and Pro_{37} in the sequence $_{32}$SNAHA↓PSVAPA$_{42}$. Cleavage at this site would yield a mature polypeptide with a theoretical molecular mass of 28.2 kDa.

The protein encoded by the PA4298 gene (HtpH) is composed of 94 amino acid residues and has a predicted molecular mass of 10.3 kDa. The protein is predicted to possess one transmembrane-spanning region and the first 20 amino acids of the protein possess characteristics of a signal peptide, with the cleavage site predicted to be between Ala_{19} and Ala_{20} in the sequence $_{14}$SGTAWA↓ADTPAV$_{25}$ to yield a mature polypeptide of 8.3 kDa. The protein was found by BLAST analysis to only display sequence homology to a probable signal peptide protein (27% sequence similarity) of *R. solanacearum*.

2.3.3.3 Proteins predicted to be localized to the inner membrane

Most of the Htp proteins are predicted to be localized to the inner membrane. These include proteins HtpC (394 amino acid residues, 42.4 kDa), HtpE (294 amino acid residues, 32.4 kDa), HtpF (303 amino acid residues, 33.6 kDa) and HtpI (556 amino acid residues, 56.4 kDa). The HtpC protein, encoded by the PA4303 gene, is predicted to possess a single transmembrane-spanning helix, but lacks an apparent signal peptide sequence. Proteins HtpE and HtpF, encoded by the PA4301 and PA4300 genes, respectively, are predicted to each possess four transmembrane-spanning helices, while protein HtpI, encoded by the PA4297 gene, is predicted to possess six transmembrane-spanning helices. Whereas the first 18 amino acids of the HtpE sequence possess characteristics of a signal peptide, with the cleavage site predicted to be between Ala_{17} and Ala_{18} in the sequence $_{12}$GVLAFA↓ALALAF$_{23}$, the first 118 amino acids of the HtpI sequence possess characteristics of a signal peptide, with the cleavage site predicted to be between Ala_{117} and Asn_{118} in the sequence $_{111}$SNAAVA↓NEAVHV$_{124}$. Cleavage at these sites would yield mature proteins of 30.7 and 44.8 kDa, respectively. By contrast, the HtpF protein may possess an N-terminal signal peptide with the cleavage site predicted to between Ala_{294} and Ile_{295} in the sequence $_{290}$PGVIA↓ITKALG$_{300}$, thus yielding a mature protein of 32.8 kDa. Both the HtpF and HtpI proteins possess several leucine zipper motifs (L-X_6-L-X_6-L-X_6-L) (Busch and Sassone-Corsi, 1990). Protein segments that contain leucine zipper motifs appear to exist in an alpha-helical conformation. The leucine side chains extending from one alpha-helix interact with those from a similar alpha-helix of a second polypeptide, facilitating dimerization of the proteins (O'Shea *et al.*, 1989). Thus, it is tempting to speculate that the HtpF and HtpI proteins may either function as homodimers or heterodimers.

2.3.3.4 Proteins predicted to be localized to the cytoplasm

The HtpD protein, encoded by the PA4302 gene, consists of 421 amino acid residues and has a predicted molecular mass of 46.2 kDa. The protein lacks apparent transmembrane-spanning helices and a signal peptide, thus supporting a probable cytoplasmic localization of the protein. Analysis of the sequence indicated that it contains a consensus motif for an ATP/GTP binding site ([AG]-X_4-G-K[ST]) which is also referred to as the Walker box A consensus sequence (Walker *et al.*, 1982) or the "P-loop" (Seraste *et al.*, 1990). The glycine residues that are present in this motif, are responsible for forming a flexible loop which then

interacts with one of the phosphate groups of the nucleotide (Moller and Amons, 1985; Seraste *et al.*, 1990). Notably, BLASTP searches revealed that the HtpD protein displays significant sequence identity to the TadA proteins of *H. ducreyi*, *P. multocida* and *A. actinomycetemcomitans* as well as the CpaF protein of *C. crescentus* and the CtpG protein of *A. tumefaciens*. These proteins all belong to the TrbB/VirBII family of proteins (Section 2.3.1.2). The Trb/VirBII family of proteins is involved in DNA uptake, extracellular secretion and pilus assembly (Hobbs and Mattick, 1993; Christie, 1997). All of these proteins contain a Walker binding motif (Walker *et al.*, 1982), as does HtpD (Fig. 2.6), suggesting that nucleotide binding or hydrolysis is critical to their function (Turner *et al.*, 1993). Although the biochemical function of the TrbB/VirBII protein family is not understood, it is believed that energy provided by ATP hydrolysis is used to provide energy for the export of proteins across the inner membrane (Christie, 1997; Russell, 1998). It can therefore be proposed that HtpD represents a NTPase of the putative *P. aeruginosa htp* pilus assembly/secretion system.

2.3.3.5 Characterization of the putative prepilin, HtpP

Analysis of the protein encoded by the PA4306 gene indicated that it consists of 72 amino acid residues and has an estimated molecular mass of 7.3 kDa. The first 22 amino acids of the protein possess characteristics of a signal peptide sequence, with the cleavage site between Gly_{21} and Ala_{22} in the sequence $_{17}$ADEEG↓ANAIE_{26}, thus yielding a mature protein with a theoretical molecular mass of 4.9 kDa. In contrast to the *pilA-cpa* and *flp-rcp-tad* systems of other bacteria, the *htpP* ORF of *P. aeruginosa* PAO1 does not form part of the putative operon structure and appears to be separately transcribed (Section 2.3.2.2). Comparison of the HtpP protein sequence to other type IV prepilin proteins revealed the presence of an Flp motif at the amino-terminus of the predicted mature peptide (Fig. 2.7). The Flp motif is characterized by adjacent glutamate and tyrosine residues in its center, and it occurs within a stretch of approximately 20 hydrophobic, non-polar, aliphatic amino acids (Kachlany *et al.*, 2000). Based on these structural similarities and significant sequence homology of HtpP to other members of the type IV prepilin family of proteins, it can be proposed that HtpP is a prepilin that is secreted by the apparatus encoded by the putative *P. aeruginosa htp* operon.

Table 2.1 Characteristics of the *P. aeruginosa htp* genes and encoded gene products, listed in conjunction with homologous genes

Gene name	PA number[a]	Nearest homologue[b]	% Identity[c]	% Similarity[d]	Protein size (amino acids)[e]	Molecular mass (in kDa)[e]	Theoretical pI[e]	Predicted function[f]
htpI	**4297**				**556**	**56.4**	**5.02**	**IM, transmembrane protein**
htpH	**4298**				**94**	**10.3**	**8.14**	**Periplasmic**
htpG	**4299**				**245**	**26.5**	**6.15**	**OM / IM lipoprotein**
htpF	**4300**				**303**	**33.6**	**9.93**	**IM, transmembrane protein**
		TadC / *A. actinomycetemcomitans*	23	48 (121 - 283)	288	32.2	9.13	
		TadC / *H. ducreyi*	20	44 (114 - 276)	281	31.7	5.61	
		TadC / *P. multocida*	24	51 (112 - 279)	284	31.9	9.50	
		TadC / *C. tepidum*	21	46 (132 - 309)	312	35.2	9.81	
		CpaA / *C. crescentus*	19	42 (194 - 360)	160	40.2	9.80	
		CtpI / *A. tumefaciens*	23	42 (161 - 324)	328	36.8	9.89	
htpE	**4301**				**294**	**32.4**	**10.33**	**IM, pilus assembly transmembrane protein**
		TadB / *A. actinomycetemcomitans*	21	42 (111 - 292)	295	34.4	9.54	
		TadB / *H. ducreyi*	22	40 (113 - 292)	295	34.4	9.84	
		TadB / *P. multocida*	21	44 (102 - 288)	291	34.1	9.86	
		TadB / *C. tepidum*	26	43 (133 - 302)	305	34.1	10.27	
		TadB / *C. crescentus*	26	46 (136 - 321)	325	34.6	10.08	
		CtpH / *A. tumefaciens*	22	51(145 - 330)	334	37.0	10.07	
htpD	**4302**				**421**	**46.2**	**5.91**	**Cytoplasmic, type IV secretion NTPase**
		TadA / *A. actinomycetemcomitans*	40	63 (7 - 377)	426	47.1	5.51	
		TadA / *H. ducreyi*	39	64 (14 - 379)	427	47.5	5.39	
		TadA / *P. multocida*	40	64 (14 - 377)	425	48.1	5.22	
		CpaF / *C. crescentus*	37	64 (76 - 462)	501	54.3	6.07	
		CtpG / *A. tumefaciens*	38	64 (67 - 446)	491	53.9	5.28	

Gene name	PA number[a]	Nearest homologue[b]	% Identity[c]	% Similarity[d]	Protein size (amino acids)[e]	Molecular mass (in kDa)[e]	Theoretical pI[e]	Predicted function[f]
htpC	**4303**				**394**	**42.4**	**6.87**	**IM, pilus assembly protein**
		CpaE / *C. crescentus*	20	37 (203 - 493)	517	55.1	4.85	
		CtpF / *A. tumefaciens*	18	30 (124 - 231)	427	46.6	5.05	
		ATPase / *C. tepidum*	16	41 (34 - 285)	292	32.5	5.84	
htpB	**4304**				**416**	**44.2**	**5.34**	**OM, channel protein (secretin)**
		RcpA / *A. actinomycetemcomitans*	28	56 (249 - 429)	460	50.2	5.65	
		RcpA / *H. ducreyi*	25	59 (249 - 411)	456	49.8	5.99	
		RcpA / *P. multocida*	27	59 (259 - 425)	470	51.1	6.21	
		RcpA / *C. tepidum*	29	46 (134 - 437)	445	46.9	6.87	
		CpaC / *C. crescentus*	27	55 (335 - 522)	560	58.1	9.17	
		CtpD / *A. tumefaciens*	25	46 (227 - 478)	527	56.8	8.44	
htpA	**4305**				**303**	**31.7**	**6.22**	**Periplasmic, pilus assembly protein**
		CpaB / *C. crescentus*	25	43 (149 - 261)	297	29.8	7.93	
		CtpC / *A. tumefaciens*	26	49 (129 - 223)	268	28.2	5.61	
htpP	**4306**				**72**	**7.3**	**4.72**	**Pilus subunit protein, secreted**
		Flp-1 / *A. actinomycetemcomitans*	31	68 (14 - 45)	76	8.1	9.6	
		Flp-2 / *A. actinomycetemcomitans*	27	63 (1 - 76)	76	8.3	7.9	
		Flp-1 / *H. ducreyi*	21	53 (23 - 79)	85	9.3	9.99	
		Flp-2 / *H. ducreyi*	22	76 (23 - 48)	81	8.9	9.87	
		Flp-3 / *H. ducreyi*	28	57 (1-89)	89	7.8	9.13	
		Flp-1 / *P multocida*	35	68 (1 - 74)	74	7.6	9.1	
		Flp-2 / *P multocida*	24	75 (1 - 60)	60	6.6	8.9	
		PilA / *C. cresentus*	35	72 (1 - 58)	59	6	9.52	
		CtpA / *A. tumefaciens*	35	76 (15 - 40)	63	6.4	9.22	

[a] The PA number corresponds to the genome annotation (http://www.pseudomonas.com)
[b] Homologues were identified using BLAST (NCBI) (Altschul *et al.*, 1997). Homologous putative and hypothetical proteins have been excluded from the analyses
[c] The percentage identity was calculated for full-length proteins using LALIGN (Pearson *et al.*, 1997)
[d] The percentage similarity is assigned to the region of the protein including the indicated amino acid residues
[e] Length of predicted proteins, molecular mass and pI were calculated for full-length proteins
[f] Predicted function of each *P. aeruginosa* protein (cytoplasmic; inner membrane (IM); periplasmic; outer membrane (OM); or secreted) was independently obtained using PSORT, SMART, SIGNALP and TMPRED

2.4 DISCUSSION

Bacteria use different secretion systems to move macromolecules across their membranes (Lory, 1998; Christie, 2001). Secretion of proteins such as proteases, pili and toxins can provide selective advantages to bacteria in various environmental niches, and many of the secreted proteins are important colonization and virulence factors released by pathogenic bacteria. Of these, pili, found on a wide variety of Gram-negative bacteria, play an important role in adhesion of pathogenic bacteria to their host (Hahn, 1997; Kang *et al.*, 1997), biofilm formation (O'Toole and Kolter, 1998b; Watnick *et al.*, 1999), twitching motility (Darzins 1993; Darzins, 1994; Wall and Kaiser, 1999), conjugative DNA transfer (Christie, 1997; Krause *et al.*, 2000; Christie, 2001) and bacteriophage infection (Rehmat and Shapiro, 1983; Roncero *et al.*, 1990; Mattick *et al.*, 1996). In Gram-negative bacteria, the pilin subunit must be secreted across both inner and outer membranes before being assembled into an extracellular filament. Although several distinct mechanisms of pilus assembly have been described (Christie, 1997; Soto and Hultgren, 1999), they all share common requirements: prepilin peptidases that process the signal peptide found on prepilin, an ATP-hydrolyzing protein that provides energy for transport of pilin across the inner membrane, and outer membrane proteins that form channels whereby the pilin subunit is translocated to the cell surface. Recently, gene clusters encoding a novel pilus assembly and secretion system have been identified in *A. actinomycetemcomitans* (Kachlany *et al.*, 2000) and *C. crescentus* (Skerker and Shapiro, 2000), and the gene order has been shown to be largely conserved in a group of diverse bacteria.

In this part of the study, open readings frames (ORFs) were identified in the genome of *P. aeruginosa* PAO1 (Stover *et al.*, 2000) with predicted products that are significantly similar to proteins involved in the formation of novel pili in *A. actinomycetemcomitans* and *C. crescentus* (Fig. 2.3 and Table 2.1). The putative *htp* gene cluster of *P. aeruginosa* bears greatest similarity to the *cpa* cluster of *C. crescentus*, both at the level of amino acid sequence and in terms of genomic organization. Whereas the *htp* cluster contains six of the seven conserved *cpa* genes, it contains only four of the 14 *tad* genes found in *A. actinomycetemcomitans* (Fig. 2.3). Given that horizontal gene transfer may play a role in the evolution of the *htp*, *tad* and *cpa* gene clusters, the *htp* gene cluster and flanking sequences (PA4306 through PA4295) were examined for the presence of insertion elements, inverted and direct repeats, and G+C content. Neither insertion sequences, nor duplicated regions were

detected (results not shown). Analysis of the G+C content of the cluster showed that the average G+C content of the nine genes spanning PA4305 through PA4297 is 68%, while the G+C content of PA4306 is 57%. These averages, as well as the average G+C content of the whole region (62.5%), are similar to the average G+C content (66.6%) of the *P. aeruginosa* PAO1 genome (Stover *et al.*, 2001). This is in contrast to *A. actinomycetemcomitans* for which the G+C content of the complete *tad* region (35%) is significantly different from the genome (48%), suggesting that the *tad* region was inserted into the genome following horizontal gene transfer from an as-yet-unidentified source (Planet *et al.*, 2001).

In silico analysis of the predicted protein products encoded by the individual ORFs in the *P. aeruginosa htp* cluster indicated several noteworthy features. The predicted product of the *htpD* (PA4302) gene, HtpD, showed significant sequence homology to the *C. crescentus* CpaF and *A. actinomycetemcomitans* TadA proteins, as well as to several members of the TrbB/VirBII family of secretion NTPases. Protein sequence alignment between HtpD and various members of this secretion NTPase family (Fig. 2.6) indicated that HtpD contains four conserved domains present in all family members. These include two canonical nucleotide-binding motifs designated as Walker boxes A and B and two conserved regions designated as the Asp and His boxes (Whitchurch *et al.*, 1991; Possot and Pugsley, 1994). Several members of the type IV family of NTPases bind and hydrolyze ATP, and mutations in the Walker A motif abolish both this activity and macromolecular secretion (Stephens *et al.*, 1995; Rivas *et al.*, 1997; Krause *et al.*, 2000). Thus, NTP-binding and/or hydrolysis very likely are essential to the function of all these proteins. Indeed, recent work has confirmed that TadA of *A. actinomycetemcomitans* hydrolyzes ATP (Battacharjee *et al.*, 2001) and a transposon insertion in the *tadA* gene has previously been reported to be defective in pilus biogenesis (Kachlany *et al.*, 2000). Furthermore, the putative NTPases of both type II and type IV secretion systems, such as the PulE protein of the type II system for pullulanase secretion (Possot and Pugsley, 1994; 1997) as well as the type IV NTPase VirBII encoded by Ti plasmids (Rashkova *et al.*, 1997), are soluble, found in the cytoplasm and have been shown to interact with other components of the secretion apparatus that are embedded in the inner membrane (Grahn *et al.*, 2000; Krause *et al.*, 2000; Rashkova *et al.* 2000). Except for HtpD, the proteins encoded by the *htp* gene cluster of *P. aeruginosa* are predicted to be integral membrane proteins (Table 2.1), and it is thus reasonable to expect that HtpD is localized to the inner membrane by interaction with one or more of these proteins. Based on the above,

the HtpD protein may be proposed to function as an energizer of the putative *P. aeruginosa htp* macromolecular secretion system.

BLASTP homology analysis indicated that the predicted protein product of the *P. aeruginosa htpB* (PA4304) ORF, HtpB, is, like both *C. crescentus* CpaC and *A. actinomycetemcomitans* RcpA, a member of the pulD/pIV family of outer membrane secretins. The HtpB protein displays similarity to the protein D (43%) of the general secretion pathway (GspD) in several Gram-negative organisms (Russel, 1998), and to PilQ (49%), an assembly protein of type IV fimbriae in *P. aeruginosa* and *N. gonorrhoeae* (Martin *et al.*, 1993; Drake and Koomey, 1995). The general secretary pathway is used to transport a variety of macromolecules across the outer membrane, including type IV fimbrae (Genin and Boucher, 1994; Hultgren *et al.*, 1996). Bacterial homologues within the GspD protein family have been proposed to function as gatekeepers by interacting with substrates bound for export through the outer membrane (Martinez *et al.*, 1998). Several of these secretins have been purified and shown to form a toroid-shaped channel, with an inner diameter of 5-10 nm (Linderoth *et al.*, 1997; Bitter *et al.*, 1998; Nouwen *et al.*, 1999). Since type IV pili have a diameter of 5-7 nm (Strom and Lory, 1993; Soto and Hultgren, 1999), the size of the channel is sufficiently large that an intact pilus filament could exit through it (Bitter *et al.*, 1998). Based on its similarity to outer membrane secretins and its predicted location in the outer membrane, the HtpB protein may thus function as a secretin in *P. aeruginosa* HtpP pili assembly.

Of the genes in the *A. actinomycetemcomitans flp-rcp-tad* gene cluster, the *flp-1* gene has been proposed to encode the major pilin subunit, Flp (Inoue *et al.*, 1998; Kachlany *et al.*, 2001). Flp is a small protein (6.5 kDa) that has a type IV-like leader sequence and the predicted mature Flp-1 protein is 40% identical to the PilA protein of *C. crescentus*. The predicted protein product encoded by the *htpP* (PA4306) ORF of *P. aeruginosa* displays significant identity to the Flp-1 (31%) and PilA (35%) proteins (Table 2.1). The HtpP protein has a leader peptide similar to that found in type IV pilin and alignment of the *P. aerugonosa* HtpP protein with other members of the Flp subfamily (Fig. 2.7) revealed the presence of an Flp motif (Kachlany *et al.*, 2001) at the amino-terminus of the predicted mature protein. This motif is characterized by adjacent glutamate and tyrosine residues in its center, and it occurs within a stretch of 20 hydrophilic non-polar, aliphatic amino acids. In addition, the predicted HtpP protein, like other members of the subfamily, contains a phenylalanine residue close to the middle of its hydrophilic carboxy-terminal domain.

The presence of a type IV-like leader peptide suggests that *P. aeruginosa* HtpP is processed by a specific prepilin peptidase. Prepilin peptidase PilD, first identified in *P. aeruginosa,* cleaves type IV prepilin after the conserved glycine residue and methylates the resulting N-terminal phenylalanine (Nunn and Lory, 1991). Although the obtained sequence for HtpP indicates that the amino-terminal residue of mature HtpP is an alanine, it is known that prepilin peptidase can methylate an alanine residue (Strom and Lory, 1992). It is, however, not known whether the *A. actinomycetemcomitans* Flp, *C. crescentus* PilA or *P. aeruginosa* HtpP proteins is methylated. Structure-function studies of *P. aeruginosa* PilD and the *V. cholerae* homologue, TcpJ, suggest that the methylase and peptidase activities reside on separate parts of the protein (Strom *et al.*, 1993; Pepe and Lory, 1998; LaPointe and Taylor, 2000). Notably, the *C. crescentus* CpaA and *A. actinomycetemcomitans* OrfB proteins, which have been proposed to be functional peptidases required for processing of the PilA and Flp prepilin subunits, are similar only to the domain required for peptidase activity and lack a conserved cytoplasmic loop implicated in methyltransferase function (Haase *et al.*, 1999; Skerker and Shapiro, 2000). In the case of the *P. aeruginosa* Htp system, no obvious candidate prepilin peptidase could be identified. Although both genes *htpF* (PA4300) and PA4295 encode proteins with homology to the CpaA and OrfB proteins (Section 2.3.2.2), only the protein encoded by PA4295 contains putative active sites of prepilin peptidases (Fig. 2.4). It is thus tempting to propose that the PA4295-encoded protein, rather than HtpF, may be the functional peptidase, but not a methylase, and could be required for the processing of the HtpP prepilin. A direct role of this protein in the processing of HtpP, however, awaits the production of an anti-HtpP pilin antibody.

In conclusion, despite several distinct mechanisms of pilus assembly having been described, they all share the common requirements of an NTP-hydrolyzing protein, a secretin protein and prepilin peptidase (Christie, 1997; Soto and Hultgren, 1999). The results of *in silico* analysis performed in this part of the study have indicated that the *P. aeruginosa* Htp system contains proteins that can be proposed to be equivalents of the required proteins. These findings therefore suggest that the *htp* locus constitutes a secretion system responsible for the export of HtpP pili, but the possibility that it may function in the secretion of other extracellular proteins cannot be excluded. The HtpA, HtpC and HtpE through HtpI proteins are unrelated proteins of unknown function, although they all contain at least one possible membrane-spanning domain and are thus predicted to be integral membrane proteins. It can thus be proposed that these Htp proteins form part of a membrane complex for the secretion

and assembly of HtpP pili. In such a model, HtpD might act as the energizing protein, while HtpB, which displays similarity to the pulD/Type IV family of secretins, may form channels in the bacterial outer membrane to allow for export of the processed pilin subunits.

CHAPTER 3

CONSTRUCTION OF *htp* MUTANT STRAINS OF *P. aeruginosa* DSM1707

3.1 INTRODUCTION

There have been numerous examples of structures that mediate attachment of bacterial cells to surfaces. Among the best characterized is the type 1 and P pili of *Escherichia coli* (Orndorff and Bloch, 1990; Pratt and Kolter, 1998) and type IV pili of *Pseudomonas aeruginosa* (Smyth *et al.*, 1996; O'Toole and Kolter, 1998b). The type IV pili have been identified in a wide spectrum of Gram-negative bacteria (Strom and Lory, 1993; Smyth *et al.*, 1996). In *P. aeruginosa*, both the flagellum and type IV pili are involved in the initiation of biofilm formation (O'Toole and Kolter, 1998a; 1998b). The type IV pili are presumably the principal adhesins mediating adherence to eukaryotic cell surfaces (Saiman *et al.*, 1990; Kang *et al.*, 1997; Hahn, 1997) as well as abiotic surfaces (O'Toole and Kolter, 1998b; Semmler *et al.*, 1999). They appear to be required for initiation of biofilm formation by promoting cell aggregation and the formation of microcolonies. Furthermore, type IV pili are also responsible for the flagellum-independent mode of surface translocation called twitching motility (Darzins and Russel, 1997; Semmler *et al.*, 1999; Wall and Kaiser, 1999).

Recently, a potentially novel class of pili was identified in *Actinobacillus actinomycetemcomitans* (Kachlany *et al.*, 2000), *Caulobacter crescentus* (Skerker and Shapiro, 2000) and *Haemophilus ducreyi* (Nika *et al.*, 2002). Fimbriae purified from *A. actinomycetemcomitans* (Inoue *et al.*, 1998) and *C. crescentus* (Skerker and Shapiro, 2000) were shown to be comprised of a major subunit with an apparent molecular mass of 5 to 6 kDa. The amino acid sequences of these small proteins display similarity to known type IV prepilin proteins, and the *A. actinomycetemcomitans* protein was designated Flp (Inoue *et al.*, 1998). The gene encoding the major fimbrial subunit proved to be part of a large gene cluster in all three organisms that may constitute a polycistronic operon encoding a novel secretion system for the assembly and release of the fimbriae (Kachlany *et al.*, 2000; Skerker and Shapiro, 2000; Nika *et al.*, 2002).

In this study, a cluster of nine genes, termed *htpABCDEFGHI*, was identified in the genomic DNA of *P. aeruginosa*. Most of these genes were found to encode predicted protein products that display homology to the proteins involved in the formation and release of novel fimbriae in other bacteria, including *A. actinomycetemcomitans*, *C. crescentus* and *H. ducreyi* (Chapter 2). Notably, a gene (*htpP*) encoding a protein that is homologous to the Flp protein of the above organisms was identified immediately upstream from the *htpA* ORF in *P. aeruginosa*.

Although the function of the *htp* genes in *P. aeruginosa* may be deduced through *in silico* predictions and by comparing these different genes, it is, however, only through the construction of mutations in *P. aeruginosa* that their actual function in this bacterium can be elucidated.

Various different strategies have been described whereby mutant bacterial strains can be generated (Maloy and Nunn, 1981; Gay *et al.*, 1985; Pelicic *et al.*, 1996; Thompson *et al.*, 1999; Espinosa-Urgel *et al.*, 2000). Of these, the use of plasmids that are conditional for their replication ("suicide plasmids") as a means of introducing defined mutations within a target genome has frequently been employed (Suh *et al.*, 1999; Dasgupta *et al.*, 2000; Nika *et al.*, 2002). In such instances, a copy of a chromosomal gene, which has been disrupted through the insertion of an antibiotic gene, is cloned into a plasmid and then introduced into a recipient strain where the plasmid cannot replicate. Since the plasmid cannot replicate, selection for some property of the plasmid, such as the newly introduced antibiotic resistance marker, results in isolates that have integrated the cloned disrupted DNA fragment into the host chromosome via homology between the DNA fragment and the corresponding region of the recipient chromosome. The construction of such mutations therefore represents a powerful approach to the definition of structure-function relationships and the identification of gene function (Hensel and Holden, 1996; Dasgupta *et al.*, 2000; Nika *et al.*, 2002).

The results obtained in Chapter 2 indicated the existence of a novel pilus biogenesis/secretion system (*htp*) in the genome of *P. aeruginosa* that bears a strong resemblance to that of the recently described *flp-rcp-tad* and *pilA-cpa* pilus biogenesis/secretion systems of *A. actinomycetemcomitans* and *C. crescentus*, respectively. Towards determining the function of the *htp* locus in *P. aeruginosa*, the aims of this part of the investigation were (i) to generate *P. aeruginosa* DSM1707 mutant strains by directed insertional mutagenesis using appropriate allelic exchange vectors, (ii) to verify the presence of the mutant alleles in the generated *P. aeruginosa* mutant strains and (iii) to construct a complementation plasmid whereby these mutations could be complemented in the *P. aeruginosa* mutant strains. As homologues of the predicted protein products encoded by the *htpD*, *htpE* and *htpF* ORFs of *P. aeruginosa* are present in both the *flp-rcp-tad* system of *A. actinomycetemcomitans* and the *pilA-cpa* system of *C. crescentus*, they were selected for insertional inactivation in this study.

3.2 MATERIALS AND METHODS

3.2.1 Bacterial strains, plasmids and culture conditions

The bacterial strains and plasmids used in this part of the study are listed in Table 3.1. *E. coli* and *P. aeruginosa* strains were routinely cultivated at 37°C with shaking at 250 rpm in Luria-Bertani (LB) broth (0.5% [w/v] yeast extract, 1% [w/v] tryptone, 1% [w/v] NaCl; pH 7.4). *P. aeruginosa* strains were maintained either on *Pseudomonas* Isolation Agar (PIA; Difco) or on *Pseudomonas* Agar Base (PAB; Sigma-Aldrich). The following antibiotics were used to maintain the plasmid DNA and chromosomal insertions in *P. aeruginosa* DSM1707 strains: neomycin at 150 µg/ml, carbenicillin at 150 µg/ml and gentamicin at 50 µg/ml. For plasmid DNA selection and maintenance in *E. coli*, the concentrations of antibiotics used were: 100 µg/ml for ampicillin, 20 µg/ml for tetracycline, 7 µg/ml for gentamicin and 100 µg/ml for kanamycin. All antibiotics were purchased from Sigma-Aldrich.

3.2.2 Genomic DNA isolation

The genomic DNA of wild-type and mutant *P. aeruginosa* DSM1707 strains was isolated using cetyltrimethylammonium bromide (CTAB) as described by Jansen (1995). Briefly, the cells from 500 µl of an overnight culture were collected by centrifugation at 10 000 rpm for 3 min and suspended in 567 µl of 1 × TE buffer (10 mM Tris-HCl, 1 mM EDTA; pH 8.0). The cells were lysed by the addition of SDS to a final concentration of 0.5% (v/v) and the proteins were digested by addition of proteinase K to a final concentration of 100 µg/ml in a total volume of 600 µl. Following incubation at 37°C for 1 h, 100 µl of 5 M NaCl and 80 µl of a CTAB/NaCl solution was added and incubation was continued for 10 min at 65°C. The CTAB-protein/polysaccharide complexes were removed by extraction with an equal volume of chloroform:isoamyl alcohol (24:1) followed by centrifugation at 10 000 rpm for 5 min. The supernatant, containing the genomic DNA, was recovered and transferred to a new microfuge tube. The remaining CTAB was removed by addition of an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) followed by centrifugation (10 000 rpm, 5 min). The chromosomal DNA was precipitated from the recovered supernatant by addition of 0.6 volume isopropanol. The precipitated chromosomal DNA was pelleted by brief centrifugation, rinsed with 70% ethanol, dried under vacuum and resuspended in 20 µl 1 ×

TE buffer or UHQ water. An aliquot of the DNA was analyzed by electrophoresis on a 1% (w/v) agarose gel.

3.2.3 DNA amplification

3.2.3.1 Oligonucleotide primers

The oligonucleotide primers used in PCR assays to amplify the genomic region containing the *htpD*, *htpE* and *htpF* ORFs from *P. aeruginosa* strain DSM1707, are indicated in Table 3.2. The primers were designed on the basis of the published complete genome sequence of *P. aeruginosa* strain PAO1 (Stover *et al.*, 2000). In order to facilitate cloning of the PCR-amplified genes, unique restriction endonuclease recognition sites were included at the 5' terminus of the respective primers (Table 3.2). The primers were synthesized by MWG Biotech or Inqaba Biotech.

3.2.3.2 Polymerase chain reaction (PCR) amplification

Each of the reaction mixtures (50 µl) contained 100 ng of *P. aeruginosa* DSM1707 genomic DNA as template, 25 pmol of each the sense and antisense primer, 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% [v/v] TritonX-100), $MgCl_2$ at 1.5 mM, each deoxynucleoside triphosphate (dNTP) at a concentration of 0.2 mM, dimethyl sulfoxide at 5% (v/v) and 1 U of *Taq* DNA polymerase (Southern Cross Biotechnology). The reaction tubes were placed in a Perkin-Elmer GeneAmp 2400 thermal cycler. Following initial denaturation of 3 min at 94°C, the reactions were subjected to 25 cycles of denaturation for 45 s at 94°C, annealing for 45 s at 63°C (primers Tad-IF - Tad-OR) or 57°C (primers Tad-NF - Tad-IR) and elongation for 2 min at 72°C. After the last cycle, the reactions were kept at 72°C for 6 min to complete synthesis of all strands. For control purposes, reaction mixtures containing UHQ water and all other reagents but no template DNA were included. Aliquots of the PCR reaction mixtures were subsequently analyzed by agarose gel electrophoresis on 1% (w/v) agarose gels in the presence of an appropriate DNA molecular weight marker.

Table 3.1 Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant properties	Reference
Strains:		
E. coli DH5α	*hsdR recA lac*ZYA ϕ 80d*lac*ZΔ*M15*	Gibco-BRL
P. aeruginosa		
DSM1707	Wild-type, Prototroph (PAO1)	DSM#
DSMHtpD	DSM1707*htpD*::Gm^r	This study
DSMHtpE	DSM1707*htpE*::Gm^r	This study
DSMHtpDEF	DSM1707*htpDEF*::Gm^r	This study
DSMDEF	Wild-type with complementation plasmid pJB-DEF-Kan	This study
Plasmids:		
pUC18	Cloning vector, ColE1, Amp^r, LacZα peptide	Stratagene
pGEM®-T Easy	Cloning vector for PCR products, ColE1, Amp^r, LacZα peptide	Promega
pBluescript SKII (+)	Cloning vector, ColE1, Amp^r, LacZα peptide	Stratagene
pJB3cT20	Derivative of pJB3, *oriV*, *oriT*, Tc^r, Amp^r	Blatney *et al.* (1997)
pRK2013	ColE1, mob$^+$tra$^+$, (RK2), Kan^r	Greener *et al.* (1992)
pGEM-Gent	pGEM®-T Easy containing gentamicin resistance cassette	Smith (2003)
pUC18-Gent	Gentamicin cassette cloned into the *Sma*I/*Sac*I sites of pUC18	This study
pGEM-IFOR	pGEM®-T Easy containing a Tad-IF - Tad-OR PCR product	This study
pUC18-IFOR	The Tad-IF - Tad-OR DNA fragment cloned into the *Kpn*I/*Hind*III sites of pUC18	This study
pGEM-NFIR	pGEM®-T Easy containing a Tad-NF - Tad-IR PCR product	This study
pUC18-DEF	pUC18-IFOR containing the Tad-NF - Tad-IR DNA fragment cloned into the *Not*I/*Kpn*I sites	This study
pUC18-htpE	*htpE* inserted as a 1.9-kb *Pst*I fragment into *Pst*I site of pUC18	This study
pUC18-htpD-Gent	pUC18-IFOR with a gentamicin cassette inserted at the *Sal*I site of the *htpD* ORF	This study
pUC18-htpE-Gent	pUC18-htpE with a gentamicin cassette inserted at the *Not*I site of the *htpE* ORF	This study
pUC18-htpDEF-Gent	pUC18-DEF with a gentamicin cassette inserted at the *Pst*I sites of *htpD* and *htpF*	This study
pUC4K	Kanamycin gene cloned into pUC18	Taylor *et al.* (1988)
pBlue-Kan	Kanamycin gene cloned into the *Eco*RI site of pBluescript SKII (+)	This study
pBlue-DEF	DNA fragment containing *htpDEF* ORFs cloned into the *Kpn*I/*Hind*III sites of pBluescript SKII (+)	This study
pJB-DEF	The *lac-htpDEF* DNA fragment cloned into the *Eco*RV/*Kpn*I sites of pJB3cT20	This study
pJB-DEF-Kan	pJB-DEF with the kanamycin gene cloned into the *Sca*I site	This study

DSM - Deutsche Sammlung von Mikroorganismen, Braunschweig, Germany

Table 3.2 Primers used in this study

Oligonucleotide primer	Nucleic acid sequence
PCR amplification*:	
Tad-1F	5' - GCGGGCGCTGGATACC**ggtacc**ACGCATCTGG - 3'; *Kpn*I site incorporated
Tad-OR	5' - CCGGATCC**aagctt**CGGCGCCTCGACCAGACCC - 3'; *Hind*III site incorporated
Tad-NF	5' - CCCAGGTCGTTGC**ggtacc**GGCTCTCGGTCGG - 3'; *Kpn*I site incorporated
Tad-IR	5' - CGG**Ttctaga**GCTGCTGGGACGCCAGGAGGC - 3'; *Xba*I site incorporated
Tad-GR	5' - CCGGACGCATAGCAGGGGTCTGCC - 3'
Gent-1	5' - CG**gatatc**CCTTCCAGAAAACCGAGG - 3'; *Bss*HI site incorporated
Gent-2	5' - **gcgcgc**TCAGTCCAGTTATGCTGTG - 3'; *Eco*RV site incorporated
Nucleic acid sequencing:	
Tad-F1	5' - CGGCCGGGCCAAGGCGGATCCGCT - 3'
Tad-F2	5' - CTCAAGAGCGTCTAGGGATCCGCC - 3'
Tad-F3	5' - GCCACTGTGCCTGAATTCCCACC - 3'
Tad-F4	5' - CCGATGCCGAATTCTGCTTTCAT - 3'
pUC/M13 Forward	5' - GTTTCCCAGTCACGAC - 3'
pUC/M13 Reverse	5' - GTAAAACGACGGCCAGT - 3'

* In primer sequences, the restriction endonuclease sites are indicated in bold lower case letters, while the annealing position of the respective primers on the *P. aeruginosa* genome are shown in Fig. 3.1.

3.2.4 Agarose gel electrophoresis

DNA was analyzed by agarose gel electrophoresis (Sambrook *et al.*, 1989). For this purpose, horizontal 1% (w/v) agarose slab gels were cast and electrophoresed at 100 V in 1 × TAE buffer (40 mM Tris-HCl, 20 mM NaOAC, 1 mM EDTA; pH 8.5). The agarose gels were supplemented with ethidium bromide (0.5 µg/ml) in order to allow visualization of the DNA on an UV transilluminator. Where appropriate, the DNA fragments were sized according to their migration in the gel as compared to that of standard DNA molecular weight markers, namely phage lambda DNA digested with *Pst*I, *Eco*RI and/or *Hind*III (Roche).

3.2.5 Purification of DNA fragments from agarose gels

DNA fragments were purified from 1% (w/v) agarose gels using a silica suspension as described by Boyle and Lew (1995). Briefly, the DNA band of interest was excised from the agarose gel and mixed with 400 µl of a 6 M NaI solution. The agarose was dissolved by incubation at 55°C for 10 min, after which 8 µl of the silica suspension was added to the sample. The DNA was allowed to bind to the silica by incubation of the samples on ice for 30

min with intermittent vortexing. The DNA-silica complex was pelleted by centrifugation (10 000 rpm for 30 s) and washed four times with Wash buffer (50 mM NaCl, 10 mM Tris-HCl (pH 7.5), 2.5 mM EDTA, 50% [v/v] ethanol). The DNA was eluted from the silica matrix in a final volume of 7 µl UHQ water by incubation at 55°C for 10 min. The purified DNA fragments were analyzed on a 1% (w/v) agarose gel to assess both their purity and concentration.

3.2.6 Restriction endonuclease digestions

All restriction endonuclease digestions were performed in sterile microfuge tubes and contained the appropriate concentration of salt (using the 10 × buffer supplied by the manufacturer) for the specific enzyme and 5-10 U of enzyme per µg of plasmid DNA. The reaction volumes were small (10-20 µl) and incubation was typically for 1-1.5 h at 37°C, except for *Sma*I, which was incubated at 25°C. Partial restriction endonuclease digestions were performed by incubating the reaction mixtures at the optimum temperature for enzymatic activity, but limiting the time of incubation to 20 min. When digestion entailed the use of two enzymes requiring different salt concentrations for optimal activity, the plasmid DNA was first digested with the enzyme requiring a lower salt concentration, after which the salt concentration was adjusted and the second enzyme added. All restriction enzymes were supplied by Roche or Promega. The digestion products were typically analyzed on a 1% (w/v) agarose gel in the presence of appropriate DNA molecular weight markers.

3.2.7 Cloning of DNA fragments into plasmid vectors

3.2.7.1 Ligation of DNA fragments

Purified restricted vector DNA and specific DNA fragments were ligated at 16°C for 16 h in a final reaction volume of 10 µl, which contained 1 µl of a 10 × DNA ligase buffer (660 mM Tris-HCl (pH 7.5), 10 mM DTT, 50 mM $MgCl_2$, 10 mM ATP) and 3 U of T4 DNA ligase (Promega, 3 U/µl). The ratio of vector to insert was typical in excess of 1:5. For cloning of PCR amplicons, the pGEM®-T Easy vector system (Promega) was used. Approximately 150 ng of the gel-purified amplicons were ligated to 50 ng of linear pGEM®-T Easy vector DNA using the conditions described above.

3.2.7.2 Preparation of competent *E. coli* DH5α cells

Competent *E. coli* DH5α cells were prepared according to the procedures described by Chung and Miller (1988). An overnight culture was prepared by inoculating 10 ml of LB-broth with a colony from a freshly streaked culture of *E. coli* DH5α. After overnight incubation at 37°C with shaking, 1 ml of the culture was inoculated into 100 ml preheated (at 37°C) sterile LB-broth and grown to an OD_{540} of 0.3 to 0.4. The cells from 30 ml of the culture were pelleted in a polypropylene tube by centrifugation at 5 000 rpm for 10 min at 4°C. The pellet was suspended in 3 ml ice-cold TSS (0.1 M $MgCl_2$, 0.1 M $MgSO_4$, 10% [w/v] PEG 8000 prepared in 93 ml LB-broth, 5% [v/v] DMSO; pH 6.5). Following incubation on ice for 15 min, the cell suspension was aliquoted into microfuge tubes and stored at -70°C until use.

3.2.7.3 Transformation of competent *E. coli* DH5α cells

After allowing the competent *E. coli* DH5α cells to thaw on ice, the cells were transformed using the method described by Chung and Miller (1988). An aliquot of the cells (100 µl) and the ligation reaction mixture (10 µl) were mixed in a pre-cooled microfuge tube and incubated on ice for 1 h. After addition of 500 µl LB-broth containing 20 mM glucose, the transformation mixtures were incubated with shaking at 37°C for 3 h. The transformed cells were selected by plating the cells in aliquots of 100-200 µl onto LB-agar plates supplemented with the appropriate antibiotic. The plates were incubated overnight at 37°C and investigated for the presence of recombinant transformants. When appropriate, the cells were plated together with 10 µl IPTG (100 mM stock solution) and 50 µl X-gal (2% [w/v] stock solution) to allow for blue/white colour selection, based on insertional inactivation of the *lacZ'* marker gene in the pUC18 (Stratagene), pGEM®-T Easy (Promega) and pBluescript SKII (+) (Stratagene) vectors.

3.2.8 Extraction and purification of plasmid DNA

3.2.8.1 Plasmid DNA extraction

Plasmid DNA was isolated from selected colonies using a modified alkaline lysis method (Sambrook *et al.*, 1989). Colonies were picked from the agar plates with sterile toothpicks, inoculated into 10 ml of LB-broth containing the appropriate antibiotic and then incubated

overnight at 37°C with shaking. After incubation, cells from 3 ml of the overnight cultures were collected by centrifugation for 4 min at 10 000 rpm. The supernatant was discarded and the bacterial cell pellet suspended in 400 µl of Solution 1 (50 mM glucose, 25 mM Tris-HCl, 10 mM EDTA, 10 mg/ml lysozyme; pH 8.0) before incubation at room temperature for 10 min. The spheroplasts were lysed following the addition of 400 µl of freshly prepared Solution 2 (0.2 N NaOH, 1% [w/v] SDS). After incubation on ice for 10 min, 300 µl of 7.5 M ammonium acetate (pH 7.6) was added and incubation was continued on ice for a further 10 min. The cellular debris was removed by centrifugation for 10 min at 10 000 rpm, after which the plasmid DNA was precipitated from the recovered supernatants by the addition of 650 µl isopropanol at room temperature for 10 min. The precipitated plasmid DNA was collected by centrifugation, as above, and resuspended in 100 µl of 2 M ammonium acetate (pH 7.4). After incubation on ice for 10 min, the excess proteins were removed by centrifugation at 10 000 rpm for 10 min and the plasmid DNA in the supernatant precipitated by the addition of 110 µl isopropanol. The plasmid DNA was pelleted by centrifugation, rinsed with 500 µl of 70% ethanol, dried under vacuum for 10 min and then resuspended in 15 µl UHQ water.

3.2.8.2 Purification of plasmid DNA

To completely remove contaminating RNA from the plasmid DNA extractions, the plasmid DNA was incubated with 0.5 µl RNase A (10 mg/ml) at 37°C for 30 min. The plasmid DNA was precipitated by the addition of 2.5 volumes 96% ethanol and NaOAc (pH 7.0) to a final concentration of 0.3 M. After incubation at -20°C for 30 min, the precipitated plasmid DNA was collected by centrifugation at 15 000 rpm for 15 min. The pelleted plasmid DNA was rinsed with 70% ethanol, dried under vacuum and resuspended in 15 µl UHQ water.

3.2.9 Nucleic acid sequencing

3.2.9.1 Sequencing reactions

The nucleotide sequence of cloned insert DNA was determined using an ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems, Perkin-Elmer). In addition to the universal pUC/M13 forward and reverse sequencing primers, internal primers were also used in the sequencing reactions. The primers that were

used in the sequencing reactions are indicated in Table 3.2, while a diagram depicting the annealing positions and orientation of the sequencing primers are indicated in Fig. 3.1. Each reaction mixture contained 80-100 ng of purified plasmid DNA, 12.5 pmol of sequencing primer, 2 µl Terminator Ready Reaction Mix and UHQ water in a final volume of 5 µl. Cycle sequencing reactions were performed in a Perkin-Elmer GeneAmp 2400 thermal cycler with 25 of the following cycles: denaturation at 96°C for 30 s, primer annealing at 50°C for 15 s and extension at 60°C for 4 min. Following brief centrifugation, the extension products were precipitated by the addition of 2 µl of 3 M NaOAC (pH 4.6) and 50 µl of 95% ethanol in a final volume of 70 µl. The tubes were incubated at room temperature for 20 min in the dark, centrifuged at 15 000 rpm for 30 min and the supernatants carefully aspirated. The pellets were rinsed twice with 250 µl of 70% ethanol, dried under vacuum for 10 min and stored at 4°C. Prior to electrophoresis, the purified extension products were resuspended in 3.5 µl Blue dextran/EDTA loading buffer, denatured for 2 min at 90°C and loaded onto a model 377 automated DNA sequencer (Perkin-Elmer).

3.2.9.2 Sequence analysis

The obtained nucleic acid sequences were assembled and analysed using DNAMAN (Lynnon Biosoft). The nuclotide and deduced amino acid sequences were compared against the *P. aeruginosa* genome database (http://www.pseudomonas.com) and also against the sequences in the GenBank Database (http://www.ncbi.nlm.nih.gov/) by using the BLASTN and BLASTP programmes (Altschul *et al.*, 1997).

3.2.10 Construction of allelic exchange vectors

All molecular cloning techniques employed in the construction of the allelic exchange vectors were performed according to the procedures described in the preceding sections. All plasmid constructs were confirmed by restriction endonuclease digestion using agarose gel electrophoresis and by nucleic acid sequencing. The cloning strategy employed in the construction of the recombinant pUC18-DEF plasmid is indicated diagrammatically in Fig. 3.2, and those used in the construction of allelic exchange vectors pUC18-htpD-Gent, pUC18-htpE-Gent and pUC18-htpDEF-Gent are indicated in Fig. 3.3.

3.2.10.1 Cloning of a gentamicin resistance cassette into pUC18

A 1.2-kb gentamicin resistance cassette, obtained by PCR amplification using the pFASTBAC™ (Gibco-BRL) plasmid DNA as template, has previously been cloned into the pGEM®-T Easy vector to create pGEM-Gent (Smith, 2003). The gentamicin resistance cassette was subsequently recovered from pGEM-Gent by digestion with *Sac*I followed by a partial digestion with *Eco*RV. The gentamicin resistance cassette was agarose gel-purified and then cloned into pUC18, which had been digested with *Sac*I and *Sma*I. One of the resulting recombinant clones was selected for further use and designated pUC18-Gent.

3.2.10.2 Cloning of the full-length *htpDEF* ORFs into pUC18

To obtain the genomic region of *P. aeruginosa* DSM1707 containing the *htpDEF* genes, two separate PCRs were performed to generate DNA fragments that represented overlapping sections of the *htpDEF* gene cluster. The DNA fragments overlapped each other by 91 nucleotides and contained a mutual unique *Not*I restriction endonuclease recognition site in the overlap region. Following restriction enzyme digestion of the respective DNA fragments, they could therefore be ligated to yield the desired full-length clone.

In the first PCR, oligonucleotide primers Tad-IF (containing a *Kpn*I site) and Tad-OR (containing a *Hind*III site) were used with chromosomal DNA from *P. aeruginosa* DSM1707 to generate a 2.0-kb product containing the full-length *htpD* ORF and 200 bp of the 5' halve of the *htpE* ORF (Section 3.2.3.2). The amplicon was purified from the agarose gel and cloned into pGEM®-T Easy to yield recombinant plasmid pGEM-IFOR. The cloned DNA fragment was subsequently recovered by *Kpn*I and *Hind*III restriction endonuclease digestion, agarose gel-purified and cloned into identically digested pUC18, yielding pUC18-IFOR. In the second PCR, oligonucleotide primers Tad-NF (containing a *Kpn*I site) and Tad-IR (containing a *Xba*I site) were used with chromosomal DNA from *P. aeruginosa* DSM1707 to generate a 2.1-kb product containing 720 bp of the 3' halve of the *htpE* ORF and the full-length *htpF* ORF (Section 3.2.3.2). The amplicon was cloned into the pGEM®-T Easy vector to create pGEM-NFIR. The cloned DNA fragment was then recovered from plasmid pGEM-NFIR by digestion with both *Kpn*I and *Not*I, gel-purified and cloned into *Kpn*I/*Not*I-digested pUC18-IFOR to yield pUC18-DEF (Fig. 3.2).

3.2.10.3 Construction of allelic exchange vector pUC18-htpD-Gent

Plasmid pUC18-IFOR was partially digested with *Sal*I, which cuts three times in the *htpD* ORF and once in the *P. aeruginosa* DNA, immediately upstream from the *htpD* ORF. A mixture of DNA restriction fragments of between 4.4 to 4.8 kb were subsequently purified from the agarose gel and then ligated with the gentamicin resistance cassette, which had been recovered from plasmid pUC18-Gent by digestion with *Sal*I. Following transformation of competent *E. coli* DH5α cells, the plasmid DNA was extracted from gentamicin-resistant transformants and characterized by restriction enzyme digestion with *Sal*I and with both *Eco*RI and *Hind*III. One of the resultant recombinant clones was selected and designated pUC18-htpD-Gent (Fig. 3.3).

3.2.10.4 Construction of allelic exchange vector pUC18-htpE-Gent

Recombinant plasmid pUC18-DEF, harboring the *P. aeruginosa* DSM1707 *htpDEF* ORFs, was digested with *Pst*I. The *Pst*I restriction enzyme cuts once in both the *htpD* and *htpF* ORFs and once in the *P. aeruginosa* DNA downstream from the *htpF* ORF. Thus, following *Pst*I-digestion of plasmid pUC18-DEF, a fragment of 1.9 kb, which harbors the full-length *htpE* ORF, was purified from the agarose gel and then cloned into pUC18, which had been linearized by *Pst*I digestion, to create pUC18-htpE. This plasmid was digested with *Not*I, which cuts once in the *htpE* ORF, and a gentamicin resistance cassette, recovered from pGEM-Gent by *Not*I digestion, was ligated into this site to construct the allelic exchange vector pUC18-htpE-Gent (Fig. 3.3).

3.2.10.5 Construction of allelic exchange vector pUC18-htpDEF-Gent

Recombinant plasmid pUC18-DEF was partially digested with *Pst*I to remove a 1.9-kb DNA fragment containing 341 bp of the *htpD* ORF, the full-length *htpE* ORF (884 bp) and 654 bp of the *htpF* ORF. The gentamicin resistance cassette was recovered from pUC18-Gent by *Pst*I digestion and then ligated into this deletion site of pUC18-DEF to produce the allelic exchange vector pUC18-htpDEF-Gent (Fig. 3.3).

3.2.11 Construction of mutant *P. aeruginosa* DSM1707 strains

3.2.11.1 Preparation of competent *P. aeruginosa* DSM1707 cells

Competent *P. aeruginosa* DSM1707 cells were prepared according to the procedures described by Olsen *et al.* (1982). A single colony of *P. aeruginosa* DSM1707, grown overnight on a TN-agar plate (0.5% [w/v] tryptone, 0.1% [w/v] dextrose, 0.25% [w/v] yeast extract, 0.4% [w/v] $NaNO_3$, 1.2% [w/v] agar) (Olsen and Shipley, 1973), was inoculated into 10 ml of TN-broth and grown overnight at 37°C with shaking. After incubation, 500 µl of the culture was used to inoculate 50 ml of TN-broth and then grown at 37°C to an OD_{540} of 0.5. The culture was incubated on ice for 20 min to inhibit further growth. The cells from 4 ml of the culture were collected in a microfuge tube by centrifugation at 5 000 rpm for 5 min at 4°C. The pellet was suspended in 2 ml ice-cold filter-sterilized 0.15 M $MgCl_2$, incubated on ice for 5 min, pelleted as before, and gently resuspended in 1 ml of the ice-cold $MgCl_2$ solution. After incubation on ice for 20 min, the cells were again collected by centrifugation and the pellet finally resuspended in 100 µl of ice-cold 0.15 M $MgCl_2$.

3.2.11.2 Transformation of competent *P. aeruginosa* cells

For introducing plasmid DNA into *P. aeruginosa* DSM1707 cells, the prepared competent cells (100 µl) were mixed with 300 ng plasmid DNA in a sterile microfuge tube. Following incubation for 1 h on ice, the transformation mixtures were incubated at 42°C for 90 s and then chilled on ice for 5 min. After addition of 500 µl of preheated (37°C) TN-broth, the transformation mixtures were incubated for 3 h at 37°C with shaking and then plated onto LB-agar plates supplemented with the appropriate antibiotic. The plates were incubated overnight at 37°C.

For introducing the allelic exchange vectors into *P. aeruginosa* DSM1707, the same procedure as above was followed, except for the following modifications. After incubation of the transformation mixtures for 3 h at 37°C, the mixtures were inoculated into 10 ml of LB-broth lacking antibiotics and incubated for 24 h at 37°C with shaking. An aliquot (50 µl) of the overnight culture was then used to inoculate 10 ml of LB-broth supplemented with 15 µg/ml gentamicin and the culture was incubated for a further 24 h at 37°C. The procedure

was repeated once more, except that the LB-broth was supplemented with gentamicin at a concentration of 50 µg/ml. Following incubation, aliquots of the culture were plated onto LB-agar plates supplemented with 50 µg/ml gentamicin. The agar plates were incubated overnight at 37°C and observed for the presence of gentamicin-resistant transformants.

3.2.11.3 Triparental conjugation

As an alternative to the above transformation procedures, the complementation plasmid, pJB-DEF-Kan, was introduced into *P. aeruginosa* DSM1707 mutant strains by triparental conjugation. The triparental matings were performed with helper plasmid pRK2013 as previously described by Greener *et al.* (1992), with the following modifications. A single colony of freshly streaked cultures of donor (*E. coli* DH5α containing the complementation plasmid), helper (*E. coli* DH5α containing pRK2013) and recipient (mutant *P. aeruginosa*) strains were mixed on a LB-agar plate with a sterile inoculation needle and then incubated overnight at 37°C. Following incubation, the mixed growth was streaked onto selective medium consisting of LB-agar supplemented with 50 µg/ml gentamicin, 100 µg/ml ampicillin and 150 µg/ml neomycin. The agar plates were then incubated at 37°C for a further 24 to 48 h.

3.2.12 Characterization of mutant *P. aeruginosa* DSM1707 strains

The replacement of the wild-type *htp* ORFs with the mutant null alleles in the putative *P. aeruginosa* mutant strains was verified by two methods, Southern blot hybridization (Southern, 1975) and PCR analysis.

3.2.12.1 Southern blot analysis

3.2.12.1.1 Preparation of labeled probes

Prior to the preparation of labeled probe DNA, pUC18 plasmid DNA was linearized by digestion with *Eco*RI, whereas the gentamicin resistance cassette was obtained by PCR amplification using pUC18-Gent as template DNA and primers Gent-1 and Gent-2 (Table 3.2). The reaction conditions were as described in Section 3.2.3.2, except that annealing was

performed at 55°C for 1 min. To prepare DIG-dUTP labeled probes, 1 µg of each the purified gentamicin amplicon and linearized pUC18 plasmid DNA was diluted in UHQ water to a final volume of 16 µl. The DNA was denatured by heating in a boiling water bath for 10 min and then quickly chilled in an ice water bath. The denatured DNA was then mixed with 4 µl DIG-High Prime (consisting of random primers, dNTPs, DIG-dUTP, Klenow polymerase and buffer; Roche) and incubated for 16 h at 37°C. Following incubation, the reactions were terminated by the addition of 2 µl of 0.2 M EDTA (pH 8.0) and heating the reaction mixtures to 65°C for 10 min.

3.2.12.1.2 Preparation of the membranes

Preparations of the chromosomal DNA of *P. aeruginosa* wild-type and mutant strains (Section 3.2.2) were digested with *Eco*RI at 37°C overnight. Recombinant pUC18-Gent and *Eco*RI-linearized pUC18 plasmid DNA were included in the analyses as positive controls. The resultant DNA fragments were separated by electrophoresis in 1% (w/v) agarose gels.

The DNA fragments were transferred from the agarose gels to Hybond™-N nylon membranes (Amersham Life Science) by capillary blotting as follows. The DNA was first denatured by soaking the gels for 45 min with constant agitation in a denaturation solution (1.5 M NaCl, 0.5 M NaOH), after which the gels were rinsed in ddH_2O and then neutralized in a neutralization solution (1 M Tris-HCl (pH 7.2), 1.5 M NaCl, 10 mM EDTA) as above. Two pieces of filter paper, soaked in 20 × SSC (3 M NaCl, 0.3 M Na.citrate; pH 7.0), were stacked on a piece of Glad Wrap, after which the inverted gels were placed onto these and the Glad Wrap folded so as to surround the gels. The gels were then overlaid with nylon membranes followed by two more pieces of filter paper, all of which were prewet in 2 × SSC. The filter papers and membranes were the same size as the gels and the orientation of the membranes were marked by cutting off matching corners of both the gel and membrane. Four additional dry filter papers and several paper towels were stacked on top of the wet filter papers and weighed down by a light weight. Transfer was allowed to proceed at room temperature for 18 h. The membranes were then rinsed in 2 × SSC and the DNA fixed to the membranes by UV irradiation for 5 min each side.

3.2.12.1.2 Nucleic acid hybridization

Each membrane was sealed in a separate plastic bag together with DIG-Easy Hyb buffer (Roche) and prehybridized for 30 min. The prehybridization buffer was then removed from the plastic bags and replaced with hybridization buffer so as to just cover the membranes, after which 25 ng/ml of the denatured labeled probe (Section 3.2.12.2.1) was added to the buffer and the bag resealed. The prehybridization and hybridization buffers were the same. Hybridization was allowed to proceed at 44°C for 4 h. After hybridization, the membranes were recovered and washed twice for 5 min each time in 2 × SSC; 0.1% SDS at room temperature, followed by 0.5 × SSC; 0.1% SDS at 65°C.

The hybridized probes were detected by rinsing the membranes in washing buffer (0.1 M maleic acid, 0.15 M NaCl, 0.3% [v/v] Tween-20; pH 7.5), followed by incubation for 30 min in 1 × blocking solution (prepared by diluting the supplied blocking solution 10-fold in maleic acid buffer [0.1 M maleic acid, 0.15 M NaCl; pH 7.5]). The membranes were incubated for 30 min at room temperature in antibody solution (alkaline phosphatase-conjugated anti-digoxigenin, diluted 1:5000 in fresh blocking solution). The unbound primary antibodies were removed by washing the membranes twice for 15 min each wash in washing buffer and equilibrated for 5 min in detection buffer (0.1 M Tris-HCl, 0.1 M NaCl; pH 9.5). The membranes were then immersed in the alkaline phosphatase enzyme substrate (NBT/BCIP stock diluted 1:50 in detection buffer) until the bands became visible.

3.2.12.2 PCR analysis

The DSMHtpD and DSMHtpE mutant strains were analyzed for the presence of the gentamicin resistance cassette within the *htp* ORFs by PCR analyses. Oligonucleotide primers (Table 3.2, Fig. 3.1) were used to amplify a hybrid amplicon consisting of either the 5' end of the interrupted *htp* ORF and the 3' end of the gentamicin resistance cassette or alternatively, to amplify a hybrid amplicon consisting of the 5' end of the gentamicin resistance cassette and the 3' end of the interrupted *htp* ORF. Appropriate oligonucleotide primers were also used to amplify the interrupted *htp* ORF plus the gentamicin resistance cassette in mutant chromosomal DNA.

3.2.12.2.1 Oligonucleotide primers

To amplify the hybrid products from the DSMHtpD strain, oligonucleotide primer pair Tad-GR and Gent-2, and primer pair Tad-F3 and Gent-1 were used, while primer pair Tad-GR and Tad-F3 was used to amplify the interrupted *htpD* ORF plus the gentamicin resistance cassette. For analysis of the DSMHtpE strain, oligonucleotide primer pair Tad-NF and Gent-2, and primer pair Tad-IR and Gent-1 was used to amplify the hybrid products, while primers Tad-NF and Tad-IR were used to amplify the interrupted *htpE* ORF.

3.2.12.2.2 PCR amplification

The PCR reaction mixtures (50 µl) contained 100 ng of chromosomal DNA, 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% [v/v] TritonX-100), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 25 pmol of each the sense and antisense primer and 1 U of *Taq* DNA polymerase (Southern Cross Biotechnology). PCR was performed in a Perkin-Elmer GeneAmp 2400 thermal cycler. For analysis of the DSMHtpE strain, the cycling profile consisted of initial denaturation of 4 min at 94°C, followed by 25 cycles of denaturation for 30 s at 94°C, primer annealing for 30 s at 55°C and elongation for 1.5 min at 72°C. For analysis of the DSMHtpD strain, the cycling profile was similar, except primer annealing was performed at 58°C for 30 s and elongation was performed at 68°C for 4 min for the first 10 cycles, after which the extension time was increased by 20 s per cycle. After the last cycle, the reactions were kept at 72°C for 6 min to complete synthesis of all strands. For all of the analyses, UHQ water served as a negative control, while chromosomal DNA extracted from the parental and mutant strains provided sample template DNA. Following PCR amplification, aliquots of the respective reaction mixtures were analyzed by electrophoresis on 1% (w/v) agarose gels.

3.2.13 Complementation analysis

To prove that any altered phenotypes that may be displayed by the mutant *P. aeruginosa* strains was due to the disruption of the respective *htp* ORFs, a complementation plasmid was constructed whereby the mutant strains could be complemented. The strategy used for

construction of the pJB-DEF-Kan complementation plasmid is indicated diagrammatically in Fig. 3.4.

3.2.13.1 Cloning of a kanamycin resistance gene into pBluescript SKII (+)

Plasmid pUC4K, containing a kanamycin resistance gene (Taylor *et al.*, 1988), was digested with *Eco*RI to excise the kanamycin gene. The 1.3-kb DNA fragment was purified from the agarose gel and cloned into an *Eco*RI-digested pBluescript SKII (+) vector to yield pBlue-Kan.

3.2.13.2 Construction of complementation plasmid pJB-htpDEF-Kan

The recombinant plasmid pUC18-DEF (Section 3.2.10.2) was digested with both *Kpn*I and *Hind*III to excise the cloned *P. aeruginosa*-specific DNA insert of 4.1 kb containing the full-length *htpD*, *htpE* and *htpF* ORFs. Following purification of the DNA fragment from an agarose gel, it was cloned into pBluescript SKII (+), which had been digested with both *Kpn*I and *Hind*III. The resultant plasmid, pBlue-DEF, was then digested with both *Pvu*II and *Kpn*I to excise the *P. aeruginosa*-specific DNA insert together with the upstream *lac*I promoter of the pBluescript SKII (+) vector. The 4.4-kb DNA fragment was cloned into the broad-host-range pJB3Tc20 vector (Blatney *et al.*, 1997), which had been digested with both *Eco*RV and *Kpn*I, to yield plasmid pJB-DEF. This cloning step, however, resulted in the inactivation of the tetracycline resistance gene on plasmid pJB3Tc20. Thus, to provide a selection marker, the 1.3-kb kanamycin resistance cassette was excised from plasmid pBlue-Kan by *Pvu*II digestion and inserted by blunt-end ligation into the unique *Sca*I site of plasmid pJB-DEF, thereby completing the construction of the complementation plasmid pJB-DEF-Kan.

3.3 RESULTS

The primary aim of this part of the study was to generate mutant strains of the wild-type *P. aeruginosa* DSM1707 strain in which the wild-type *htpD*, *htpE* and *htpDEF* open reading frames (ORFs) on the *P. aeruginosa* genome were replaced with *in vitro*-modified alleles. Based on the results obtained by homology searches (Chapter 2), it is possible that *htpD* may encode for a cytoplasmic NTPase, while both *htpE* and *htpF* may encode proteins that are

localized to the inner membrane. Since the predicted protein products of these *P. aeruginosa* genes appear to be conserved in both of the recently described *flp-rcp-tad* (Kachlany *et al.*, 2000; 2001) and *pilA-cpa* (Skerker and Shapiro, 2000) pilus assembly and secretion systems, allelic exchange vectors were thus constructed in which either the *htpD*, *htpE* or all three *htp* ORFs was disrupted by the insertion of a cassette encoding gentamicin resistance. Mutants of *P. aeruginosa* DSM1707 were constructed by introducing the allelic exchange vectors into the wild-type strain and selecting for subsequent homologous recombination events between the *htp* DNA flanking the antibiotic cassette in each of the vectors and the wild-type *htp* locus in the genome. The resultant mutant strains were finally characterized by Southern blot hybridization and by PCR analysis to verify the presence of the mutant alleles. Furthermore, to investigate whether any altered phenotypes that may be displayed by the mutant *P. aeruginosa* strains were due to the disruption of the respective *htp* ORFs, a plasmid was constructed whereby the mutant strains could be complemented. Thus, a transcriptional fusion between the strong *lac* promoter, obtained from the pBluescript SKII (+) vector, and the intact full-length *htpDEF* ORFs were constructed and subsequently cloned into the broad-host-range plasmid pJB3Tc20. The following sections will aim to provide more detailed information regarding the strategies that were used in this part of the study, as well as the results that were obtained.

3.3.1 Construction of pUC18-Gent

Since double crossover events that incorporate a cloned gene from a plasmid into the chromosome of an organism are rare, it is not feasible to screen for such events if the cloned gene cannot be directly selected. However, by inactivating the cloned gene with a readily selectable marker, such as an appropriate antibiotic resistance gene, it is possible to directly screen for potential mutants based on their newly acquired resistance to the antibiotic. The gentamicin resistance cassette that was used in this study to disrupt the *htp* ORFs of *P. aeruginosa* consists of the gene *aacC1*, which encodes the enzyme 3-N-aminoglycoside acetyltransferase, and is flanked by transcriptional and translational stop signals (Luckow *et al.*, 1993). By making use of the *aacC1* gene, it would thus be possible to rapidly and directly screen for potential *htp* mutant strains based on their newly acquired resistance to gentamicin.

A recombinant plasmid, pGEM-Gent, had previously been constructed by cloning a PCR-amplified gentamicin resistance cassette into the pGEM®-T Easy vector (Smith, 2003). However, this vector lacked the appropriate restriction endonuclease recognition sites that would be required for the cloning strategies used in this study. Thus, the gentamicin resistance cassette was recloned into the pUC18 vector using a strategy whereby the gentamicin cassette would be flanked by *Sal*I and *Pst*I restriction endonuclease recognition sites in the newly constructed pUC18-Gent vector. To construct plasmid pUC18-Gent, the cloned gentamicin resistance cassette was to be recovered from pGEM-Gent by digestion with both *Sac*I and *Eco*RV. This should yield a 1.2-kb DNA fragment consisting of the gentamicin cassette flanked at its 5' end with DNA sequences derived from the multiple cloning site (MCS) of the pGEM®-T Easy vector containing, amongst other, the *Sal*I and *Pst*I recognition sites. However, initial attempts at recovering the gentamicin resistance cassette using these restriction endonucleases consistently yielded a DNA fragment that was slightly smaller than the expected full-length gentamicin cassette. Subsequent nucleic acid sequence analysis of the cloned full-length gentamicin cassette indicated that it contained an *Eco*RV restriction endonuclease recognition site located 100 bp from the 5' end of the gene. This site was in addition to an *Eco*RV recognition site that had been incorporated at the 3' end of the gentamicin cassette during the PCR amplification of the cassette, using the Gent-2 primer (Table 3.2).

Thus, to avoid internal digestion of the gentamicin cassette, the pGEM-Gent plasmid was digested with *Sac*I and, after verifying complete digestion of the plasmid DNA by agarose gel electrophoresis of an aliquot of the reaction mixture, the DNA was then subjected to a partial restriction digestion with *Eco*RV. Following agarose gel electrophoresis of the reaction mixture, three digestion products of *ca.* 3.0, 1.2 and 1.1 kb, respectively, could be observed in the gel. Whereas the 3.0-kb DNA fragment corresponded to the size of the pGEM®-T Easy vector, the 1.1-kb DNA fragment represented a truncated version of the gentamicin cassette, while the 1.2-kb DNA fragment corresponded to the expected size of the full-length gentamicin resistance cassette. This DNA fragment was therefore purified from the agarose gel and ligated into pUC18, which had been digested with both *Sma*I and *Sac*I. Although *Eco*RV and *Sma*I generate blunt ends so that the termini are compatible, the hybrid site does, however, not constitute a target for either of the restriction endonucleases following ligation. After transformation of competent *E. coli* DH5α cells, recombinant plasmid DNA was extracted from gentamicin-resistant transformants and characterized by restriction enzyme

digestion. Digestion of the plasmid DNA with *Pst*I resulted in DNA fragments corresponding to the size of the pUC18 vector (2.7 kb) and gentamicin resistance cassette (1.2 kb) (Fig. 3.5, lane 5). One of the recombinant clones was selected, designated pUC18-Gent and used in all subsequent DNA manipulations.

3.3.2 Construction of a recombinant pUC18 vector containing the full-length *htpD*, *htpE* and *htpF* ORFs

The high G + C content (66.6%) of the *P. aeruginosa* genomic DNA (Stover *et al.*, 2000) and the large segment of genomic DNA to be amplified, necessitated the use of two separate PCR reactions, using four different primers, to generate two overlapping sections of the genomic region containing the full-length *htpDEF* ORFs. Thus, primers Tad-IF and Tad-OR were used to amplify a 2.057 kb region containing the full-length *htpD* ORF and 200 bp of the 5' halve of the *htpE* ORF, whereas primers Tad-NF and Tad-IR were used to amplify a 2.113 kb region containing 720 bp of the 3' halve of the *htpE* ORF and the full-length *htpF* ORF (Fig. 3.1). The primers were designed in such a way that the amplified DNA fragments would overlap each other by 91 nucleotides in the *htpE* ORF. This overlap region contained an unique *Not*I restriction endonuclease recognition site located in the *htpE* ORF. The DNA fragments could consequently be digested and ligated to construct a clone containing the full-length *htpD*, *htpE* and *htpF* ORFs, without the incorporation of additional sequences. The strategy for the construction of pUC18-DEF is indicated in Fig. 3.2.

3.3.2.1 Construction of plasmid pUC18-IFOR

Oligonucleotide primers Tad-IF (containing a *Kpn*I site) and Tad-OR (containing a *Hind*III site) were used in a PCR with chromosomal DNA of *P. aeruginosa* DSM1707 as described under Materials and Methods (Section 3.2.3.2). An aliquot of the reaction mixture was analyzed by agarose gel electrophoresis and a single discreet amplicon of the expected size (*ca.* 2.0 kb) was observed (Fig. 3.6a, lane 2). By contrast, no amplification products were observed in the negative control in which template DNA was omitted.

The amplicon was gel-purified and ligated into pGEM®-T Easy vector DNA. Following transformation of competent *E. coli* DH5α cells, recombinant transformants with a Lac$^-$ phenotype were selected from X-gal containing indicator plates and cultured in LB-broth

supplemented with ampicillin. The extracted plasmid DNA was analyzed by agarose gel electrophoresis. Plasmid DNA migrating slower than the parental pGEM®-T Easy vector DNA were selected and analyzed for the presence of a cloned insert DNA by using restriction endonucleases of which the recognition sites had been incorporated during the design of the primers. The putative recombinant plasmid DNA was therefore digested with both *Kpn*I and *Hind*III. Following agarose gel electrophoresis, restriction fragments of *ca.* 3.0 kb and 2.0 kb, respectively, were observed, which is in agreement with the expected size of the pGEM®-T Easy vector (3.0 kb) and insert DNA (2.0 kb) (Fig. 3.6a, lane 7). A recombinant clone, designated pGEM-IFOR, was selected and the integrity of the cloned insert DNA was verified by nucleic acid sequence analysis of both the terminal ends prior to it being used in further DNA manipulations. To construct plasmid pUC18-IFOR, the insert DNA was recovered from pGEM-IFOR by digestion with both *Kpn*I and *Hind*III and cloned into similarly prepared pUC18 vector DNA. A recombinant plasmid from which an insert of the expected size was excised by digestion with both *Kpn*I and *Hind*III (Fig. 3.6b, lane 4) was designated pUC18-IFOR and used in the construction of pUC18-DEF.

3.3.2.2 Construction of plasmid pGEM-NFIR

By making use of oligonucleotide primers Tad-NF (containing a *Kpn*I site) and Tad-IR (containing a *Xba*I site) and chromosomal DNA extracted from *P. aeruginosa* DSM1707, PCR amplification was carried out using the conditions described under Materials and Methods (Section 3.2.3.2). Following agarose gel electrophoresis of the reaction mixture, an amplicon of the expected size (*ca.* 2.1 kb) was obtained when compared to the size of the DNA molecular weight marker. No amplification products were observed in the negative control in which template DNA was omitted (Fig. 3.7, lanes 2 and 3, respectively).

The agarose gel-purified amplicon was subsequently cloned into the pGEM®-T Easy vector, as described in the previous section, and restriction of the derived recombinant plasmids with both *Kpn*I and *Xba*I resulted in the excision of a 2.1-kb DNA fragment, indicating that the amplicon was successfully cloned into the pGEM®-T Easy vector (Fig. 3.7, lane 7). A recombinant clone, designated pGEM-NFIR, was selected and the integrity of the cloned insert DNA was verified by nucleic acid sequence analysis of both terminal ends. This recombinant clone was used in all subsequent DNA manipulations.

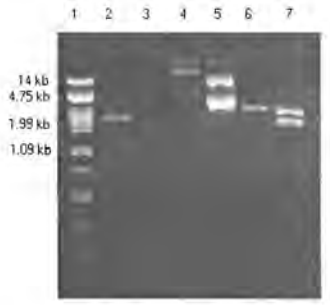

Fig. 3.6a Agarose gel electrophoretic analysis of the recombinant plasmid pGEM-IFOR constructed by cloning the Tad-IF - Tad-OR amplicon into pGEM®-T Easy vector DNA. Lane 1, DNA molecular weight marker; lane 2, DNA product obtained by PCR amplification using *P. aeruginosa* chromosomal DNA as template and primers Tad-IF and Tad-OR; lane 3, negative control PCR reaction mixture lacking template DNA; lane 4, uncut recombinant plasmid pGEM-IFOR; lane 5, uncut parental pGEM®-T Easy vector DNA; lane 6, *Eco*RI-linearized pGEM®-T Easy vector DNA; lane 7, recombinant plasmid pGEM-IFOR digested with both *Kpn*I and *Hind*III. The sizes of the DNA molecular weight marker, phage λ DNA digested with *Pst*I, are indicated to the left of the figure.

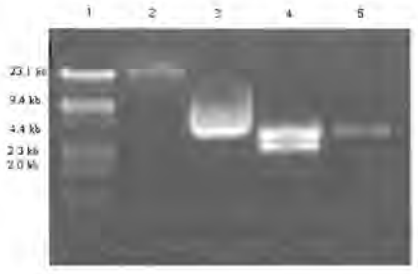

Fig. 3.6b Agarose gel electrophoretic analysis of the recombinant plasmid pUC18-IFOR. Lane 1, DNA molecular weight marker; lane 2, uncut parental pUC18 vector DNA; lane 3, uncut recombinant plasmid pUC18-IFOR; lane 4, recombinant plasmid pUC18-IFOR digested with both *Kpn*I and *Hind*III; lane 5, pUC18 vector DNA digested with *Kpn*I and *Hind*III. The sizes of the DNA molecular weight markers, phage λ DNA digested with *Hind*III, are indicated to the left of the figure.

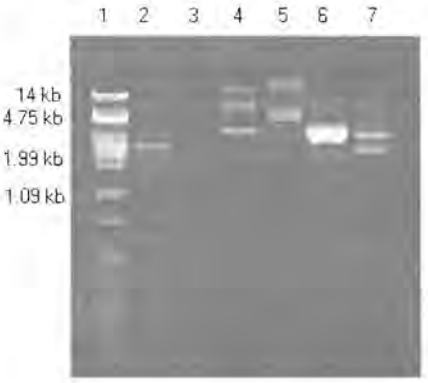

Fig. 3.7 Agarose gel electrophoretic analysis of the recombinant plasmid pGEM-NFIR constructed by cloning the Tad-NF - Tad-IR amplicon into pGEM®-T Easy vector DNA. Lane 1, DNA molecular weight marker; lane 2, DNA product obtained by PCR amplification using *P. aeruginosa* chromosomal DNA as template and primers Tad-NF and Tad-IR; lane 3, negative control PCR reaction mixture lacking template DNA; lane 4, uncut parental pGEM®-T Easy vector DNA; lane 5, uncut recombinant plasmid pGEM-NFIR; lane 6, *Eco*RI-linearized pGEM®-T Easy vector DNA; lane 7, recombinant plasmid pGEM-NFIR digested with both *Kpn*I and *Xba*I. The sizes of the DNA molecular weight marker, phage λ DNA digested with *Pst*I, are indicated to the left of the figure.

3.3.2.3 Construction of recombinant plasmid pUC18-DEF

Having successfully constructed recombinant plasmids pUC18-IFOR and pGEM-NFIR that harbor two overlapping sections of the *htpDEF* gene cluster, the final step was to join these fragments so as to obtain a single clone, pUC18-DEF, containing the full-length *htpD*, *htpE* and *htpF* ORFs. To construct plasmid pUC18-DEF (Fig. 3.8a), the cloned insert DNA was recovered from pGEM-NFIR by digestion with both *Not*I and *Kpn*I, gel-purified and then cloned into similarly prepared pUC18-IFOR vector DNA. As a consequence of this cloning strategy, all the transformants displayed a Lac$^-$ phenotype. Plasmid DNA was therefore extracted from a number of randomly selected ampicillin-resistant transformants and analyzed by agarose gel electrophoresis. Plasmid DNA migrating slower than the parental pUC18-IFOR vector DNA was selected and characterized by restriction enzyme digestion.

Digestion of the recombinant plasmid DNA with *Sma*I, which cuts once in the *htpE* ORF and once in the *P. aeruginosa* DNA, immediately downstream from the *htpF* gene, yielded expected bands corresponding to *ca.* 5.4 and 1.4 kb, respectively (Fig. 3.8b, lane 2). Whereas digestion of the recombinant plasmid DNA with both *Not*I and *Kpn*I yielded two DNA fragments corresponding in size to the pUC18-IFOR vector DNA (4.7 kb) and the cloned insert DNA (2.1 kb) (Fig. 3.8b, lane 3), digestion of the recombinant plasmid DNA with both *Eco*RI and *Hind*III, which flank the insert DNA in the MCS of pUC18, excised a DNA fragment with an expected size of *ca.* 4.1 kb (Fig. 3.8b, lane 4). To furthermore confirm the successful construction plasmid pUC18-DEF, the recombinant plasmid DNA was digested with *Pst*I, which cuts once in both the *htpD* and *htpF* ORFs and once in the *P. aeruginosa* DNA, immediately downstream from the *htpF* ORF (Fig. 3.8a). Agarose gel electrophoresis of the digestion products indicated the presence of three DNA fragments of 4.5, 1.9 and 0.368 kb (Fig. 3.8b, lane 5). The sizes of these DNA fragments were in agreement with those predicted from a map of the recombinant plasmid DNA. The 368-kb band was barely visible due to diffusion of the small DNA fragment from the agarose gel during electrophoresis.

Based on the above results, it was concluded that plasmid pUC18-DEF harbors a 4.1-kb insert DNA containing the intact *htpDEF* ORFs together with 417 bp of upstream and 595 bp of downstream *P. aeruginosa* DNA. This plasmid was used in the construction of allelic exchange vectors pUC18-htpE-Gent and pUC18-htpDEF-Gent (Section 3.3.3), as well as in the construction of the complementation plasmid pJB-DEF-Kan (Section 3.3.6).

3.3.2.4 Sequence analysis

The integrity of the insert DNA cloned into pUC18-DEF was verified by determining the nucleotide sequence using automated DNA sequencing procedures as described in Section 3.2.9. To facilitate sequencing of the insert DNA, internal primers were designed (Table 3.2 and Fig. 3.1) and used in addition to the universal pUC/M13 forward and reverse sequencing primers. Analysis of the deduced amino acid sequence of the cloned 4.1-kb insert DNA indicated that the HtpD and HtpF proteins each displayed single amino acid dissimilarities when compared to the corresponding published sequences. In the case of HtpD, a polar uncharged threonine residue was replaced with a nonpolar isoleucine residue at position 37, while in HtpF, a nonpolar alanine residue was replaced with a nonpolar proline residue at position 294. No other differences were noted.

3.3.3 Construction of different allelic exchange vectors

With a view towards elucidating the effects caused by the lack of the HtpD, HtpE and HtpDEF proteins, attempts were made to construct *P. aeruginosa* strains mutagenized in their *htpD*, *htpE* or *htpDEF* ORFs. For this purpose, pUC18-based allelic exchange vectors, harboring the *htp* ORFs disrupted through the insertion of a gentamicin resistance cassette, were constructed. As pUC18 carries the ColEI origin of replication, the plasmid is therefore unable to replicate in *P. aeruginosa*.

3.3.3.1 Construction of the allelic exchange vector pUC18-htpD-Gent

The previously constructed recombinant pUC18-Gent and pUC18-IFOR plasmids served as sources for the construction of the allelic exchange vector pUC18-htpD-Gent. The strategy for constructing the pUC18-htpD-Gent vector is indicated in Fig. 3.3, while a plasmid map of the constructed allelic exchange vector is supplied in Fig. 3.9a.

Recombinant plasmid pUC18-IFOR was partially digested with *Sal*I, which cuts three times in the *htpD* ORF and once in the *P. aeruginosa* DNA, immediately upstream from the *htpD* ORF. Thus, partial restriction digestion of pUC18-IFOR with *Sal*I resulted in several DNA fragments corresponding in size to *ca.* 4.8, 4.6, 4.5, 4.4, 4.1, 3.9 and 3.8 kb. The restriction fragments of between 4.4 to 4.8 kb, representing plasmid pUC18-IFOR DNA that had been

cut in the *htpD* ORF only, were purified from the agarose gel by making use of a silica suspension. The resulting fragment mixture was then ligated with the gentamicin resistance cassette, which had been recovered from plasmid pUC18-Gent by digestion with *Sal*I. Following transformation of competent *E. coli* DH5α cells with the ligation reaction mixture, plasmid DNA was extracted from the gentamicin-resistant transformants and characterized by agarose gel electrophoresis and restriction enzyme digestion.

Digestion of the recombinant plasmid DNA with *Sal*I yielded DNA fragments of 4.8, 1.2, 0.6 0.2 and 0.1 kb, respectively, of which the latter two bands were not clearly visible on the agarose gel. The size of the 1.2-kb DNA fragment corresponded to the expected size of the gentamicin cassette, thus confirming the successful cloning of the gentamicin cassette (Fig. 3.9b, lane 3). Digestion of the recombinant plasmid DNA with both *Eco*RI and *Hind*III yielded DNA fragments of 2.7, 2.2 and 1.1 kb (Fig. 3.9b, lane 4). The latter two bands represent DNA fragments resulting from digestion of the 3.3-kb insert DNA at an *Eco*RI recognition site located at the 5'end of the gentamicin resistance cassette. Based on the size of these restriction DNA fragments, it was also possible to conclude that the gentamicin resistance cassette was indeed cloned into the *Sal*I site located at nucleotide position 866 in the *htpD* ORF (Fig. 3.9a). The recombinant plasmid was designated pUC18-htpD-Gent. Insertion of the gentamicin cassette into the *htpD* ORF of pUC18-htpD-Gent resulted in 949 bp of upstream and 1.2 kb of downstream *P. aeruginosa* DNA flanking the insertion.

3.3.3.2 Construction of the allelic exchange vector pUC18-htpE-Gent

The recombinant pGEM-Gent and pUC18-DEF plasmids were used as the sources for constructing the pUC18-htpE-Gent allelic exchange vector. The strategy for construction of this allelic exchange vector is indicated in Fig. 3.3, while a plasmid map of the constructed pUC18-htpE-Gent vector is supplied in Fig. 3.10a.

To obtain a full-length copy of the *htpE* ORF, plasmid pUC18-DEF was digested with *Pst*I, which cuts once in the *htpD* and *htpF* ORFs and once in the *P. aeruginosa* DNA, immediately downstream from the *htpF* ORF (Fig. 3.8a). As expected, complete digestion of pUC18-DEF yielded three DNA fragments of 4.5, 1.9 and 0.368 kb, respectively. The 1.9-kb DNA fragment, containing the full-length *htpE* ORF, was purified from the agarose gel and

cloned into *Pst*I-digested pUC18 vector DNA. A recombinant clone from which an insert of the expected size was excised by *Pst*I digestion was designated pUC18-htpE. To construct the allelic exchange vector, pUC18-htpE was linearized by digestion with *Not*I, which cuts once in the *htpE* ORF (Fig. 3.10, lane 4). The linearized vector DNA was purified from the agarose gel and ligated to the gentamicin resistance cassette, which had been recovered from pGEM-Gent by *Not*I restriction enzyme digestion. Following transformation of competent *E. coli* DH5α cells with the ligation reaction mixture, the plasmid DNA extracted from gentamicin-resistant transformants was characterized by agarose gel electrophoresis and restriction enzyme digestion.

The recombinant plasmid DNA was characterized by digestion with *Not*I, and with both *Xba*I and *Hind*III. Whereas digestion of the recombinant plasmid DNA with *Not*I resulted in the excision of a 1.2-kb DNA fragment corresponding in size to the gentamicin cassette (Fig. 3.10b, lane 5), digestion with both *Xba*I and *Hind*III, which flank the insert DNA in the MCS of pUC18, excised a 3.1-kb DNA fragment (Fig. 3.10b, lane 6). This corresponded to the size of the gentamicin cassette together with 483 bp of upstream and 1.403 kb of downstream *P. aeruginosa* DNA flanking the insertion. One of the recombinant clones was selected for further use and designated pUC18-htpE-Gent.

3.3.3.3 Construction of the allelic exchange vector pUC18-htpDEF-Gent

The recombinant pUC18-Gent and pUC18-DEF plasmids also served as sources for the construction of the allelic exchange vector pUC18-htpDEF-Gent, containing interrupted *htpD*, *htpE* and *htpF* ORFs. The strategy for construction of pUC18-htpDEF-Gent is indicated in Fig. 3.3, while a map of the constructed vector is supplied in Fig. 3.11a.

To construct the allelic exchange vector, plasmid pUC18-DEF was subjected to a partial restriction digestion with *Pst*I to yield two DNA fragments of *ca.* 4.8 and 1.9 kb, respectively. By making use of a partial digestion reaction, as opposed to the above complete digestion reaction, it was ensured that there would be sufficient *htpD* and *htpF* sequences flanking the insertion to allow for homologous recombination to occur efficiently. Thus, the larger of the two DNA fragments, corresponding to the size of pUC18-DEF that had been cut in the *htpD* and *htpF* ORFs only, was purified from the agarose gel and used in subsequent ligation reactions. The 1.2-kb gentamicin resistance cassette was recovered from pUC18-Gent by

digestion with *Pst*I, gel-purified and ligated to the deletion site of the pUC18-DEF vector DNA. Following transformation of competent *E. coli* DH5α cells with the ligation reaction mixture, the plasmid DNA extracted from gentamicin-resistant transformants was characterized by agarose gel electrophoresis and restriction enzyme digestion.

To verify the successful cloning of the gentamicin cassette, the recombinant plasmid DNA was digested with *Pst*I. This yielded three DNA fragments corresponding to the size of the recombinant vector DNA (4.4 kb), the gentamicin cassette (1.2 kb) and to a DNA fragment of 368 bp, respectively (Fig. 3.11b, lane 3). Digestion of the recombinant plasmid DNA with both *Kpn*I and *Hind*III resulted in the excision of a DNA fragment of 3.4 kb (Fig. 3.11b, lane 4). The size of the DNA fragment corresponded with the size of the gentamicin resistance together with 1.620 kb of upstream and 674 bp of downstream *P. aeruginosa* DNA flanking the insertion. One of the recombinant clones was selected for further use and designated pUC18-htpDEF-Gent.

3.3.4 Construction of mutant strains of *P. aeruginosa* DSM1707

Mutant strains of the wild-type *P. aeruginosa* DSM1707 strain were constructed by transforming competent cells prepared of the wild-type DSM1707 strain with the different allelic exchange plasmids (Section 3.2.11). *P. aeruginosa* strains harbouring an integrated copy of the mutant allele were selected by culturing the transformed cells in LB-broth supplemented with increasing concentrations of gentamicin prior to plating onto selective medium. The gentamicin-resistant *P. aeruginosa* mutant strains were designated DSMHtpD, DSMHtpE and DSMHtpDEF, respectively. In contrast to the DSMHtpD and DSMHtpE strains, the DSMHtpDEF strain was found to be severely growth-impaired and thus excluded from further analyses. The presence of the mutant null alleles in the DSMHtpD and DSMHtpE strains was subsequently verified by two methods, Southern blot hybridization and PCR analysis.

3.3.5 Characterization of the putative *P. aeruginosa htp* mutant strains

3.3.5.1 Southern blot analysis

To determine whether the gentamicin resistance cassette was present in the genome of the mutant DSMHtpD and DSMHtpE strains, and whether integration of the mutant alleles occurred by means of a single or double cossover event, Southern blot analysis was performed (Section 3.2.12). The chromosomal DNA of strains DSMHtpD and DSMHtpE was thus isolated, digested with *Eco*RI and separated by agarose gel electrophoresis. The DNA fragments were transferred to nylon membranes by capillary blotting and the membranes were then hybridized with DIG-dUTP labeled DNA probes specific for either the gentamicin gene (Fig. 3.12a) or pUC18 vector DNA (Fig. 3.12b). In these analyses, recombinant plasmid pUC18-Gent and *Eco*RI-linearized pUC18 plasmid DNA were included as positive hybridization controls, while *Eco*RI-digested chromosomal DNA of the wild-type *P. aeruginosa* DSM1707 strain was included as a negative hybridization control.

The results indicated that the probe specific for the gentamicin gene (Fig. 3.12a) hybridized with plasmid pUC18-Gent, which harbours a cloned copy of the gentamicin resistance cassette, as well as with a DNA restriction fragment from each the DSMHtpD and DSMHtpE chromosomal DNA. The probe did, however, not hybridize with the *Eco*RI-digested chromosomal DNA of the *P. aeruginosa* DSM1707 strain. Probing of a similarly prepared membrane indicated that the probe specific for the pUC18 vector DNA hybridized with *Eco*RI-linearized pUC18 as well as with a DNA fragment of the *Eco*RI-digested chromosomal DNA prepared from DSMHtpD (Fig. 3.12b, lane 3). However, the labeled probe did not hybridize with the digested chromosomal DNA of either the wild-type DSM1707 or mutant DSMHtpE strains (Fig. 3.12b, lanes 2 and 4, respectively). From these results it was thus concluded that a single copy of the mutant *htpD*::*Gent*^{R} allele was integrated into the chromosomal DNA of DSMHtpD by a single crossover event, while a single copy of the mutant *htpE*::*Gent*^{R} allele was integrated into the chromosomal DNA of DSMHtpE by means of a double crossover event.

3.3.5.2 PCR analysis of *P. aeruginosa* mutant strains

The presence of integrated copies of the mutant alleles in the respective mutant *P. aeruginosa* strains was also investigated by different PCR analyses. Different pairs of oligonucleotide primers were used to amplify hybrid products only if the gentamicin resistance cassette was located within the chromosomal-borne inactived *htp* ORFs. These products represented the junction between the gentamicin gene and the disrupted *htp* ORF. In the second analysis, primers were used that annealed to sequences flanking the ORF in which the mutant allele was integrated.

3.3.5.2.1 Analysis of DSMHtpD

Southern blot analysis using pUC18 as labeled probe indicated that the pUC18-htpD-Gent allelic exchange vector integrated into the genome of DSMHtpD via a single crossover event. To determine whether integration indeed occurred within the *htpD* ORF, primers Tad-F3 and Gent-1, as well as primers Tad-GR and Gent-2 were used. These primer sets were expected to amplify hybrid products of *ca.* 3.1 kb and 7.5 kb, respectively. In the final analysis, primers Tad-F3 and Tad-GR were used to amplify either a 3.3-kb product in the absence of the gentamicin resistance cassette, or a 9.4-kb product in the presence of the integrated allelic exchange vector. However, except for being able to PCR-amplify the 3.1-kb hybrid product, attempts at amplifying the latter products were unsuccessful when chromosomal DNA of DSMHtpD was used as template. This may have been due to limited processivity of the *Taq* polymerase enzyme used in this analysis, which was not a polymerase enzyme specifically developed for long and accurate (LA) PCR technology. As expected, a 3.3-kb product indicative of the absence of the gentamicin cassette within the *htpD* ORF was generated, and no hybrid products were amplified when chromosomal DNA of wild-type *P. aeruginosa* DSM1707 was used as template in the respective PCR reactions (results not shown).

3.3.5.2.2 Analysis of DSMHtpE

Primers Tad-IR and Gent-1, as well as primers Tad-NF and Gent-2 were used to amplify a 1.3-kb and 3.2-kb hybrid product, respectively, only if the gentamicin resistance cassette was located within the disrupted *htpE* ORF. The respective products were produced when DSMHtpE chromosomal DNA was used as template (Fig. 3.13, lanes 3 and 4), but these

products were absent when wild-type *P. aeruginosa* DSM1707 chromosomal DNA were used as template in the respective PCR reactions (Fig. 3.13, lanes 6 and 7). In the final analysis, primers Tad-NF and Tad-IR were used to amplify either a 2.1-kb *htpE* product in the absence of the gentamicin resistance cassette, or a 3.3-kb product in the presence of the 1.2-kb cassette. As expected, a 3.3-kb product was produced when DSMHtpE chromosomal DNA were used as template. Template DNA from wild-type DSM1707 generated the 2.1-kb product indicative of the absence of the gentamicin cassette within the *htpE* gene (Fig. 3.13, lanes 2 and 5, respectively).

3.3.6 Construction of complementation plasmid pJB-DEF-Kan

To investigate whether any altered phenotypes that may be displayed by the *P. aeruginosa htp* mutant strains were due to the disruption of the respective wild-type *htp* ORFs, a complementation plasmid, containing intact copies of the *htpDEF* ORFs under transcriptional control of a *lac*I promoter, was constructed whereby the mutant strains could be complemented. The strategy used for construction of the complementation plasmid pJB-DEF-Kan is indicated diagrammatically in Fig. 3.4, and a map of the recombinant plasmid is provided in Fig. 3.14a.

Nucleic acid sequence analysis of the region in the genome of *P. aeruginosa* spanning the *htp* gene cluster indicated that the *htpDEF* genes lacked individual promoter sequences and consequently they may rather form part of a polycistronic operon (Section 2.3.2). To provide a promoter that would allow efficient transcription of the *htpDEF* genes in *P. aeruginosa* during complementation studies, the 4.1-kb DNA insert, containing the full-length *htpDEF* ORFs, was recovered from pUC18-DEF by digestion with both *Kpn*I and *Hind*III. The DNA fragment was purified from the agarose gel and ligated into identically prepared pBluescript SKII (+) vector. Following transformation of competent *E. coli* DH5α cells with the ligation reaction mixture, plasmid DNA was extracted from transformants diplaying a Lac$^-$ phenotype and analyzed by agarose gel electrophoresis and restriction enzyme digestion. A recombinant plasmid from which an insert of the expected size (4.1 kb) was excised by digestion with both *Kpn*I and *Hind*III (Fig. 3.14b, lane 3) was selected. This plasmid was designated pBlue-DEF and contained the *htpDEF* ORFs in the correct transcriptional orientation relative to the *lac*I promoter of the pBluescript SKII (+) vector.

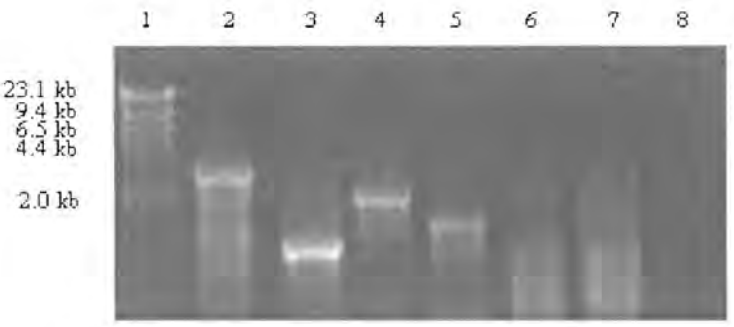

Fig. 3.13 Agarose gel electrophoretic analysis of the amplification products obtained following PCR analysis of DSMHtpE and *P. aeruginosa* DSM1707 using primers Tad-NF and Tad-IR (lanes 2 and 5), Tad-IR and Gent-1 (lanes 3 and 6), and Tad-NF and Gent-2 (lanes 4 and 7). Lanes 2, 3 and 4 represent genomic DNA from mutant strain DSMHtpE, while lanes 5, 6 and 7 represent genomic DNA from wild-type *P. aeruginosa* DSM1707. A control PCR reaction lacking template DNA was included (lane 8). The sizes of the molecular weight marker, phage λ DNA digested with *Hind*III, are indicated to the left of the figure.

Since complementation studies would also require that the plasmid DNA is capable of replicating in *P. aeruginosa*, a different plasmid to pBluescript SKII (+) had to be used, as this plasmid contains an ColEI origin of replication and is therefore unable to replicate in *P. aeruginosa*. The broad-host-range plasmid pJB3Tc20 was selected for construction of the complementation plasmid, as it contains both a vegetative origin of replication (*oriV*) that permits replication in wide variety of Gram-negative bacteria, including *E. coli* and *P. aeruginosa*, and an *oriT* that permits introduction of recombinant DNA into *P. aeruginosa* by triparental mating procedures. In addition, the plasmid also harbors tetracycline and ampicillin resistance markers (Blatney *et al.*, 1997).

Towards construction of the complementation plasmid, plasmid pBlue-DEF was digested with *Pvu*II, which cuts at the 5' end of the *lac*I promoter, and *Kpn*I, which cuts at the 3' end of the cloned insert DNA. A DNA fragment of 4.4 kb, consisting of the cloned *htpDEF* fragment (4.1 kb) and *lac*I promoter (307 bp), was purified from the agarose gel and cloned into pJB3Tc20, which had been digested with both *Eco*RV and *Kpn*I. After transformation of competent *E. coli* DH5α cells, plasmid DNA from a number of transformants was analyzed by agarose electrophoresis and restriction enzyme digestion. Since *Eco*RV and *Pvu*II generate blunt ends, the termini are compatible but the resultant hybrid site does not constitute a target for either of the restriction endonucleases. Therefore, the recombinant plasmid DNA was screened by digestion with *Sma*I, which cuts once in the tetracycline gene of the vector and three times in the cloned DNA fragment; once in the *htpE* ORF, once in the region between the *lac*I promoter and the *htpDEF* ORFs and once in the *P. aeruginosa* DNA immediately downstream from *htpF*. Digestion of the recombinant plasmid DNA yielded four DNA fragments of the expected sizes, namely 5.0, 2.7, 1.4 and 0.5 kb (Fig. 3.14c, lane 2), thus confirming the successful cloning of the *lac*I-*htpDEF* DNA fragment. A recombinant plasmid was designated pJB-DEF and used in subsequent DNA manipulations.

Since the above cloning strategy resulted in the inactivation of the vector-borne tetracycline resistance gene, and taking into account that *P. aeruginosa* is resistant towards ampicillin, a selectable marker had to be provided. A kanamycin resistance gene was subsequently cloned into plasmid pJB-DEF, which would serve as a marker to verify successful transformation of *P. aeruginosa* cells with the complementation plasmid. As the pUC4K plasmid (Taylor *et al.*, 1988), which served as source of the kanamycin gene, did not have the appropriate restriction enzyme recognition sites that would permit construction of the desired clone, the kanamycin

gene was recovered from pUC4K by digestion with *Eco*RI and cloned into pBluescript SKII (+) to yield pBlue-Kan (Fig. 3.14b, lane 4). To complete the construction of the complementation plasmid pJB-DEF-Kan, the kanamycin resistance gene was subsequently excised from pBlue-Kan by digestion with *Pvu*II and cloned by blunt-end ligation into pJB-DEF that had been digested with *Sca*I, which cuts once only in the vector-borne ampicillin resistance gene (*bla*). Following transformation of competent *E. coli* DH5α cells, plasmid DNA was extracted from kanamycin-resistant transformants and characterized by agarose gel electrophoresis and restriction enzyme digestion. Digestion of the recombinant plasmid DNA with *Eco*RI yielded DNA fragments of 4.1, 3.7, 2.2 and 1.3 kb, respectively (Fig. 3.14c, lane 3). These results confirmed that the 1.3-kb kanamycin gene was successfully cloned.

The complementation plasmid pJB-DEF-Kan therefore harbours a 4.4-kb DNA fragment containing the intact *htpDEF* ORFs from the wild-type *P. aeruginosa* strain under transcriptional control of a strong constitutive *lac*I promoter from the pBluescript SKII (+) vector, as well as a kanamycin resistance gene as selectable marker. This plasmid was consequently used as a vector equivalent for pUC18-htpD-Gent, pUC18-htpE-Gent and pUC18-htpDEF-Gent. The plasmid was introduced into the constructed *P. aeruginosa* DSMHtpD and DSMHtpE mutant strains by triparental mating as previously described (Section 3.2.11.3). Strains that displayed resistance to both kanamycin and gentamicin were consequently selected for further use in this study.

3.4 DISCUSSION

The construction of isogenic mutant strains from which specific functions have been eliminated is central to the analysis of various questions in microbiology. This approach has been used successfully to determine the genetics of biofilm formation in different bacterial species (Heilmann *et al.*, 1996; O'Toole and Kolter, 1998; Pratt and Kolter, 1998; Loo *et al.*, 2000), aspects relating to bacterial pathogenesis (Hensel and Holden, 1996) and to determine the precise function of undefined ORFs (Nika *et al.*, 2002). To generate such complete loss-of-function mutations, both random transposon insertion mutagenesis and allelic exchange methods have been useful.

Transposons, being mobile genetic elements, have the capability of inserting themselves into genes on a bacterial chromosome or plasmid, thereby disrupting the gene itself and sometimes additional genes that are encoded downstream of the mutated gene. Nevertheless, this represents a powerful approach towards identifying genes involved in a specific function provided that an appropriate high-throughput screen is available. Using such an approach, Kachlany *et al.* (2000) identified a cluster of seven genes (*tadA-F*) as being required for tight adherence of *A. actinomycetemcomitans* to abiotic surfaces. Loss of function of any of the *tad* genes severely diminished adherence to surfaces and the mutant cells no longer produced bundles of pili and failed to autoaggregate. Furthermore, mutations within the *flp-1* gene, which encodes the major pilus subunit (Inoue *et al.*, 1998), caused a similar phenotype (Kachlany *et al.*, 2001).

Open reading frames with predicted products that are significantly similar to those of *A. actinomycetemcomitans* genes *flp*, *tadA*, *tadB* and *tadC* were identified in the genome of *P. aeruginosa* PAO1 using an *in silico* approach (Chapter 2). The homologous genes were termed *htpP*, *htpD*, *htpE* and *htpF*, respectively, and appeared to form part of a larger polycistronic operon consisting of nine ORFs that excludes the putative pilus subunit-encoding gene, *htpP*. Towards determining the importance of the *P. aeruginosa htp* gene cluster in the ability of *P. aeruginosa* to attach to surfaces, allelic exchange was used in this study as an alternative to transposon mutagenesis for constructing mutant strains. Allelic exchange involves using a suicide plasmid that is unable to replicate in the studied strain to deliver an *in vitro*-inactivated or -modified allele of the gene in the chromosome. Mutations made by allelic exchange are thus targeted, therefore making it a more attractive method of mutagenesis than random transposon insertion mutagenesis. However, as with transposon mutations, it is possible to introduce polar mutations downstream of the insertion site.

The *htpD*, *htpE* as well as *htpDEF* ORFs were targeted for insertional inactivation as homologues of these ORFs are consistently found in similar pilus assembly and secretion systems of various different Gram-negative bacteria (Kachlany *et al.*, 2000; Skerker and Shapiro, 2000; Planet *et al.*, 2001; Nika *et al.*, 2002). Whereas HtpD may function as an NTPase that could play a role as the energizer of macromolecular transport in type II and type IV secretion systems (Hobbs and Mattick, 1993; Possot and Pugsley, 1994; 1997; Burns, 1999; Chapter 2 of this study), HtpE and HtpF are proteins of unknown function, although they are both predicted to be integral membrane proteins. To enable the generation of the

desired mutant strains, allelic exchange vectors were constructed by disrupting the *htpD*, *htpE* and *htpDEF* ORFs through insertion of a gentamicin resistance cassette flanked by transcriptional and translational stop signals (Luckow *et al.*, 1993). The allelic exchange vectors were subsequently introduced into *P. aeruginosa* DSM1707 by transformation of competent cells and presumptive mutant strains were selected following culturing on a selective medium, and finally characterized by Southern blot hybridization and by PCR analyses.

The results obtained from these analyses indicated that the insertional inactivation of the *htpD* ORF in the mutant DSMHtpD strain occurred by means of a single crossover event, thereby resulting in the integration of the complete allelic exchange vector. In contrast, insertional inactivation of the *htpE* ORF in the mutant DSMHtpE strain occurred by means of a double crossover event and consequently resulted in the integration of only the mutant *htpE* allele. Due to the strong likelihood of polar mutations occurring, the mutant DSMHtpE strain would thus be expected to successfully express the *htpABCD*, but not the *htpE* through *htpI* ORFs. In the case of the mutant DSMHtpD strain, polar mutations downstream of the insertion site would result in a lack of expression of the *htpD* through *htpI* ORFs, but successful expression of the *htpABC* ORFs. Nevertheless, it should be noted that polar mutations, should it occur, would have no effect on the outcome tested in this study, *i.e.* determining the importance of the *htp* gene cluster in biofilm formation by *P. aeruginosa* by making use of mutant strains lacking a functional *htp* system. However, information regarding the importance of HtpD, a probable type II/type IV NTPase, in the putative pilus biogenesis system may be obtained by comparing the DSMHtpD to the DSMHtpE strain. Although a third mutant strain, DSMHtpDEF, was also constructed, it was severely growth-impaired, requiring up to 48 h of culturing prior to obtaining visible growth in broth cultures. The mutant DSMHtpDEF strain was thus excluded from all further investigations since such slow growth, rather than inactivation of the *htpDEF* ORFs, may account for potentially altered phenotypes displayed by the mutant strain.

Not only is the construction of mutant strains required for investigating the functional importance of the *htp* gene cluster, but also an important step in such investigations would be to complement the mutations in the constructed *P. aeruginosa* mutant strains by providing the wild-type ORFs *in trans* on a recombinant plasmid. Should the mutant phenotypes be complemented to wild-type levels, it would provide strong evidence that the altered

phenotype is due to disruption of the specific wild-type *htp* ORF under investigation. Consequently, a complementation plasmid was constructed by cloning intact copies of the wild-type *htpDEF* ORFs under transcriptional control of a strong constitutive promoter into plasmid pJB3Tc20, a broad-host-range plasmid that would permit extrachromosomal replication in *P. aeruginosa* to a high copy number (Blatney *et al.*, 1997), thereby resulting in possible overexpression of the HtpD, HtpE and HtpF proteins.

The construction of the *P. aeruginosa* DSMHtpD and DSMHtpE mutant strains as well as complementation plasmid pJB-DEF-Kan, as detailed in this Chapter, provided the tools necessary to undertake further studies regarding the functional importance of the *P. aeruginosa htp* gene cluster. The details of these investigations and the results that were obtained during the course of these investigations are given in the following chapter.