JOURNAL OF BACTERIOLOGY, Aug. 1988, p. 3738–3741
0021-9193/88/083738-04$02.00/0

Vol. 170, No. 8

NOTES

Two Unusual Pilin Sequences from Different Isolates of *Pseudomonas aeruginosa*

BRITTAN L. PASLOSKE, PARIMI A. SASTRY, B. BRETT FINLAY,† AND WILLIAM PARANCHYCH*

Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received 16 February 1988/Accepted 25 April 1988

The pilin genes of two *Pseudomonas aeruginosa* strains isolated from two different patients with cystic fibrosis were cloned and sequenced. The predicted protein sequences of these two pilins had several unusual features compared with other published *P. aeruginosa* pilin sequences.

Pseudomonas aeruginosa is an opportunistic pathogen which readily infects hosts who have been immunocompromised by cancer, thermal injury, or cystic fibrosis (3). Polar pili, produced by this organism, are long, filamentous appendages which have been shown to be virulence factors by virtue of their ability to promote adherence to epithelial tissue and thus establish colonization (18, 26). The subunit monomer of pili, pilin, has a molecular weight of about 15,000. These subunits are arranged in a helical configuration of five subunits per turn in the assembled pilus (25). The *P. aeruginosa* pilins appear to be part of a conserved family of pilins which share extensive N-terminal sequence homology, and they have an unusual *N*-methylphenylalanine (NMePhe) (7, 9) as the first amino acid in the mature pilin. This class of pilin is often referred to as NMePhe (17) or type 4 (13) pili and includes the pili of *P. aeruginosa*, *Moraxella bovis*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *Bacteroides nodosus*. We now present the DNA sequences and the predicted primary structures of two other *P. aeruginosa* NMePhe pilins which have several different features compared with any of the other published *P. aeruginosa* pilin sequences.

P. aeruginosa P1 was originally cultured from the sputum of a patient with cystic fibrosis in the University of Minnesota Hospital, Minneapolis (24). It was generously provided by D. P. Speert, British Columbia Children's Hospital Research Centre, Vancouver, British Columbia, Canada. *P. aeruginosa* K122-4 was isolated from a cystic fibrosis patient in Toronto and was provided by R. T. Irvin, Department of Botany and Microbiology, Erindale College, University of Toronto, Missisauga, Ontario, Canada.

The chromosomal DNAs from these two strains were isolated, subjected to several restriction enzyme digestions, fractionated, and blotted onto nitrocellulose filter paper as previously described (20). The blots were allowed to hybridize to the nick-translated 610-base-pair *Pst*I fragment from strain PAK which codes for the first 95 amino acids of mature pilin and also has about 320 base pairs of DNA upstream of the pilin gene (20). High-stringency hybridization conditions were used, and the washes were performed at room temperature before the blots were subjected to autoradiography (12).

Interestingly, strain P1 was the first isolate which did not contain the pilin gene in a 1.2-kilobase (kb) *Hin*dIII fragment. The *Hin*dIII fragment was greater than 21 kb and therefore was not useful for cloning and sequencing. Instead, a 3.5-kb *Hin*dIII-*Bam*HI fragment was cloned into the vector pUC19 (chimera pP1-001) by shotgun cloning methodology (20). *Escherichia coli* DH5α [K-12 F$^-$ *endA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1* λ$^-$ *recA1 gyrA96 relA1* φ80d*lacZ* ΔM15] was purchased from Bethesda Research Laboratories, Burlington, Ontario, Canada, and was the strain used for DNA transformations. A 1.5-kb *Sph*I-*Pst*I fragment containing the entire pilin gene was subcloned into pUC19 (chimera pP1-005) with *E. coli* JM109 (27) used for the transformations. A 730-base-pair *Sph*I fragment directly upstream of the *Sph*I-*Pst*I fragment was subcloned from pP1-001 into pUC19 to create pP1-006. This clone was constructed to obtain sequence information upstream of the P1 pilin gene.

On the other hand, strain K122-4 had a 1.2-kb *Hin*dIII fragment which hybridized to the PAK probe. The shotgun strategy used to clone the P1 pilin gene was also used to clone the 1.2-kb *Hin*dIII fragment with the vector pUC19 (chimera pK122-4) and *E. coli* JM109.

The recombinant plasmids pP1-005, pP1-006, and pK122-4 were isolated in large quantities by CsCl density centrifugation, and the pilin genes were sequenced by the Sanger dideoxy method (20). Both strands of both pilin genes were sequenced. The enzymes *Sau*3A and *Sph*I were used to subclone the P1 pilin gene, and two synthetic oligonucleotides encompassing nucleotide positions 263 to 279 and 664 to 680 (Fig. 1) were used as primers to obtain overlaps and the sequence of the opposite strand. *Alu*I, *Sau*3A, and *Hin*dIII were the only restriction enzymes required to subclone the K122-4 pilin gene for the complete sequence. The DNA sequences and the predicted protein sequences of the P1 and K122-4 pilin genes are shown in Fig. 1. Within the two DNA sequences, starting at approximately nucleotide position 120, are the predicted NtrA-dependent promoters as previously recognized by Johnson et al. (11). The putative NtrA-dependent promoter is a region 10 to 26 base pairs upstream of the transcription start site, where the *ntrA* gene product, acting as a sigma factor, is believed to be required for transcription initiation. *Pseudomonas* NtrA-dependent promoters have also been discovered for the *xylABC* operon

* Corresponding author.
† Present address: Department of Medical Microbiology, School of Medicine, Stanford, CA 94301.

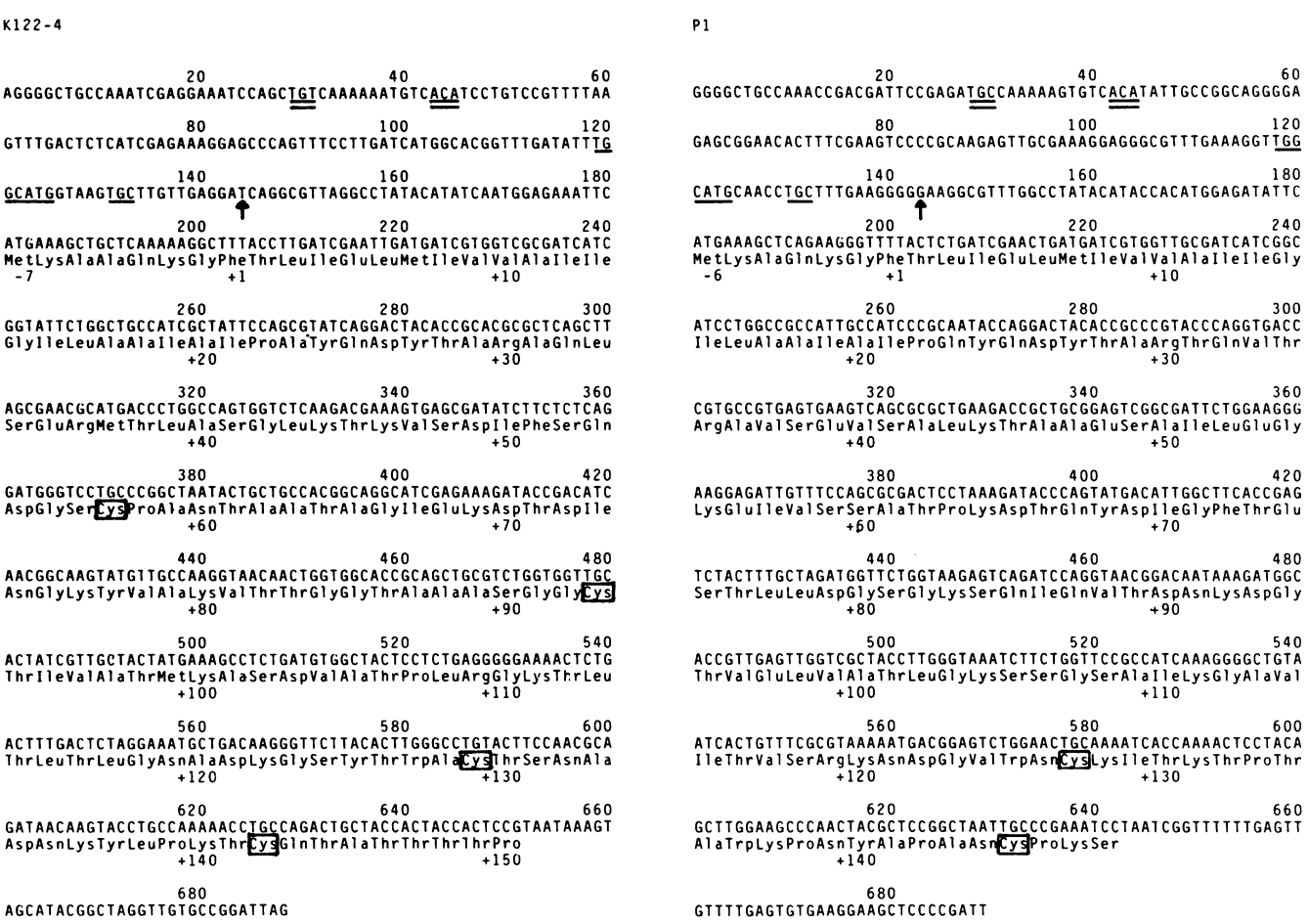

FIG. 1. Nucleotide sequences and the predicted amino acid sequences of the pilins from *P. aeruginosa* P1 and K122-4. Nucleotides are numbered along the top of the sequences, and amino acids are numbered along the bottom of the sequences. Negative numbering indicates the amino acid positions within the putative leader peptides of the two pilins. Arrows indicate the putative transcription initiation sites, and single-underlined sequences are the consensus sequences for the putative NtrA-dependent promoters (positions 119 to 133 for K122-4 and 118 to 132 for P1) (11). Double-underlined regions indicate the putative NifA recognition sites (positions 29 to 45 for K122-4 and 28 to 44 for P1) (8). The predicted cysteine residues of the two pilins are boxed.

of the TOL plasmid (6) and the carboxypeptidase G2 gene (15) of *Pseudomonas putida*. In many enteric bacteria, genes involved in nitrogen assimilation, such as *glnA*, as well as the nitrogen fixation genes in organisms like *Klebsiella pneumoniae*, have NtrA-dependent promoters, and these genes are regulated by the gene products encoded by *ntrC* and *nifA*, respectively (8). NtrC and NifA both recognize specific sequences upstream of the genes that they positively regulate. Analysis of the DNA upstream of the pilin genes from *P. aeruginosa* PAK (22), PAO (22), PA103 (11), P1, and K122-4 revealed a conserved sequence among all five strains which was located approximately 114 nucleotides upstream of their putative transcription initiation sites. This sequence had a strong homology with the NifA recognition sites of the nitrogen fixation genes which are normally located 100 to 150 nucleotides upstream of the transcription initiation site (Fig.

TABLE 1. Putative NifA recognition sites from the pilin genes of five *P. aeruginosa* isolates

Strain (reference)	Sequence[a]	Position of consensus G[b]
Consensus (4)	T G T . . . . . A C A	−100 to −150
PAK (22)	G G A G T C C G G C T G T C A A A A A G T G T C A C A T C C T G T C C G T	−113
PAO (22)	G A A A T C C A G C T G T C A A A A A A T G T C A C A T C C T G T C C G T	−114
PA103 (11)	G A A A T C C A G C T G T C A A A A A A T G T C A C A T C C T G T C C G T	−114
K122-4	G A A A T C C A G C T G T C A A A A A A T G T C A C A T C C T G T C C G T	−114
P1	G A T T C C G A C A T G C C A A A A A G T G T C A C A T A T T G C C G G C	−114

[a] Underline nucleotides indicate differences from the predominant nucleotide of that position.
[b] Relative to the putative transcriptional initiation sites at positions 143 and 144 for P1 and K122-4, respectively (Fig. 1).

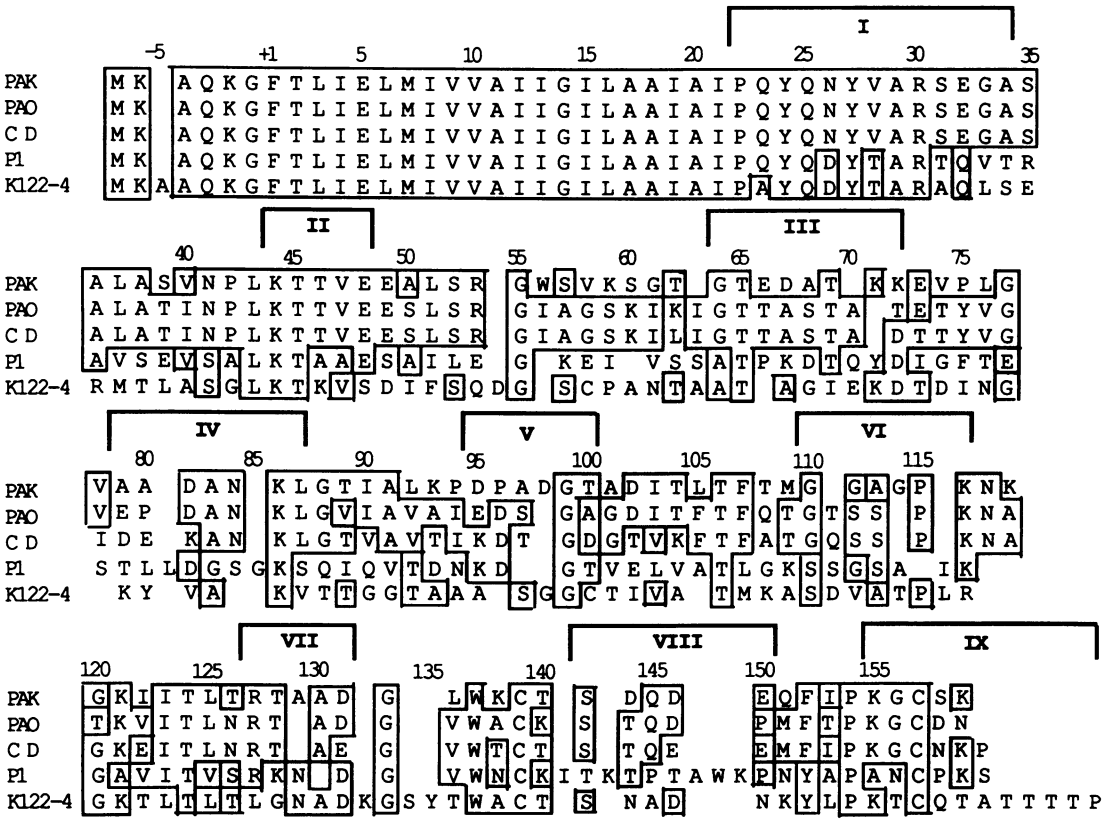

FIG. 2. Comparison of the five pilin sequences of *P. aeruginosa* PAK (22), PAO (22), CD (21), P1, and K122-4 and their predicted surface domains. The pilin sequences were aligned to obtain the greatest possible homology by visual examination on the basis of complete identity and by using the amino acid similarity values of Bacon and Anderson (2). Empty spaces were allowed for best possible alignment, and positions occupied by identical amino acids are boxed. The numbering of the amino acids was based on the total number of positions including spaces rather than on the numbering of any particular pilin as a standard. The overlined sequences indicate predicted surface domains I through IX (see text) as determined by the program of Parker et al. (19).

1) (4). Table 1 outlines the hypothetical NifA recognition sites associated with each pilin gene. These sequences are identical to the NifA-binding site consensus sequence of TGT-N_{10}-ACA (4) except for the difference in the spacing between the conserved trinucleotides, so that N equals 11 instead of 10. Since *P. aeruginosa* is not known to have a *nifA* gene, an *ntrC*-type regulatory gene may be responsible for this regulation. Pilin expression and nitrogen assimilation (10) in *P. aeruginosa* may both be regulated by an NtrC-like protein which associates with the NifA-like recognition site, since the *ntrC* and *nifA* genes are often interchangeable for gene activation (8).

In Fig. 2, the predicted amino acid sequences of the five known *P. aeruginosa* pilins have been aligned to obtain the best homology possible. There are some features in the K122-4 and P1 pilin sequences which make them distinctly different from the other published pilin sequences. Both of these sequences diverge from each other and the other pilin sequences at amino acid position 31, while pilin sequences of strains PAK, PAO, and CD differ after position 55. This point of divergence at position 31 is similar to that of the NMePhe pilin sequences of other genera (5). The strain P1 pilin is different from the other four pilins in that it contains the largest putative C-terminal disulfide loop, which contains 17 residues, compared with 12 in other strains.

The K122-4 pilin has three characteristics which make it unusual: (i) it has seven residues instead of six as a putative leader peptide; (ii) it is the longest NMePhe pilin found in *P. aeruginosa* to date, with 150 residues in its putative mature form; and (iii) it has two additional cysteines at positions 58 and 100. Similarly situated cysteines within the pilin of *B. nodosus* 198 were shown to be in the reduced state (14), suggesting that the analogous cysteines of K122-4 pilin may also be in the reduced state.

It is worth noting that certain amino acids besides the hydrophobic stretch at the N terminus appear to be conserved in all five sequences: Leu-43; Lys-44 and -86; Thr-45, -65, -106, and -124; Gly-55, -99, and -133; Trp-136; Cys-139 and -157; Pro-154; and the aromatic amino acids at position 152. These residues may be important in maintaining a basic tertiary structure required for pilus assembly.

The pilins of strains PA103 (11) and PA1244 (J. Sadoff, personal communication) have sequences almost identical to those of CD (21) and P1 pilins, respectively. Moreover, we have sequenced other isolates with pilins nearly identical to those of strains PAK and K122-4. Interestingly, the strains which had closely related pilin sequences were isolated from patients in widely separated locations throughout North America. These data suggest that the number of unique pilin types may be fairly limited, perhaps 10 to 15. This number would be in keeping with results for *B. nodosus* (1) and *M. bovis* (16), in which there are fewer than 10 pilus serogroups for each pathogen.

The predicted surface domains for each of the five pilins were determined by the program of Parker et al. (19) to establish the potential antigenic sites within each pilin sub-

unit. These surface plots were then studied to locate regions where the surface residues of all five pilins overlapped (Fig. 2). The surface epitopes predicted on this basis were the residues at positions 22 to 34 (Fig. 2, domain I), 44 to 48 (II), 64 to 72 (III), 79 to 87 (IV), 95 to 100 (V), 110 to 117 (VI), 127 to 131 (VII), 142 to 150 (VIII), and 155 to the C terminus (IX). It is of interest that the strongest predicted surface epitopes (domains III, V, and VII) were in similar regions for all five pilins despite significant variations in amino acid sequence. Domains III and V, located within the central, hypervariable region of the pilins, fall within the antigenic determinants predicted by a sequence comparison of strains PAK, PAO, and CD (21), and experimental data gathered from pilin peptides confirm these predictions as well (23; P. A. Sastry and W. Paranchych, unpublished results).

This work was supported by the Medical Research Council of Canada. B.L.P. and B.B.F. were recipients of a studentship from the Alberta Heritage Foundation for Medical Research.

We thank Kathy Volpel for technical assistance, K. K. Lee for running the surface plot program, and Laura Frost for valuable discussion.

LITERATURE CITED

1. **Anderson, B. J., J. S. Mattick, P. T. Cox, C. L. Kristo, and J. R. Egerton.** 1987. Western blot (immunoblot) analysis of the fimbrial antigens of *Bacteroides nodosus*. J. Bacteriol. **169**:4018–4023.
2. **Bacon, D. J., and W. F. Anderson.** 1986. Multiple sequence alignment. J. Mol. Biol. **191**:153–161.
3. **Bodey, G. P., R. Bolivar, V. Fainstein, and L. Jadeja.** 1983. Infections caused by *Pseudomonas aeruginosa*. Rev. Infect. Dis. **5**:279–313.
4. **Buck, M., S. Miller, M. Drummond, and R. Dixon.** 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature (London) **320**:374–378.
5. **Dalrymple, B., and J. S. Mattick.** 1987. An analysis of the organization and evolution of type 4 fimbrial (MePhe) subunit proteins. J. Mol. Evol. **25**:261–269.
6. **Dixon, R.** 1986. The *xylABC* promoter from the *Pseudomonas putida* TOL plasmid is activated by nitrogen regulatory genes in *Escherichia coli*. Mol. Gen. Genet. **203**:129–136.
7. **Frost, L. S., M. Carpenter, and W. Paranchych.** 1978. N-Methylphenylalanine at the N-terminus of *Pseudomonas aeruginosa* K pili. Nature (London) **271**:87–89.
8. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20**:567–591.
9. **Hermodson, M. A., K. C. S. Chen, and T. M. Buchanan.** 1978. *Neisseria* pili proteins: amino-terminal amino acid sequences and identification of an unusual amino acid. Biochemistry **17**:442–445.
10. **Janssen, D. B., P. M. Herst, H. M. L. J. Joosten, and C. van der Drift.** 1981. Nitrogen control in *Pseudomonas aeruginosa*: a role for glutamine in the regulation of the synthesis of NADP-dependent glutamate dehydrogenase, urease and histidase. Arch. Microbiol. **128**:398–402.
11. **Johnson, K., M. L. Parker, and S. Lory.** 1986. Nucleotide sequence and transcriptional initiation site of two *Pseudomonas aeruginosa* pilin genes. J. Biol. Chem. **261**:15703–15708.
12. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 387–389. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Mattick, J. S., M. M. Bills, B. J. Anderson, B. Dalrymple, M. R. Mott, and J. R. Egerton.** 1987. Morphogenetic expression of *Bacteroides nodosus* fimbriae in *Pseudomonas aeruginosa*. J. Bacteriol. **169**:33–41.
14. **McKern, N. M., I. J. O'Donnell, A. S. Inglis, D. J. Stewart, and B. L. Clark.** 1983. Amino acid sequence of pilin from *Bacteroides nodosus* (strain 198), the causative organism of ovine footrot. FEBS Lett. **164**:149–153.
15. **Minton, N. P., and L. E. Clarke.** 1985. Identification of the promoter of the *Pseudomonas* gene coding for carboxypeptidase G2. J. Mol. Appl. Genet. **3**:26–35.
16. **Moore, L. J., and J. M. Rutter.** 1987. Antigenic analysis of fimbrial proteins from *Moraxella bovis*. J. Clin. Microbiol. **25**:2063–2070.
17. **Paranchych, W., and L. S. Frost.** 1988. The physiology and biochemistry of pili. Adv. Microb. Physiol. **29**:53–114.
18. **Paranchych, W., P. A. Sastry, K. Volpel, B. A. Loh, and D. P. Speert.** 1986. Fimbriae (pili): molecular basis of *Pseudomonas aeruginosa* adherence. Clin. Invest. Med. **9**:113–118.
19. **Parker, J. M. R., D. Guo, and R. S. Hodges.** 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray derived accessible sites. Biochemistry **25**:5425–5432.
20. **Pasloske, B. L., B. B. Finlay, and W. Paranchych.** 1985. Cloning and sequencing of the *Pseudomonas aeruginosa* PAK pilin gene. FEBS Lett. **183**:408–412.
21. **Pasloske, B. L., A. M. Joffe, Q. Sun, K. Volpel, W. Paranchych, F. Eftekhar, and D. P. Speert.** 1988. Serial isolates of *Pseudomonas aeruginosa* from a cystic fibrosis patient have identical pilin sequences. Infect. Immun. **56**:665–672.
22. **Sastry, P. A., B. B. Finlay, B. L. Pasloske, W. Paranchych, J. R. Pearlstone, and L. B. Smillie.** 1985. Comparative studies of the amino acid and nucleotide sequences of pilin derived from *Pseudomonas aeruginosa* PAK and PAO. J. Bacteriol. **164**:571–577.
23. **Sastry, P. A., J. R. Pearlstone, L. B. Smillie, and W. Paranchych.** 1985. Studies on the primary structure and antigenic determinants of pilin isolated from *Pseudomonas aeruginosa* K. Can. J. Biochem. Cell Biol. **63**:284–291.
24. **Speert, D. P., F. Eftekhar, and M. L. Puterman.** 1984. Nonopsonic phagocytosis of strains of *Pseudomonas aeruginosa* from cystic fibrosis patients. Infect. Immun. **43**:1006–1011.
25. **Watts, T. H., C. M. Kay, and W. Paranchych.** 1983. Spectral properties of three quaternary arrangements of *Pseudomonas* pilin. Biochemistry **22**:3640–3646.
26. **Woods, D. E., D. C. Strauss, W. G. Johanson, Jr., V. K. Berry, and J. A. Bass.** 1980. Role of pili in adherence of *Pseudomonas aeruginosa* to mammalian buccal epithelial cells. Infect. Immun. **29**:1146–1151.
27. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.