

Sequence Polymorphism in the Glycosylation Island and Flagellins of *Pseudomonas aeruginosa*

Shiwani K. Arora,¹ Matthew C. Wolfgang,² Stephen Lory,² and Reuben Ramphal^{1*}

Department of Medicine and Infectious Diseases, University of Florida, Gainesville, Florida 32610,¹ and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115²

Received 23 September 2003/Accepted 11 December 2003

A genomic island consisting of 14 open reading frames, *orfA* to *orfN* was previously identified in *Pseudomonas aeruginosa* strain PAK and shown to be essential for glycosylation of flagellin. DNA microarray hybridization analysis of a number of *P. aeruginosa* strains from diverse origins showed that this island is polymorphic. PCR and sequence analysis confirmed that many *P. aeruginosa* strains carry an abbreviated version of the island (short island) in which *orfD*, *-E* and *-H* are polymorphic and *orfI*, *-J*, *-K*, *-L*, and *-M* are absent. To ascertain whether there was a relationship between the inheritance of the short island and specific flagellin sequence variants, complete or partial nucleotide sequences of flagellin genes from 24 a-type *P. aeruginosa* strains were determined. Two distinct flagellin subtypes, designated A1 and A2, were apparent. Strains with the complete 14-gene island (long island) were almost exclusively of the A1 type, whereas strains carrying the short island were associated with both A1- and A2-type flagellins. These findings indicate that *P. aeruginosa* possesses a relatively low number of distinct flagellin types and probably has the capacity to further diversify this antigenic surface protein by glycosylation.

Pseudomonas aeruginosa is a motile opportunistic Gram-negative bacterium that possesses a single polar flagellum. The flagellum of this organism is structurally similar to those of other Gram-negative bacteria and contains a relatively complex basal body and hook structure attached to a filament primarily consisting of the assembled flagellin subunit protein. Although the role of flagellar motility in virulence of many pathogenic microorganisms including *P. aeruginosa* has been well established (9, 18, 19), the recent discovery that flagellin is one of the most potent stimuli of the host innate immune response gives added significance to the expression of flagella in the context of infection. Flagellin-based signaling in a variety of cells is mediated by Toll-like receptor 5 (11), and in *P. aeruginosa*, it has been shown to be more potent than *Pseudomonas* lipopolysaccharide (17).

The flagellin proteins of *P. aeruginosa* can be classified as a-type or b-type, based on their reactions with specific polyclonal antibodies and their molecular weights (1, 2, 16). The a-type flagellins are reported to be heterologous and have variable molecular masses ranging from 45 to 52 kDa, whereas b-type flagellins are essentially conserved in sequence and have an almost invariant molecular mass of 53 kDa. However, between the two types, the N- and C-terminal domains are conserved while the central domain of the a-type is said to be hypervariable, leading to antigenic or serological variations (21, 29). The a-type flagellins have been further classified into several subtypes based on their H-antigenic components (2), nucleotide sequence variations (23), or restriction fragment length polymorphisms (20).

Totten and Lory first noted anomalous migration and heterogeneity of the a-type flagellins during sodium dodecyl sulfate-polyacrylamide gel electrophoresis that did not correspond to the predicted molecular mass (27). Similar discrepancies in the electrophoretic mobility of flagellins in several other bacteria, including *Campylobacter coli* (6), *Campylobacter jejuni* (26), and *Helicobacter felis* (14), have been also observed. These mass discrepancies have now been shown to be due to glycosylation of the proteins. Brimer and Montie (5) first demonstrated that the flagellins of some a-type *P. aeruginosa* strains were glycosylated. The genetic basis of this observation was recently demonstrated by the identification of a cluster of 14 genes localized on a glycosylation island (GI) found in a *P. aeruginosa* strain (PAK) that expresses the a-type flagellins (3).

However, other laboratories as well as ours have observed a wide range of apparent molecular weights for a-type flagellins (5; S. K. Arora, unpublished data), whose bases were unknown. Theoretically, this could be due to variability in protein sequence, since a-type flagellins have been reported to have variable molecular masses and/or differences in the extent of glycosylation. To understand the basis of this variability and its clinical significance, we analyzed the sequences of the GIs and flagellins of *P. aeruginosa* strains isolated from diverse sources. We report that the GIs of *P. aeruginosa* strains, which possess a-type flagellins, are polymorphic. A number of the *P. aeruginosa* strains that were analyzed lack approximately 5.4 kb of DNA in their GIs. In addition, sequencing of variable domains of flagellins from 24 different *P. aeruginosa* strains led to the conclusion that the a-type flagellins can be divided into two major subtypes, which we call A1 and A2 based on their amino acid sequences. This apparent limited structural repertoire is then possibly modified by glycosylation through the expression of enzymes encoded in one of the two different flagellin GIs.

* Corresponding author. Mailing address: Department of Medicine/Infectious Diseases, P.O. Box 100277, JHMC, University of Florida, Gainesville, FL 32610. Phone: (352) 392-2932. Fax: (352) 392-6481. E-mail: ramphr@medmac.ufl.edu.

MATERIALS AND METHODS

Bacterial strains, plasmids, and primers. The 100 *P. aeruginosa* isolates from common human infections were randomly selected from R. Jones' collection of *P. aeruginosa* strains that were utilized in the Global SENTRY Antimicrobial Surveillance Program. Each of these strains was isolated from individual patients. Some of the isolates used in this study and the plasmids and primers used in this study are listed in Table 1.

DNA microarray analysis. Chromosomal DNA hybridizations were performed as previously described by using GeneChip *P. aeruginosa* genome arrays (Affymetrix) (31). Hybridization intensity data were extracted from the scanned array images, and the presence or absence of each gene was determined based on the detection *P* value (α) generated by the Affymetrix Microarray Suite software, version 5.0. *P* value cutoffs were determined empirically based on the results obtained from hybridization of strain PAO1 DNA to the microarray. Probe sets with hybridization *P* values less than 0.01 were called present (black). Probe sets with hybridization *P* values greater than 0.04 were designated absent (light gray). Values between 0.01 and 0.04 were designated uncertain (darker gray).

PCR amplification and plasmid constructions. PCR was performed by using a two-step method with the *Taq*Plus Long PCR system (Stratagene). In brief, 35 cycles of 95°C for 30 s and 70°C for variable lengths of time (3 to 8 min) were run. A final extension of 10 min at 72°C was done followed by holding the temperature at 4°C. The nucleotide sequences of various primers used in this study are given in Table 1. Specifically, primer pair RER142-RER144 was used for amplifying the GI from different *P. aeruginosa* strains. Primer pair RER142-RER172 was used to amplify the 3-kb region of *P. aeruginosa* strain CF5, which appeared identical to *P. aeruginosa* strain PAK on the DNA microarray analysis. This PCR product contained *orfH* and part of *orfJ*. Similarly, a 3.68-kb PCR product was obtained by using primer pair RER171-RER144 and the DNA template from strain CF5. This PCR product contained *orfK*, *-L*, and *-M* and incomplete *orfJ* and *-N*. Both the above PCR products were gel purified and were directly utilized for sequencing part of the GI of *P. aeruginosa* strain CF5. Primer pairs AV1-RER190 and RER166-RER167 were used for amplifying 6.4-kb (containing *orfA*, *-B*, *-C*, *-D*, *-E*, *-F*, and *-G*) and 5.1-kb (containing *orfH* and *-N*) DNA fragments spanning the complete GI of *P. aeruginosa* strain JJ692. These PCR products were cloned into the TA cloning vector PCR2.1-TOPO (Invitrogen), resulting in the construction of plasmids PCR2.1AV1190 and PCR2.15kbΔGI, which were utilized for sequencing the complete GI of *P. aeruginosa* strain JJ692. For cloning these PCR products, 5 U of *Taq* polymerase was added per reaction mixture at the completion of the PCR to add additional A's which are normally removed by the proofreading activity of PFU polymerase in the *Taq*Plus Long system. This reaction was carried out at 72°C for 20 min. The treated PCR products were then directly used in the clonings. For amplifying the complete flagellin genes from different *P. aeruginosa* strains, primer pair RER150-RER170 was used and the PCR products carrying flagellins were used directly for sequencing after gel purification. Primer RER154 was used for sequencing partial flagellins from additional *P. aeruginosa* strains.

Sequencing. DNA sequencing was performed by using *Taq* DyeDeoxy Terminator and Dye Primer Cycle Sequencing protocols developed by Applied Biosystems (Perkin-Elmer Corp., Foster City, Calif.). Fluorescently labeled dideoxynucleotides and primers were used, respectively. The labeled extension products were analyzed on an Applied Biosystems model 373A DNA sequencer. Single- or double-stranded sequences were aligned and assembled by using programs in the Sequencer software package (Gene Codes Corp., Ann Arbor, Mich.).

Nucleotide sequence accession number. The GenBank accession numbers for the nucleotide sequences of the complete flagellin genes from 6 *P. aeruginosa* strains were AY275674 to AY275679, accession numbers for sequences of partial central regions of flagellins from 18 *P. aeruginosa* strains were AY278531 to AY278547, and accession numbers for sequences for the complete or partial GIs of strains JJ692 and CF5 were AY280452 and AY280453, respectively.

RESULTS

Western blot analysis of *P. aeruginosa* strains of diverse origins. *P. aeruginosa* strains isolated from cystic fibrosis (CF) patients, urinary tract infections (UTI), and the blood of patients infected with *P. aeruginosa* were analyzed by Western blotting to assess the apparent size of their flagellins and possible extent of glycosylation of their flagellins. As shown in Fig. 1, the b-type flagellins of *P. aeruginosa* strains X24509 and

UDL migrated as a sharp band of approximately 53 kDa while a-type flagellins displayed a range of apparent molecular weights which could be attributed to the variability of their flagellin sequences and/or to the extents of glycosylation. Strain CF5 did not show any immunoreactive flagellin band, consistent with the observation that it was nonmotile, as has been observed with many CF strains. To explore the above-mentioned possibilities, these *P. aeruginosa* strains were further analyzed by DNA microarrays, PCR, and DNA sequencing of selected flagellins and GIs.

DNA microarray analysis of the GIs in different *P. aeruginosa* strains. The genome content of a diverse panel of *P. aeruginosa* strains was recently examined (31) by using a whole-genome DNA microarray based on the sequenced genome of strain PAO1. This study identified a remarkably high level of conservation of genes, including virulence factors, among all isolates examined. The microarray used in this study also included probes for the sequenced GI identified in strain PAK; however, hybridization to probes for GI genes was not examined, as the island is not present in strain PAO1. We have reanalyzed the chromosomal DNA hybridization data for the set of strains that were studied by Western blotting and found that the GIs of these *P. aeruginosa* strains demonstrated a considerable degree of polymorphism (Fig. 2). Strains of the b-type, including PAO1, apparently lacked the entire GI while a-type strains had complete or partial islands. Those with partial islands showed no hybridization with probes from *orfH*, *-I*, *-J*, *-K*, *-L*, and *-M*. In addition, *orfD* and *orfK* were designated uncertain in some strains based on their *P* values. Comparison of the b-type *P. aeruginosa* strain PAO1 genome sequence with the PAK GI sequence had previously shown that PAO1 had an open reading frame (ORF) homologous to the *orfN* of the PAK GI (3), but the level of homology was too low for it to be detected by the DNA microarray hybridization with the *orfN*-specific probes.

PCR analysis of the polymorphic GIs of *P. aeruginosa* strains. The above results suggested either that a deletion exists between *orfH* and *orfM* in certain strains or that the genes spanning this region are polymorphic at the nucleotide sequence level. We explored these possibilities by PCR analysis of a number of *P. aeruginosa* strains. Two primers (RER142 and RER144) were designed by using the sequence from the GI of strain PAK to amplify a piece of DNA spanning *orfH*, *-I*, *-J*, *-K*, *-L*, and *-M* of strain PAK and other a-type strains used in the DNA microarray analysis. Consistent with the DNA microarray data, all the a-type strains which did not hybridize with *orfH*, *-I*, *-J*, *-K*, *-L*, and *-M*-specific probes from strain PAK generated a 1.4-kb PCR product, as opposed to the strains carrying the complete island, which gave a 6.8-kb PCR product, as expected. These results indicate that there was a deletion of approximately 5.4 kb in the GI of some *P. aeruginosa* strains, consistent with the existence of a shorter GI in *P. aeruginosa* (Fig. 3). Furthermore, since all the *P. aeruginosa* strains with the short GI generated a PCR product of the same size (1.4 kb), the observation of hybridizing sequences similar to *orfK* (Fig. 2) in strain U2504 suggests that a DNA segment with homology to *orfK* in this strain must be located elsewhere on its chromosome and a similar scenario is possible in strain PAO1.

Next, we addressed the question of whether *P. aeruginosa*

TABLE 1. Strains, plasmids, and primers used

Strain, plasmid, or primer	Description, origin, or sequence	Source, reference, or description
<i>P. aeruginosa</i> strains		
PAK	Wild-type laboratory strain	D. Bradley
PAO1	Wild-type laboratory strain	P. Phibbs
170002	Clinical isolate	R. Ansorg
ATCC 15691	B. W. Holloway	23
DSM 1128	Ear infection	23
JJ692	UTI isolate	S. Lory
S54485	UTI isolate	S. Lory
X24509	UTI isolate	S. Lory
UDL	UTI isolate	S. Lory
0252509	UTI isolate	R. Jones
043618	UTI isolate	R. Jones
0551701	UTI isolate	R. Jones
0031261	UTI isolate	R. Jones
025669	UTI isolate	R. Jones
0572	UTI isolate	R. Jones
0581854	UTI isolate	R. Jones
01991	UTI isolate	R. Jones
S35004	Blood isolate	S. Lory
X13273	Blood isolate	S. Lory
CS2	Blood isolate	This study
B1	Blood isolate	This study
H21651	Blood isolate	S. Lory
S29712	Blood isolate	S. Lory
CF18	CF isolate	S. Lory
CF5	CF isolate	J. Burns
CF27	CF isolate	J. Burns
CF127	CF isolate	S. Lory
CF262	CF isolate	R. Jones
CF282	CF isolate	R. Jones
CF304	CF isolate	R. Jones
CF4	CF isolate	M. Vasil
CF283	CF isolate	R. Jones
CF265	CF isolate	R. Jones
CF280	CF isolate	R. Jones
CF29	CF isolate	M. Vasil
Plasmids		
PCR2.1-TOPO	TA cloning vector, ampicillin resistance, kanamycin resistance	Invitrogen, Inc.
PCR2.1AV1190	PCR2.1-TOPO vector carrying 6.4-kb PCR fragment from JJ692 containing <i>orfA</i> to <i>orfG</i>	This study
PCR2.15KBΔGI	PCR2.1-TOPO vector carrying 5.1-kb PCR fragment from JJ692 containing <i>orfH</i> and <i>orfN</i>	This study
Primers		
RER142	GCGAATGGTTGTTGCCGGGCAGGATT	5' primer for GI
RER144	GGTCGGCATGACGATGCTGACGAGAG	3' primer for GI
RER150	ATGCGAACCTTTAGGGGTGCCTGGGGCG	5' primer for flagellin
RER170	TCGGAAGAAGACGCCGAGCCGAACAG	3' primer for flagellin
RER154	GGCACCGTCGACATCGCGAT	Primer for sequencing variable domains of flagellins
RER171	GGCTACTGTGGCATCGCCGACGAAAC	Primer for sequencing GI from CF5
RER172	GTTTCGTCGGCGATGCCACAGTAGCC	Primer for sequencing GI from CF5
RER166	CCCAAAGAGCTCGTTTGCTGAGCTT	Primer for sequencing GI from JJ692
RER167	GAACTCATGCCGATGGGGCACGCAT	Primer for sequencing GI from JJ692
AV1	AAACCCGAATTCTACGCCGGTGTGACGCGTGACGATCGA	Primer for sequencing GI from JJ692
RER190	GTTTCCCGCCATCATCGACAGCTCCC	Primer for sequencing GI from JJ692

strains of any specific clinical origin preferably had this deletion. To answer this question, we analyzed a total of 100 a-type *P. aeruginosa* isolates from some of the most common human sites of infection by PCR. These strains were isolated from CF

patients (38 isolates), UTIs (46 isolates), or the blood of patients infected with *P. aeruginosa* (16 isolates). We found that 70% of the UTI isolates and 56% of the blood isolates had the short GI while only 47% of the CF isolates had the short GI.

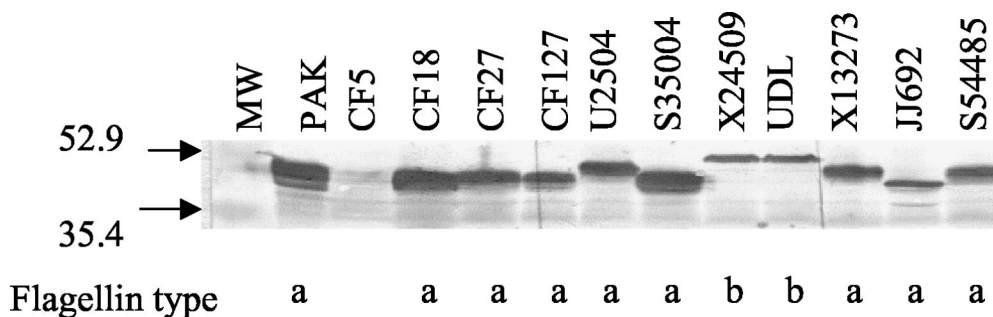


FIG. 1. Western blot analysis of *P. aeruginosa* strains. Whole-cell lysates from different *P. aeruginosa* strains were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and were subjected to Western blot analysis with FliC-specific antibodies. These antibodies were raised against purified a-type flagellar preparations, and they cross-react with the b-type flagellins. The flagellin type for each strain is indicated at the bottom of the panel. Lane 1 contained the broad-range prestained molecular weight markers which were purchased from Bio-Rad Laboratories, Richmond, Calif.

Whether the differences noted provide some advantage for urinary tract strains is a matter for further work.

Nucleotide sequence of GIs of *P. aeruginosa* strains JJ692 and CF5. To confirm the DNA microarray hybridization results and PCR analysis, we determined either partial or complete nucleotide sequences of the GIs of two *P. aeruginosa* strains, CF5, which appeared identical to strain PAK by DNA

microarray hybridization and PCR data, and strain JJ692, which showed a deletion of approximately 5.4 kb in the GI. In the case of strain CF5, a 6.8-kb DNA fragment carrying *orfH*, *-I*, *-J*, *-K*, *-L*, and *-M* was sequenced. This DNA fragment spanned the area where some *P. aeruginosa* strains had a deletion based on the DNA microarray and PCR data (Fig. 3). As shown in Table 2, *orfH*, *-I*, *-J*, *-K*, *-L*, and *-M* of CF5 were nearly identical to their corresponding homologs in PAK. In contrast, considerable polymorphism was observed in the short-island sequence of strain JJ692 compared to PAK, specifically in *orfD* and *orfH* (Table 2). Based on the DNA microarray hybridization data for JJ692, *orfH* was called absent while *orfD* was designated present. Sequencing data showed that both of the ORFs were actually present but polymorphic to different degrees, as shown in Table 2. In general, this result indicates that hybridization detection is very stringent and that the lack of hybridization signal can mean that a gene is either absent or polymorphic. On the other hand, hybridization of *orfE* was read as present, but sequencing data showed that it was prematurely terminated at 30 amino acids due to the deletion of a nucleotide resulting in a frameshift. Partial sequencing of *orfE* in another *P. aeruginosa* strain (X13273) containing a short GI showed that this strain did not have the frameshift in *orfE*, thus suggesting that the frameshift in *orfE* is not a general characteristic of *P. aeruginosa* strains carrying a short GI.

The limitation of the microarray analysis was also demonstrated by *orfH*, which was read as absent from all strains with a short island but in fact was detected by PCR and sequence analysis. Comparison of its sequence from strains PAK and JJ692 showed extensive divergence, explaining poor hybridization to the corresponding probes on the microarray. To confirm this difference in the two types of GI, we determined the nucleotide sequence of *orfH* from another *P. aeruginosa* strain (X13273), whose GI appeared to be identical to JJ692 based on the DNA microarray data. The two nucleotide sequences were 99.4% identical (data not shown), suggesting that other *P. aeruginosa* strains carrying the short GI probably retained *orfH*, but the DNA microarray data showed no hybridization due to low levels of homology.

Amino acid sequence of a-type flagellins. To examine whether there was coinherence of GI type with any specific flagellin sequences, the flagellin genes of 7 *P. aeruginosa* strains

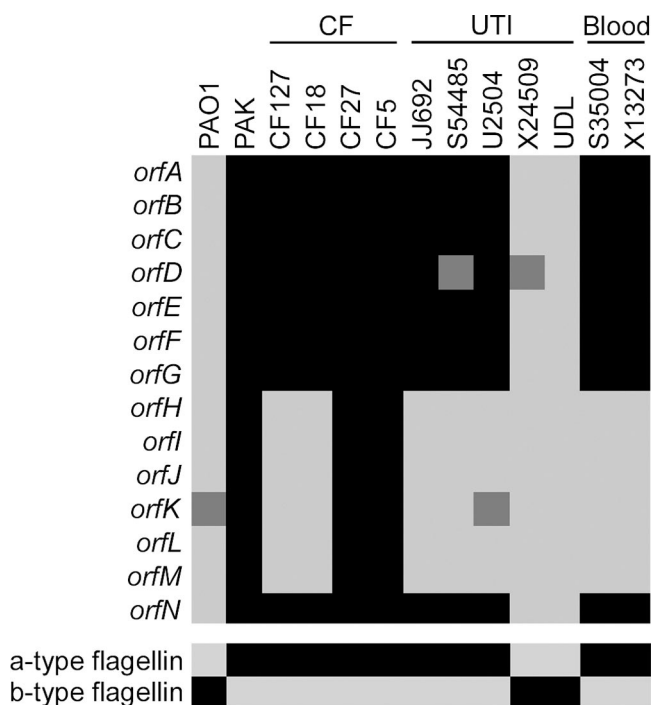


FIG. 2. DNA microarray analysis of the GI of different *P. aeruginosa* strains. The diagram suggests the presence or absence of genes in the GI of different *P. aeruginosa* strains isolated from CF patients, patients with UTIs, or the blood of infected patients. PAK and PAO1 were used as reference strains. The bottom panel shows the hybridization with the probes for the two flagellin types, a-type or b-type. The genes designated as being present are shown in black, and the genes considered to be absent are shown in light gray. The genes shown in darker gray are designated uncertain. Please note that these results only suggest the presence or absence of genes based on the various degrees of homology with the probes.

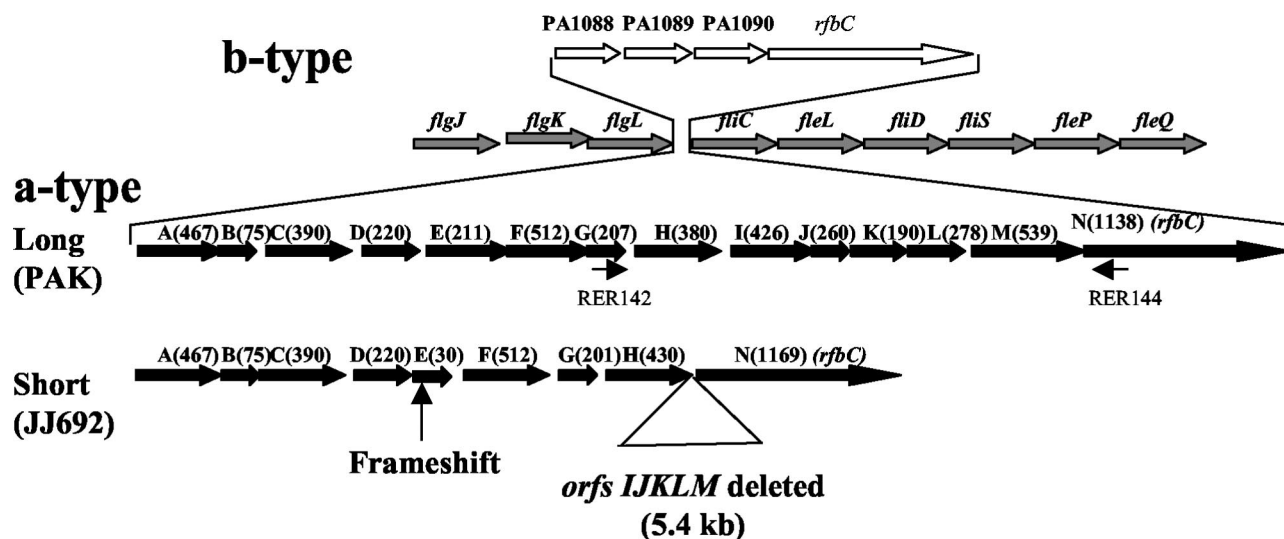


FIG. 3. Schematic diagram showing the structure of GIs of a- and b-type *P. aeruginosa* strains. The diagram shows the region of the *P. aeruginosa* chromosome where the GIs are located. The location of the GI insertion is shown to be in the middle of flagellar genes (shown by arrows with gray filling) *flgL* on the 5' end and *fljC* at the 3' end. Two a-type GIs, long (PAK) with all 14 ORFs and short (JJ692) with a 5.4-kb sequence containing deletions of *orfI*, -J, -K, -L, and -M, are shown by black filled arrows. The PAO1 GI with 3 ORFs of unknown function and an ORF which had homology to *rfbC* gene (*orfN*) is shown by arrows with no fill for comparison. The locations of the primers RER142 and RER144 used for PCR analysis of different *P. aeruginosa* strains are shown by small arrows.

containing either the complete (long) or short island were PCR amplified, and the nucleotide sequence of the complete flagellin gene was determined. These nucleotide sequences were translated, and the deduced amino acid sequences were aligned by using the PILEUP program from the Genetics Computer Group (GCG) Wisconsin package. Additionally, two more complete a-type flagellin sequences present in the GenBank database were included in this analysis. Two distinct groups could be easily identified based on this analysis (data not shown). A GROWTREE analysis of these sequences was used to generate a phylogenetic tree, clearly showing the two

flagellin subtypes as two clusters, which we called A1 and A2 (Fig. 4). One peculiar strain DSM 1128 was more like the A2 flagellins with respect to the two conserved deletions but was placed between the A1 and A2 clusters because of some similarities with the A1 flagellins (Fig. 5). The N and C termini of these flagellins were conserved, and the differences in the sequences among the two subtypes were mainly limited to the central variable domain, as reported previously for most *P. aeruginosa* strains (5, 24). The variability between the two sequences appeared to be limited; therefore, we determined the sequence of the variable domains of flagellins from an addi-

TABLE 2. Nucleotide and amino acid similarities and identities between GIs of *P. aeruginosa* strains PAK, CF5, and JJ692

ORF	Result for ^a :					
	PAK vs CF5			PAK vs JJ692		
	% Nucleotide identity	% Amino acid similarity	% Amino acid identity	% Nucleotide identity	% Amino acid similarity	% Amino acid identity
A	ND	ND	ND	99.3	99.1	99.1
B	ND	ND	ND	100	100	100
C	ND	ND	ND	99.3	100	100
D	ND	ND	ND	84.8	87.7	84.1
E	ND	ND	ND	98.6	— ^b	— ^b
F	ND	ND	ND	98.6	98.8	98.8
G	ND	ND	ND	98.5	98.0	97.5
H	99.9	99.7	99.7	75.7	82.1	77.6
I	99.5	99.5	99.5		Absent	
J	100	100	100		Absent	
K	99.8	98.9	98.9		Absent	
L	99.9	100	100		Absent	
M	99.8	99.6	99.6		Absent	
N	ND	ND	ND	98.3	82.9 ^c	81.6 ^c

^a The main differences are highlighted by boldface type. ND, sequence not determined.

^b A nucleotide deletion caused a frameshift so that *orfE* in JJ692 was prematurely terminated and consisted of only 30 amino acids instead of the 211 amino acids in PAK and CF5.

^c A nucleotide insertion caused a frameshift at the 3' end of *orfN* of *P. aeruginosa* strain JJ692, leading to a drop in homology after amino acid 867 of the PAK *orfN*.

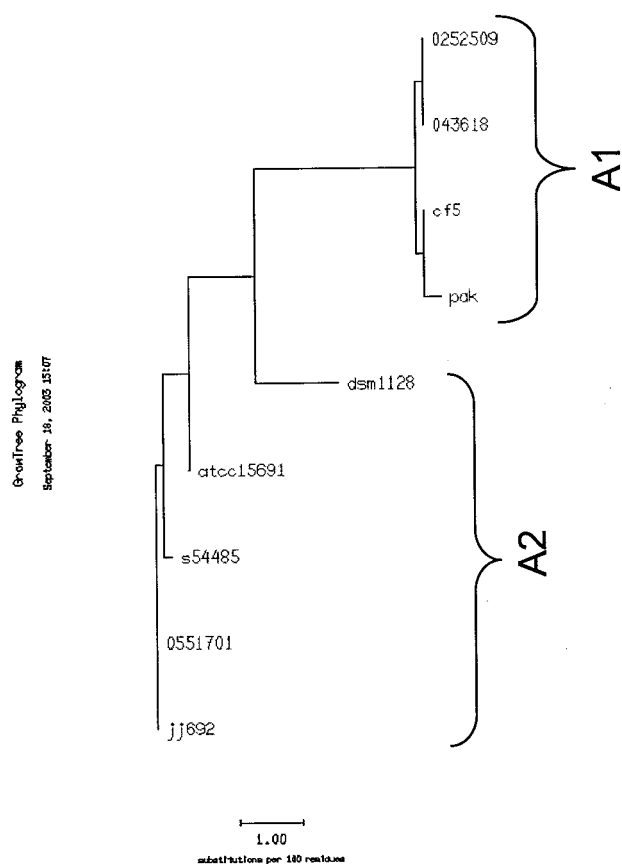


FIG. 4. GROWTREE phylogenetic analysis of a-type flagellins. The dendrogram was created with the GROWTREE program of the GCG Wisconsin package. Amino acid sequences of a-type flagellins from 7 *P. aeruginosa* strains including PAK and two a-type flagellin sequences already existing in the databank (accession numbers L81146 and L81147) were used to generate the dendrogram. The two subtypes of a-type flagellins, A1 and A2, are evident.

tional 17 a-type strains from diverse sources. A PILEUP alignment of these partial flagellin amino acid sequences from 24 *P. aeruginosa* strains, two a-type flagellin sequences already present in the databank, and a flagellin sequence from a strain with H-antigen serotype a (obtained from a U.S. investigator) is shown in Fig. 5. Again, two main subtypes of flagellins were evident. Specifically, there were different amino acid substitutions at several places, as highlighted in Fig. 5, between the two types, but the main differences were two small deletions of codons for 3 and 4 amino acids, respectively, at positions 254 and 279 relative to the PAK flagellin sequence. A hydrophobicity plot of the PAK (A1) and JJ692 (A2) flagellin amino acid sequences showed no apparent differences (data not shown), suggesting that the two small deletions and several amino acid substitutions in the JJ692 flagellin did not significantly affect its secondary structure.

We then attempted to correlate the inheritance of the short or long islands with A1 or A2 flagellins. However, no strict correlation could be established regarding the coinheritance of A1 or A2 flagellins with the short or long GI. Table 3 shows the distribution of short and long GIs and A1 and A2 flagellins in *P. aeruginosa* strains of different origins. The *P. aeruginosa*

strains carrying the short GI had either A1 or A2 flagellins without any preferences for either one. But, all the strains carrying the long GI had A1 flagellins, with the exception of one isolate (CF29). The physiological significance of this finding is not clear at present. In addition, there was no preference for A1 or A2 flagellins in CF, UTI, or blood isolates.

DISCUSSION

A genomic island carrying 14 ORFs present in *P. aeruginosa* strain PAK was previously implicated in the glycosylation of a-type flagellins (3). This island is inserted into a cluster of flagellar genes, and this region of the *P. aeruginosa* chromosome is highly polymorphic. Here we report further the presence of additional polymorphism within the GIs of *P. aeruginosa* strains expressing a-type flagellins. Many *P. aeruginosa* strains possessed a truncated version of the island lacking the downstream *orfI*, *-J*, *-K*, *-L*, and *-M*. In addition, *orfD*, *-E*, and *-H*, although present in short GIs, are also polymorphic compared to their homologues in long GIs. There was no strict correlation between site of infection and island type; however, the fraction of strains carrying the short island was higher in UTI isolates than was among CF or blood isolates.

Spencer et al. (25) compared three *P. aeruginosa* isolates, two from clonal infections of CF patients and one from an aquatic environment, with the genomic sequence of the reference strain PAO1 by shotgun sequencing. They concluded that these strains possessed at least 10% more genetic material than PAO1. They also inferred that their flagellar regions were different from that of PAO1, but no indication of the polymorphisms of the glycosylation region of the two strains that were different from PAO1 was provided. Polymorphism has also been reported in other prokaryotic glycosylation systems, e.g., the pilin glycosylation locus of *Neisseria meningitidis* expressing class II pili (15). The biological significance of this polymorphism and genetic diversity is not clear at present, but one may speculate that they may at least affect the immunological recognition of specific epitopes of these proteins. It is, however, possible that the *P. aeruginosa* strains carrying long or short GIs display different glycan structures on their flagella, which might provide them with survival advantages in different environments.

P. aeruginosa a-type flagellins have been classified into several subtypes based either on their antigenic components (2, 22) or on nucleotide sequence variations (4, 23, 30). The classification based on nucleotide variations in flagellin has minimal functional implications. An alternative classification scheme based on the amino acid sequences of flagellins may offer some serological specificities not previously recognized as such. Based on the amino acid sequences of flagellins from 24 a-type *P. aeruginosa* strains, the largest number sequenced to date, we were able to identify two broad subtypes, A1 and A2. Such structural differences of flagellins have been displayed in several other bacteria, including *Escherichia coli* (28), *C. jejuni* (10), and *Helicobacter pylori* (13). The mechanism of flagellin polymorphism in *C. jejuni* was suggested to be both intergenomic and intragenomic recombination (10). Similar recombination mechanisms may be responsible for flagellin polymorphism in *P. aeruginosa*.

Further, we explored whether there was any correlation be-



FIG. 5. PILEUP analysis of the partial central domains of 27 *P. aeruginosa* strains. Amino acid (aa) sequences from amino acids 228 to 297 of a-type flagellins of 24 *P. aeruginosa* strains analyzed in this study, one a-type strain obtained from R. A. Ansorg, and two complete a-type flagellin sequences already present in the database were aligned by using the PILEUP program of the GCG Wisconsin package. Amino acid substitutions at positions 237, 240, 241, 249, 250, 253, 266, 273, 283, 285, 287, 288, and 289 and two small deletions of 3 and 4 amino acids at positions 254 and 279, respectively, are highlighted in different shades of gray. The A2 flagellin of strain DSM 1128 was peculiar in having additional amino acid substitutions (shown in bold italics).

tween long or short GI and the coinheritance of A1 or A2 flagellins. As shown in Table 3, *P. aeruginosa* strains carrying the short GI had either A1 or A2 flagellins while the strains carrying the long island preferred the A1 flagellin. Although there were significant differences between the amino acid sequences of A1 and A2 flagellins, it seemed that these variations had no significant effect on their secondary structure. These data are consistent with the results of previous analyses of a- and b-type flagellin structures which showed very similar structural characteristics in spite of 37 to 38% divergence at the amino acid level (24), suggesting a very low impact of sequence diversity on secondary and tertiary structure of *P. aeruginosa* flagellins. With this minimal structural diversity, the host immunological response to A1 and A2 flagellins may not be too different, and therefore, the antigenic diversity among these flagellins may not be too large. Thus, given that there are two main subtypes of a-type flagellins in *P. aeruginosa* in addition to the b-type flagellin, it may prove feasible to design a more

encompassing vaccine candidate against *P. aeruginosa*, since *P. aeruginosa* flagellar preparations have been shown to be protective in animals (8, 12). One such trial began in 1997, and preliminary results showed protection against *P. aeruginosa* colonization in CF patients (7). One considerable interest is whether these differences in structure may lead to differences in the response of the innate immune system to flagellin via Toll-like receptor 5 (11), and it will be important to ascertain this for specific flagellin types as well as different glycosylation types. Further efforts are geared towards answering such vital questions.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI 47852 (to R.R.). M.C.W. was supported by a Cystic Fibrosis Foundation Postdoctoral Research Fellowship.

REFERENCES

- Allison, J. S., M. Dawson, D. Drake, and T. C. Montie. 1985. Electrophoretic separation and molecular weight characterization of *Pseudomonas aeruginosa* H-antigen flagellins. *Infect. Immun.* **49**:770-774.
- Ansorg, R. 1978. Flagella specific H antigenic schema of *Pseudomonas aeruginosa*. *Zentbl. Bakteriol. Parasitenkd. Infekkrankh. Hyb. Abt. 1.* **224**: 228-238.
- Arora, S., M. Bangera, S. Lory, and R. Ramphal. 2001. A genomic island in *Pseudomonas aeruginosa* carries the determinants of flagellin glycosylation. *Proc. Natl. Acad. Sci. USA* **98**:9342-9347.
- Bellingham, N. F., J. A. Morgan, J. R. Saunders, and C. Winstanley. 2001. Flagellin gene sequence variation in the genus *Pseudomonas*. *Syst. Appl. Microbiol.* **24**:157-165.
- Brimer, C. D., and T. C. Montie. 1998. Cloning and comparison of *fliC* genes and identification of glycosylation in the flagellin of *Pseudomonas aeruginosa* a-type strains. *J. Bacteriol.* **180**:3209-3217.
- Doig, P., N. Kinsella, P. Guerry, and T. J. Trust. 1996. Characterization of a post-translational modification of *Campylobacter* flagellin: identification of a sero-specific glycosylation moiety. *Mol. Microbiol.* **19**:379-387.

TABLE 3. Distribution of short and long GIs and A1 and A2 flagellins in 24 *P. aeruginosa* strains of different origins

Origin (n) of <i>P. aeruginosa</i> isolates	No. of isolates with GI and flagellin subtype					
	Short			Long		
	Total	A1	A2	Total	A1	A2
UTI (10)	6	2	4	4	4	0
CF (9)	5	3	2	4	3	1
Blood (4)	2	1	1	2	2	0
Lab strain PAK (1)	0	0	0	1	1	0

7. Doring, G. 2003. Relevant issues in bacterial vaccine development for patients with cystic fibrosis. *Pediatr. Pulmonol. Suppl.* **25**:128–129.
8. Drake, D., and T. C. Montie. 1987. Protection against *Pseudomonas aeruginosa* infection by passive transfer of anti-flagellar serum. *Can. J. Microbiol.* **33**:755–763.
9. Feldman, M., R. Bryan, S. Rajan, L. Scheffler, S. Brunner, H. Tang, and A. Prince. 1998. Role of flagella in pathogenesis of *Pseudomonas aeruginosa* pulmonary infection. *Infect. Immun.* **66**:43–51.
10. Harrington, C. S., F. M. Thomson-Carter, and P. E. Carter. 1997. Evidence for recombination in the flagellin locus of *Campylobacter jejuni*: implications for the flagellin gene typing scheme. *J. Clin. Microbiol.* **35**:2386–2392.
11. Hayashi, F., K. D. Smith, A. Ozinsky, T. R. Hawn, E. C. Yi, D. R. Goodlett, J. K. Eng, S. Akira, D. M. Underhill, and A. Aderem. 2001. The innate immune response to bacterial flagellin is mediated by toll-like receptor 5. *Nature* **410**:1099–1103.
12. Holder, I. A., R. Wheeler, and T. C. Montie. 1982. Flagellar preparations from *Pseudomonas aeruginosa*: animal protection studies. *Infect. Immun.* **35**:276–280.
13. Ji, W., J. Hu, J. Qiu, D. Peng, B. Shi, S. Zhou, K. Wu, and D. Fan. 2001. Polymorphism of flagellin A gene in *Helicobacter pylori*. *World J. Gastroenterol.* **7**:783–787.
14. Josenhans, C., R. L. Ferrero, A. Labigne, and S. Suerbaum. 1999. Cloning and allelic exchange mutagenesis of two flagellin genes of *Helicobacter felis*. *Mol. Microbiol.* **33**:350–362.
15. Kahler, C. M., L. E. Martin, Y. Tzeng, Y. K. Miller, K. Sharkey, D. S. Stephens, and J. K. Davies. 2001. Polymorphisms in pilin glycosylation locus of *Neisseria meningitidis* expressing class II pili. *Infect. Immun.* **69**:3597–3604.
16. Lanyi, B. 1970. Serological properties of *Pseudomonas aeruginosa*. II. Type-specific thermostable (flagellar) antigens. *Acta Microbiol. Acad. Sci. Hung.* **17**:35–48.
17. Liaudet, L., C. Szabo, O. V. Eygenov, K. G. Murthy, P. Pacher, L. Virag, J. G. Mabley, A. Marton, F. G. Soriano, M. Y. Kirov, L. J. Bjertnaes, and A. L. Salzman. 2003. Flagellin from gram-negative bacteria is a potent mediator of acute pulmonary inflammation in sepsis. *Shock* **19**:131–137.
18. Montie, T. C., D. Doyle-Huntzinger, R. C. Craven, and I. A. Holder. 1982. Loss of virulence associated with absence of flagellum in an isogenic mutant of *Pseudomonas aeruginosa* in the burned-mouse model. *Infect. Immun.* **38**:1296–1298.
19. Montie, T. C., D. Drake, H. Sellin, O. Slater, and S. Edmonds. 1987. Motility, virulence, and protection with a flagella vaccine against *Pseudomonas aeruginosa* infection. *Antibiot. Chemother.* **39**:233–248.
20. Morgan, J. A. W., N. F. Bellingham, C. Winstanley, M. A. Ousley, C. A. Hart, and J. R. Saunders. 1999. Comparison of flagellin genes from clinical and environmental *Pseudomonas aeruginosa* isolates. *Appl. Environ. Microbiol.* **65**:1175–1179.
21. Namba, K., I. Yamashita, and F. Vonderviszt. 1989. Structure of the core and central channel of bacterial flagella. *Nature* **342**:648–654.
22. Pitt, T. L. 1981. A comparison of flagellar typing and phage typing as means of subdividing the O groups of *Pseudomonas aeruginosa*. *Med. Microbiol.* **14**:261–270.
23. Spangenberg, C., T. Heuer, C. Burger, and B. Tummeler. 1996. Genetic diversity of flagellins of *Pseudomonas aeruginosa*. *FEBS Lett.* **396**:213–217.
24. Spangenberg, C., T. C. Montie, and B. Tummeler. 1998. Structural and functional implications of sequence diversity of *Pseudomonas aeruginosa* genes, *oriC*, *ampC* and *fliC*. *Electrophoresis* **19**:545–550.
25. Spencer, D. H., A. Kas, E. E. Smith, C. K. Raymond, E. H. Sims, M. Hastings, J. L. Burns, R. Kaul, and M. V. Olson. 2003. Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.* **185**:1316–1325.
26. Thibault, P., S. M. Logan, J. F. Kelly, J. Brisson, C. P. Ewing, T. J. Trust, and P. Guerry. 2001. Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **276**:34862–34870.
27. Totten, P. A., and S. Lory. 1990. Characterization of the a-type flagellin gene from *Pseudomonas aeruginosa* PAK. *J. Bacteriol.* **172**:7188–7199.
28. Wang, L., D. Rothmund, H. Curd, and P. R. Reeves. 2003. Species-wide variation in the *Escherichia coli* flagellin (H-antigen) gene. *J. Bacteriol.* **185**:2936–2943.
29. Wilson, D. R., and T. J. Beveridge. 1993. Bacterial flagellar filaments and their component flagellins. *Can. J. Microbiol.* **39**:451–472.
30. Winstanley, C., M. A. Coulson, B. Wepner, J. A. W. Morgan, and C. A. Hart. 1996. Flagellin gene and protein variation amongst clinical isolates of *Pseudomonas aeruginosa*. *Microbiology* **142**:2145–2151.
31. Wolfgang, M. C., B. R. Kulasekara, X. Liang, D. Boyd, K. Wu, Q. Yang, C. G. Miyada, and S. Lory. 2003. Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **100**:8484–8489.