JOURNAL OF BACTERIOLOGY, Feb. 1991, p. 1357–1359
0021-9193/91/031357-03$02.00/0

Vol. 173, No. 3

N-Terminal Amino Acid Sequences and Amino Acid Compositions of the *Spirochaeta aurantia* Flagellar Filament Polypeptides

JUAN PARALES, JR., AND E. P. GREENBERG*

Department of Microbiology, University of Iowa, Iowa City, Iowa 52242

Received 17 September 1990/Accepted 24 November 1990

The amino-terminal sequences and amino acid compositions of the three major and two minor polypeptides constituting the filaments of *Spirochaeta aurantia* periplasmic flagella were determined. The amino-terminal sequence of the major 37.5-kDa outer layer polypeptide is identical to the sequence downstream of the proposed signal peptide of the protein encoded by the *S. aurantia flaA* gene. However, the amino acid composition of the 37.5-kDa polypeptide is not in agreement with that inferred from the sequence of *flaA*. The 34- and 31.5-kDa major filament core polypeptides and the 33- and 32-kDa minor core polypeptides show a striking similarity to each other, and the amino-terminal sequences of these core polypeptides show extensive identity with homologous proteins from members of other genera of spirochetes. An additional 36-kDa minor polypeptide that occurs occasionally in preparations of *S. aurantia* periplasmic flagella appears to be mixed with the 37.5-kDa outer layer polypeptide or a degradation product of this polypeptide.

Spirochetes are thin, helical bacteria that are motile by means of flagella that are enclosed entirely within the periplasmic space. One end of each periplasmic flagellum is inserted near one pole of the cylinder, and the other end is free (15). *Spirochaeta aurantia*, a facultatively anaerobic spirochete isolated from freshwater sediments, has two periplasmic flagella, each inserted at opposite ends of the cell. The filaments of these periplasmic flagella wrap around the helical protoplasmic cylinder, which consists of the cytoplasmic region, the cytoplasmic membrane, and the peptidoglycan (5). The *S. aurantia* flagellar filaments are ultrastructurally and biochemically complex, consisting of an outer layer thought to be composed of repeating subunits of a polypeptide with an apparent molecular weight of 37,500 (the 37.5K polypeptide) and a core that contains one or all of five antigenically related polypeptides, i.e., the abundant or major 34K and 31.5K polypeptides and the minor 36K, 33K, and 32K polypeptides (3). It is not known whether individual filaments contain multiple core polypeptide species or whether different filaments contain different core polypeptides. The structural and biochemical complexity of the flagellar filament is a characteristic shared by a variety of spirochetes, including several species of *Treponema* and *Leptospira* (6, 8, 23, 24). Apparently, however, the periplasmic flagella of species in the genus *Borrelia* do not possess an outer layer (1).

Genes encoding flagellar filament outer layer antigens (*flaA* genes) have been cloned from *S. aurantia* and *Treponema pallidum* and have been sequenced (4, 18, 19). The first halves of the two *flaA* genes and their products show significant sequence similarity (13a). Unlike flagellins from other bacteria, the proteins encoded by these genes appear to possess N-terminal signal peptides (4, 19, 24). In the case of *T. pallidum*, amino-terminal sequencing of the *flaA* product isolated from purified periplasmic flagella has confirmed that the signal peptide is cleaved from the polypeptide before flagellar assembly in the periplasm (2, 24).

The amino termini of each of the three core polypeptides from *T. pallidum* periplasmic flagella, the two core polypeptides from *Treponema phagedenis*, and a filament polypeptide from *Borrelia burgdorferi* (2, 9, 13, 24) have been sequenced; all show a considerable degree of conservation, i.e., 50 to 95% identity in the amino-terminal 20 amino acid residues (13). Although the amino termini of the core polypeptides within a treponemal species are similar to each other, they are not identical, thus indicating that each polypeptide is the product of a distinct gene (24). In fact, recent publications describing the cloning and sequencing of three genes, *flaB1*, *flaB2*, and *flaB3*, which encode the three different flagellar filament core polypeptides from *T. pallidum*, FlaB1, FlaB2, and FlaB3, prove that in this organism at least, each of the core polypeptides is a distinct gene product (7, 25).

Although *S. aurantia* has served as a model for studies of the mechanism of spirochete motility and chemotaxis (3, 4, 11, 12, 14), the N-terminal amino acid sequences of the flagellar filament polypeptides of this spirochete have not been reported. We have performed this analysis, and we have also determined the amino acid compositions of the *S. aurantia* flagellar filament polypeptides. These analyses provide information about the nature of the FlaA polypeptide on assembled flagellar filaments and about the relationship of the *S. aurantia* core polypeptides to each other and to those of other spirochetes.

The bacterial strain used, *S. aurantia* M1 (5), was grown as described elsewhere (3). Intact periplasmic flagella were isolated by the method of Brahamsha and Greenberg (3). The different species of flagellar filament polypeptides were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis as described elsewhere (3, 21), except that recrystallized sodium dodecyl sulfate was used (17) and the running buffer contained 0.1 mM sodium thioglycolate. After gel electrophoresis, gels were treated and the polypeptides were electroblotted on polyvinylidene difluoride membranes (Immobilon Transfer Membranes; Millipore, Bedford, Mass.) as described by Matsudaira (22). The electroblotted polypeptides were stained with Coomassie blue and cut out with a clean razor. The N-terminal amino acid sequences of the polypeptides were determined as described elsewhere (22), and the amino acid compositions were determined by

* Corresponding author.

```
MKRFFAILGAALFVGNSGAFAEQATLIDFSKLVGEGNTGLHAPTTIDYS
                EQATLIDFSKLVGEGNTGLHAPTTIDYS
```

FIG. 1. Comparison and alignment of the deduced N-terminal amino acid sequence of the *S. aurantia flaA* product (top) and the N-terminal amino acid sequence of the FlaA polypeptide from purified *S. aurantia* flagella (bottom).

Ka-Leung Ngai at the Northwestern University Biotechnology Facility as described previously (20).

Analysis of the 37.5K flagellar filament outer layer polypeptide. A comparison of the N-terminal amino acid sequence of the *S. aurantia* FlaA polypeptide and the sequence deduced from the DNA sequence of *flaA* (4) is shown in Fig. 1. The region of identity between the predicted sequence and the actual sequence begins after residue 21 of the predicted sequence, and the predicted sequence is identical to the N-terminal amino acid sequence determined for the FlaA polypeptide over the next 28 amino acid residues. This supports the hypothesis that *flaA* encodes a polypeptide with a 21-amino-acid residue signal sequence that is cleaved during secretion of this protein through the cytoplasmic membrane into the periplasm, where it assembles onto flagellar filaments (4). Furthermore, the identity between the predicted and actual FlaA sequences is as expected if *flaA* does indeed code for the 37.5K flagellar filament polypeptide. The amino acid composition of the FlaA polypeptide from purified flagella differs, however, from that predicted from the *flaA* DNA sequence (Table 1). This taken together with the fact that the *S. aurantia flaA* sequence and the sequence of the encoded polypeptide show considerable identities with those of the *T. pallidum flaA* gene (18, 19) over the first 508 bp of the *S. aurantia flaA* open reading frame, corresponding to the first 170 amino acid residues of the encoded polypeptide only (13a), suggests that after this point of divergence from the *T. pallidum flaA* gene, the *S. aurantia flaA* sequence is incorrect. This was confirmed by cloning a large *flaA*-containing fragment of *S. aurantia* DNA and sequencing the *flaA* gene on this fragment. The *S. aurantia flaA* sequence seems correct through the *Sau*3AI site starting at bp 647 in the published sequence (4). This *Sau*3AI site appears to be a junction of two different fragments of DNA (26). It is not clear why this was not revealed by previously reported Southern analyses (2a, 4). The corrected *flaA* sequence and the *flaA* open reading frame from the *T. pallidum* sequence show significant identity, and the polypeptides encoded by these homologous genes show a 41% identity (26).

TABLE 1. Amino acid compositions of the 37.5K flagellar polypeptide and the predicted *flaA* product

Amino acid residue	Amino acid composition[a]	
	37.5K polypeptide	*flaA* product[b]
D	1.12	0.81
E	0.97	1.23
S	0.71	0.58
G	0.88	1.54
H	0.0	0.38
R	0.50	0.42
T	0.43	0.62
A	1.00	1.00
P	0.25	0.85
Y	0.32	0.15
V	0.49	0.96
M	0.0	0.15
C	0.0	0.0
I	0.46	0.50
L	0.70	0.85
F	0.27	0.35
K	0.41	0.35
W	ND[c]	0.04

[a] Expressed as micromoles per micromole of alanine.
[b] Values are those for the predicted *flaA* product less the residues in the 21-amino-acid residue signal peptide (Fig. 1).
[c] ND, Not determined.

```
Major 34K   MIINHNMSAI NAQRVQGBVT GVTKNMV
Minor 33K   MIINHNMSAI NAQXVQGBVT
Minor 32K   MIINHNMSAI NANRVLGBT
Major 31.5K MIINHNMSAI NANRVLGBTN ADITKDL
```

FIG. 2. N-terminal amino acid sequences of the 34K major, 33K and 32K minor, and 31.5K major core polypeptides from purified *S. aurantia* flagella.

Analysis of the flagellar filament core polypeptides. It has been demonstrated previously by Western immunoblotting that there are five antigenically related core polypeptides that correspond to the 34K and 31.5K major bands and the 36K, 33K, and 32K minor bands on sodium dodecyl sulfate-polyacrylamide gels. Of these polypeptide bands we were able to determine the N-terminal amino acid sequences and the amino acid compositions of all but the 36K band, which, as determined by N-terminal sequencing and Western immunoblotting, was always contaminated with the 37.5K outer layer polypeptide or a degradation product of this protein (data not shown). The N-terminal amino acid sequences of all four core polypeptides analyzed were identical in the first 12 positions (Fig. 2). After position 12 the 34K major and 33K minor polypeptides showed a divergence from the 31.5K major and 32K minor polypeptides, but within those two pairings the N-terminal amino acid sequences could not be distinguished (Fig. 2). The 34K major and 33K minor polypeptides were, however, distinguishable with respect to amino acid composition, as were the 31.5K major and 32K minor polypeptides (Table 2). With respect to the 34K and 33K polypeptides there are notable differences in the numbers of aspartate plus asparagine residues, histidine residues, and threonine residues. With respect to the 31.5K and 32K polypeptides the most notable differences are in the aspartate plus asparagine residues, the proline residues, and the phenylalanine residues. It thus appears that each of the four *S. aurantia* core polypeptides is the product of a different gene; however, this remains to be demonstrated, as it has been for the three *T. pallidum* flagellar core polypeptides (7, 25).

The first 20 amino acid residues of the *S. aurantia* core polypeptides not only showed extensive similarity to each other (Fig. 2) but also were similar to the first 20 residues in the homologous polypeptides from *T. pallidum*, *T. phagedenis*, and *B. burgdorferi*; of the eight homologs compared, the identities over this 20-amino-acid residue region ranged from 50 to 95%. Over the first 10 residues, all of the *Treponema* sequences are identical, and they differ from the *S. aurantia* sequences (Fig. 2) only by the substitution of a methionine for isoleucine at position 10 (24). The N-terminal 10 residues of the *B. burgdorferi* flagellin are identical to those of the *S. aurantia* core polypeptides, with the exception of a threonine in place of methionine at position 7 (9, 13).

For all of the *S. aurantia* core polypeptides analyzed (Fig. 2), as well as their homologs from *B. burgdorferi* (9, 13) and

TABLE 2. Amino acid compositions of *S. aurantia* flagellar filament core polypeptides

Amino acid residue	Amino acid composition[a]			
	34K major polypeptide	33K minor polypeptide	32K minor polypeptide	31.5K major polypeptide
D	1.61	2.24	2.03	1.33
E	1.17	1.03	0.97	1.00
S	0.82	0.71	0.61	0.59
G	1.08	0.49	1.52	0.64
H	0.18	0.94	0.17	0.12
R	0.58	0.48	0.45	0.39
T	0.17	0.41	0.36	0.56
A	1.00	1.00	1.00	1.00
P	0.17	0.19	0.33	0.08
Y	0.18	0.17	0.15	0.13
V	0.62	0.56	0.47	0.49
M	0.15	0.16	0.14	0.14
C	0	0	0	0
I	0.36	0.35	0.33	0.42
L	0.70	0.52	0.48	0.55
F	0.18	0.06	0	0.10
K	0.21	0.21	0.26	0.21
W	ND[b]	ND	ND	ND

[a] Expressed as micromoles per micromole of alanine.
[b] ND, Not determined.

the treponemes (2, 24), the N-terminal amino acid is methionine. At least for *T. pallidum*, from which the genes encoding each of the three core polypeptides have been cloned and sequenced, it is known that the N terminus is not modified; there is no signal peptide (7, 24, 25). Presumably these polypeptides are secreted in a fashion analogous to secretion of *Escherichia coli* and *Salmonella typhimurium* flagellins, which apparently migrate through the core of a growing filament to the growing tip, where they polymerize (10, 16).

Nucleotide sequence accession number. The corrected *flaA* sequence has been deposited in GenBank (accession number M24459).

This work was supported by a grant from the National Science Foundation (DCB 8913528).

REFERENCES

1. **Barbour, A. G., S. F. Hayes, R. A. Heiland, M. E. Schrumpf, and S. L. Tessier.** 1986. A *Borrelia*-specific monoclonal antibody binds to a flagellar epitope. Infect. Immun. **52**:549–554.
2. **Blanco, D. R., C. I. Champion, J. N. Miller, and M. A. Lovett.** 1988. Antigenic and structural characterization of *Treponema pallidum* (Nichols strain) endoflagella. Infect. Immun. **56**:168–175.

2a. **Brahamsha, B.** 1988. Ph.D. thesis. Cornell University, Ithaca, N.Y.

3. **Brahamsha, B., and E. P. Greenberg.** 1988. A biochemical and cytological analysis of the complex periplasmic flagella from *Spirochaeta aurantia*. J. Bacteriol. **170**:4023–4032.
4. **Brahamsha, B., and E. P. Greenberg.** 1989. Cloning and sequence analysis of *flaA*, a gene encoding a *Spirochaeta aurantia* flagellar filament surface antigen. J. Bacteriol. **171**:1692–1697.
5. **Breznak, J. A., and E. Canale-Parola.** 1975. Morphology and physiology of *Spirochaeta aurantia* strains isolated from aquatic habitats. Arch. Microbiol. **105**:1–12.
6. **Canale-Parola, E.** 1978. Motility and chemotaxis of spirochetes. Annu. Rev. Microbiol. **32**:69–99.
7. **Champion, C. I., J. N. Miller, M. A. Lovett, and D. R. Blanco.** 1990. Cloning, sequencing, and expression of two class B endoflagellar genes of *Treponema pallidum* subsp. *pallidum* encoding the 34.5- and 31.0-kilodalton proteins. Infect. Immun. **58**:1697–1704.
8. **Cockayne, A., M. J. Bailey, and C. W. Penn.** 1987. Analysis of sheath and core structures of the axial filament of *Treponema pallidum*. J. Gen. Microbiol. **113**:1397–1407.
9. **Coleman, J. L., and J. L. Benach.** 1989. Identification and characterization of an endoflagellar antigen of *Borrelia burgdorferi*. **84**:322–330.
10. **Emerson, S. V., K. Tokuyasu, and M. I. Simon.** 1970. Bacterial flagella: polarity of elongation. Science **169**:190–192.
11. **Fosnaugh, K., and E. P. Greenberg.** 1988. Motility and chemotaxis of *Spirochaeta aurantia*: computer-assisted motion analysis. J. Bacteriol. **170**:1768–1774.
12. **Fosnaugh, K., and E. P. Greenberg.** 1989. Chemotaxis mutants of *Spirochaeta aurantia*. J. Bacteriol. **171**:606–611.
13. **Gasmann, G. S., R. Deutzmann, A. Vogt, and V. B. Göbel.** 1989. N-terminal amino acid sequence of the *Borrelia burgdorferi* flagellin. FEMS Microbiol. Lett. **60**:101–106.

13a. **Greenberg, E. P.** Unpublished data.

14. **Greenberg, E. P., B. Brahamsha, and K. Fosnaugh.** 1985. The motile behaviour of *Spirochaeta aurantia*: a twist to chemosensory transduction in bacteria, p. 107–118. *In* M. Eisenbach and M. Balaban (ed.), Sensing and response in microorganisms. Elsevier Science Publishing Inc., New York.
15. **Holt, S. C.** 1978. Anatomy and chemistry of spirochetes. Microbiol. Rev. **42**:114–160.
16. **Homma, M., H. Fujita, S. Yamaguchi, and T. Iino.** 1987. Regions of *Salmonella typhimurium* flagellin essential for its polymerization and excretion. J. Bacteriol. **169**:291–296.
17. **Hunkapillar, M. W., E. Lujan, F. Ostrander, and L. E. Hood.** 1983. Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. Methods Enzymol. **91**:227–236.
18. **Issacs, R. D., J. H. Hanke, L.-M. Guzman-Verduzco, G. Newport, N. Agabian, M. V. Norgard, S. A. Lukehart, and J. D. Radolf.** 1989. Molecular cloning and DNA sequence analysis of the 37-kilodalton endoflagellar sheath protein gene of *Treponema pallidum*. Infect. Immun. **57**:3403–3411.
19. **Issacs, R. D., and J. D. Radolf.** 1990. Expression in *Escherichia coli* of the 37-kilodalton endoflagellar sheath protein of *Treponema pallidum* by use of the polymerase chain reaction and a T7 expression system. Infect. Immun. **58**:2025–2034.
20. **Kaplan, H. B., and E. P. Greenberg.** 1987. Overproduction and purification of the *luxR* gene product: the transcriptional activator of the *Vibrio fischeri* luminescence system. Proc. Natl. Acad. Sci. USA **84**:6639–6643.
21. **Kropinski, A. M., T. R. Parr, Jr., B. L. Angus, R. E. W. Hancock, W. G. Ghiorse, and E. P. Greenberg.** 1987. Isolation of the outer membrane and characterization of the major outer membrane protein from *Spirochaeta aurantia*. J. Bacteriol. **169**:171–179.
22. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262**:10035–10038.
23. **Nauman, R. K., S. C. Holt, and C. D. Cox.** 1969. Purification, ultrastructure, and composition of axial filaments from *Leptospira*. J. Bacteriol. **98**:264–280.
24. **Norris, S. J., N. W. Charon, R. G. Cook, M. D. Fuentes, and R. J. Limberger.** 1988. Antigenic relatedness and N-terminal sequence homology define two classes of periplasmic flagellar proteins of *Treponema pallidum* subsp. *pallidum* and *Treponema phagedenis*. J. Bacteriol. **170**:4072–4082.
25. **Pallesen, L., and P. Hindersson.** 1989. Cloning and sequencing of a *Treponema pallidum* gene encoding a 31.3-kilodalton endoflagellar subunit (FlaB2). Infect. Immun. **57**:2166–2172.
26. **Parales, J., and E. P. Greenberg.** Unpublished data.