THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 263, No. 26, Issue of September 15, pp. 13246-13251, 1988
Printed in U.S.A.

Halobacterial Flagellins Are Encoded by a Multigene Family

CHARACTERIZATION OF FIVE FLAGELLIN GENES*

(Received for publication, December 24, 1987)

Lydia Gerl and Manfred Sumper‡

From the Institut für Biochemie, Genetik und Mikrobiologie, Universität Regensburg, Universitätsstrasse 31, 8400 Regensburg, Federal Republic of Germany

Purified flagellar filaments of *Halobacterium halobium* contain three different protein species based on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. These proteins were designated as flagellins Fla I, Fla II, and Fla III and were characterized as sulfated glycoproteins with *N*-glycosidically linked oligosaccharides of the type GlcA-(1→4)-GlcA-(1→4)-GlcA-(1→4)-Glc. All halobacterial flagellin polypeptides are immunologically cross-reactive.

A gene fragment of one flagellin was isolated in an expression vector using antibody probes. Using this gene fragment as probe, we identified, subcloned, and determined the nucleotide sequences of five different but highly homologous flagellin genes. Two flagellin (flg) genes are arranged tandemly at one locus (flg A1 and -2), and the other three in a tandem arrangement at a different locus (flg B1, -2, and -3), Two flg mRNAs were detected, one from the A genes and the other from the B genes. Based on immunological analysis, the products of the flg A1 and A2 are Fla II and Fla I, respectively.

The first true glycoprotein discovered in a procaryotic organism was the cell surface glycoprotein of halobacteria, a main constituent of the halobacterial S layer (1). The saccharide structures of this protein have been analyzed in detail (2), the gene encoding this procaryotic glycoprotein has been cloned, and its nucleotide sequence has been determined (3). Halobacteria synthesize at least three more glycoproteins. A set of heterogenous sulfated glycoproteins is halobacterial flagellins (4), originally purified by Alam and Oesterhelt (5) and designated as Fla I, Fla II, and Fla III. On SDS[1]-polyacrylamide gel electrophoresis these flagellins display a ladder-like pattern of different bands with three centers of intensity.

The halobacterial flagellum exhibits unusual biological properties (5). Halobacteria swim forward by clockwise and backward by counterclockwise rotation of their right-handed flagellar bundles. These flagellar bundles do not fly apart when the sense of rotation changes. The relationship of structure to function in this unique flagellum is of obvious interest. A starting point for such an investigation is the establishment of the primary structure of all the halobacterial flagellins involved in formation of the flagellum. In this paper, we describe the isolation and characterization of five related genes encoding the halobacterial flagellins.

* This work was supported by the Deutsche Forschungsgemeinschaft (SFB 43). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J03942.

‡ To whom correspondence should be addressed.

[1] The abbreviations used are: SDS, sodium dodecyl sulfate; Mops, 4-morpholinepropanesulfonic acid; flg, flagellin; bp, base pair(s); kb, kilobase pain(s); ORF, open reading frame.

MATERIALS AND METHODS

Strains, Plasmids, DNA and RNA Preparations—*Halobacterium halobium* strain R_1M_1 was grown in complex medium (6). DNA and RNA preparations, construction of libraries of halobacterial genomic DNA in the expression vector pIN-III A, and EMBL4 were previously described (3). Oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer.

DNA Sequencing—DNA sequencing was performed by the dideoxy chain termination method (7) using [α-*thio*-^{35}S]dATP and pUC8 as a vector (8).

Purification of Flagellin Proteins—Isolation of flagellin proteins was performed according to Ref. 9.

Labeling of DNA Fragments and Oligonucleotides—Double-stranded DNA fragments were labeled with [α-*thio*-^{35}S]dATP either by nick translation (10) or by using random 6-mer oligonucleotides and the Klenow polymerase to a specific activity of 10^8 cpm/μg DNA (11). Oligonucleotides were labeled with [γ-^{32}P]ATP using polynucleotide kinase to give 6×10^6 cpm/pmol.

Southern Blot Hybridizations—After electrophoresis and denaturation, the DNA fragments were transferred to nitrocellulose membranes (12) and hybridized (10). Using ^{35}S-labeled DNA as probes (10^6 cpm/ml), the hybridization solution contained 50% formamide and 50 mM dithiothreitol. The blots were incubated and washed under standard conditions (10). Hybridization with labeled ^{32}P-oligonucleotides (10^6 cpm/ml) was in the absence of formamide for 15 h at 42 °C. The blots were washed twice in 2 × SSC, 0.1% SDS at room temperature followed by a wash in 1 × SSC, 0.1% SDS at 40 °C.

Northern Blot Analysis—RNA and DNA size markers were denatured in 50% formamide containing 2.2 M formaldehyde and 40 mM Mops, pH 7, at 56 °C for 15 min, followed by 97 °C for 3 min (10). Total RNA from *H. halobium* was subjected to electrophoresis in 1.2% agarose gels containing 2.2 M formaldehyde and blotted to Zeta probe (Bio-Rad). Hybridization (5×10^6 cpm/ml) was in 50% formamide (omitted for oligonucleotide probes), 1% SDS, 0.5% nonfat powdered milk, and 0.01% NaN_3. All other conditions were as described for Southern blot hybridizations.

Primer Extension—Total RNA (10 μg) was incubated with 10 fmol of oligonucleotide primer in 7 μl of RT buffer (50 mM Tris, pH 8.0, 8 mM $MgCl_2$, and 4 mM $MnCl_2$) under two different annealing conditions: 10 min at 70 °C followed by 20 min at 42 °C, or 15 min at 56 °C and 40 min at room temperature. TTP, dGTP, and dCTP were then added to final concentrations of 1 mM each, and dithiothreitol to 10 mM. After addition of [α-*thio*-^{35}S]dATP, the volume was adjusted to 15 μl with RT buffer and the reaction started with 100 units of Molony murine leukemia virus reverse transcriptase. After incubation for 30 min at 42 °C the products of the primer extension reaction were analyzed on a 7% sequencing gel.

Production of Antibodies against Synthetic Peptides—Synthetic peptides (5 μmol), prepared according to the Fmoc synthesis method on a Labortec SP 640 peptide synthesizer, were coupled to 10 mg of hemocyanin according to the method described in Refs. 13 and 14. Peptide-hemocyanin complex (1 mg) dissolved in 500 μl of buffer was

mixed with 500 μl of Freund's complete adjuvant and injected subcutaneously into a rabbit. The rabbit was reinjected 2 and 4 weeks later with 1 mg of peptide adduct/injection. Antiserum was collected 10 days after the final injection. IgG was purified by ammonium sulfate fractionation.

Western Blot Analysis—Purified flagellins from *H. halobium* strain R_1M_1 (9) were separated on 12% SDS-polyacrylamide gels and electrophoretically transferred in buffer containing 25 mM Tris, 114 mM glycine, pH 8.3 (15), to nitrocellulose sheets at 50 V, 4 °C for 4 h. The nitrocellulose sheets were dried and stained with antibodies and fluorescein isothiocyanate-marked goat anti-rabbit IgG, as described for the immunological screening of the halobacterial gene library (3).

RESULTS

Identification of Flagellin Genes—Polyclonal antibodies raised against deglycosylated Fla I polypeptide cross-react with both the Fla II and Fla III polypeptides, indicating a close relationship of all halobacterial flagellins (4). This antibody was used to screen a genomic library of *H. halobium* DNA constructed in the high level expression vector pIN-III A (16). The vector has the following features: the *Escherichia coli* lipoprotein promoter and the 95-bp lac UV 5 promoter-operator region are inserted in tandem, so that a cloned gene is expressed only in the presence of lac inducer. Expression of cloned DNA results in a fusion protein consisting of the amino-terminal prolipoprotein amino acids followed by the amino acids encoded by the insert DNA. Screening of 40,000 clones resulted in the identification of a single immunopositive clone (clone 129). The halobacterial insert DNA of this clone was transferred into the pUC8 vector and its nucleotide sequence determined. Since amino acid sequences of peptides from purified flagella are known (4), the cloned DNA fragment was identified definitively as part of a flagellin gene. However, the cloned DNA encoded only 181 amino acids and lacked the information for the N-terminal portion of the flagellin.

To isolate a complete flagellin gene, a second genomic library of halobacterial DNA constructed in the phage EMBL4 vector (containing 14–15-kb inserts of halobacterial DNA (3)) was screened with insert DNA from clone 129. Out of 5000 phage, 2 clones gave a hybridization signal and were subsequently found to contain identical DNA inserts. *Eco*RI cleaved the insert DNA once, resulting in 4.5- and 11-kb fragments. Only the 4.5-kb fragment hybridized with the flagellin DNA probe and was subsequently subcloned into the pUC8 vector and sequenced. The strategy of sequencing and the restriction map are summarized in Fig. 1*A*. The dideoxy chain termination method was used throughout, using mainly synthetic 17-mer oligonucleotides and the universal pUC primers.

The 1983-bp segment of this 4.5-kb *Eco*RI fragment shown in Fig. 2*A* contains an open reading frame (ORF 1) starting at position 402 and ending at position 992 with a TAA codon. Three stretches of the predicted amino acid sequence were confirmed by sequencing of peptides derived from purified flagella (*underlined* in Fig. 2). Thus the translation of ORF 1 indeed represents the primary structure of a halobacterial flagellin molecule.

Surprisingly, ORF 1 is immediately followed by a second open reading frame (ORF 2) which starts only 11 nucleotides downstream from the stop codon of ORF 1. The predicted amino acid sequence from ORF 2 is nearly identical with that of ORF 1, with 85% of the amino acids conserved in ORF 2. Thus, ORF 2 encodes another flagellin molecule. Subsequently, these tandem genes are called flg A1 and flg A2.

Since purified flagella appear to be composed of three related glycoproteins, at least one additional flagellin gene remained to be identified. A comparison of the nucleotide sequence of original flagellin gene fragment (clone 129) with flagellin genes A1 and A2 revealed an unexpected result: the flagellin gene fragment was not derived from gene flg A1 nor from gene flg A2, because its variable regions did not match the corresponding nucleotide sequences. Therefore, the insert DNA of clone 129 is part of a third flagellin gene. To locate this additional flagellin gene, genomic DNA of *H. halobium* was digested with *Bam*HI and analyzed by a Southern blot experiment, using the 350-bp *Eco*RI fragment of clone 129 as the radioactive probe. Two strong hybridization signals at 5.5 and 15 kb were detected (Fig. 3, *lane 2*). The DNA was reprobed with the synthetic oligonucleotides 5′ CGGTCCC-ACTAGCGGTC 3′ and 5′ CGGTTGAGCCAGACGTG 3′ (sequences unique to flagellin gene flg A1 and flagellin fragment of clone 129, respectively). The 5.5-kb fragment gave a positive signal with the clone 129-specific probe (Fig. 3, *lane 5*); the 11-kb fragment did not. Therefore, the 5.5-kb fraction from a *Bam*HI digest of genomic DNA was ligated directly into the vector pUC8 and the resulting transformants probed

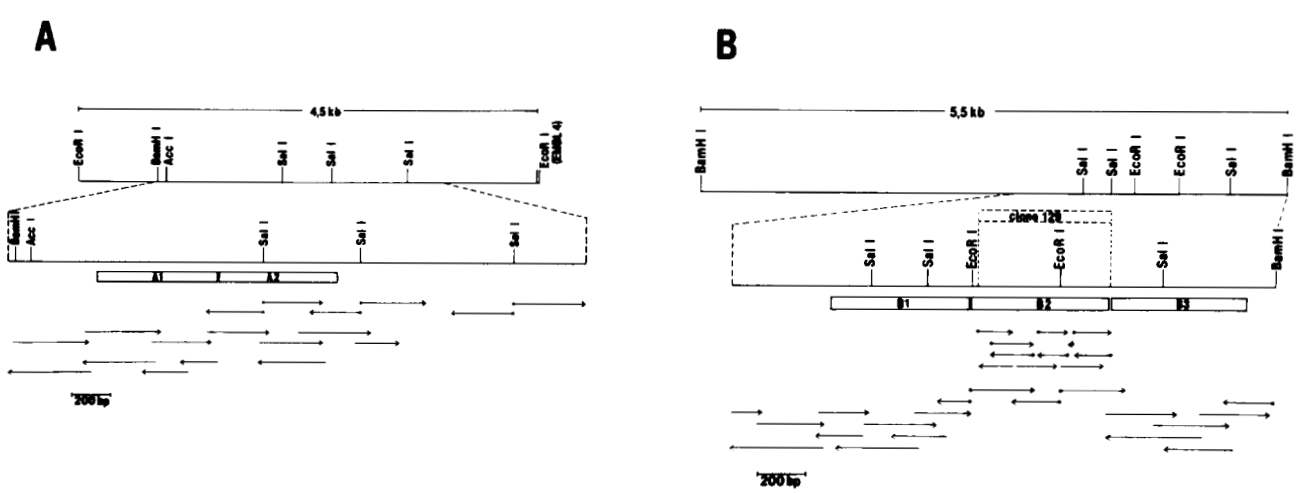

FIG. 1. **Sequencing strategy and restriction maps of the flg A and flg B loci, encoding a total of five related flagellin proteins.** *A*, the 4.5-kb *Eco*RI fragment derived from the EMBL 4 genomic library carries two tandemly arranged flagellin genes, flg A1 and flg A2. *B*, the 5.5-kb *Bam*HI fragment obtained by direct cloning of size-fractionated *Bam*HI-digested genomic DNA. This fragment carries a cluster of three flagellin genes, flg B1, B2, and B3. Sequencing reactions using the universal pUC primers are marked by a *dot*; all other sequencing reactions were primed with synthetic 17-mer oligonucleotides.

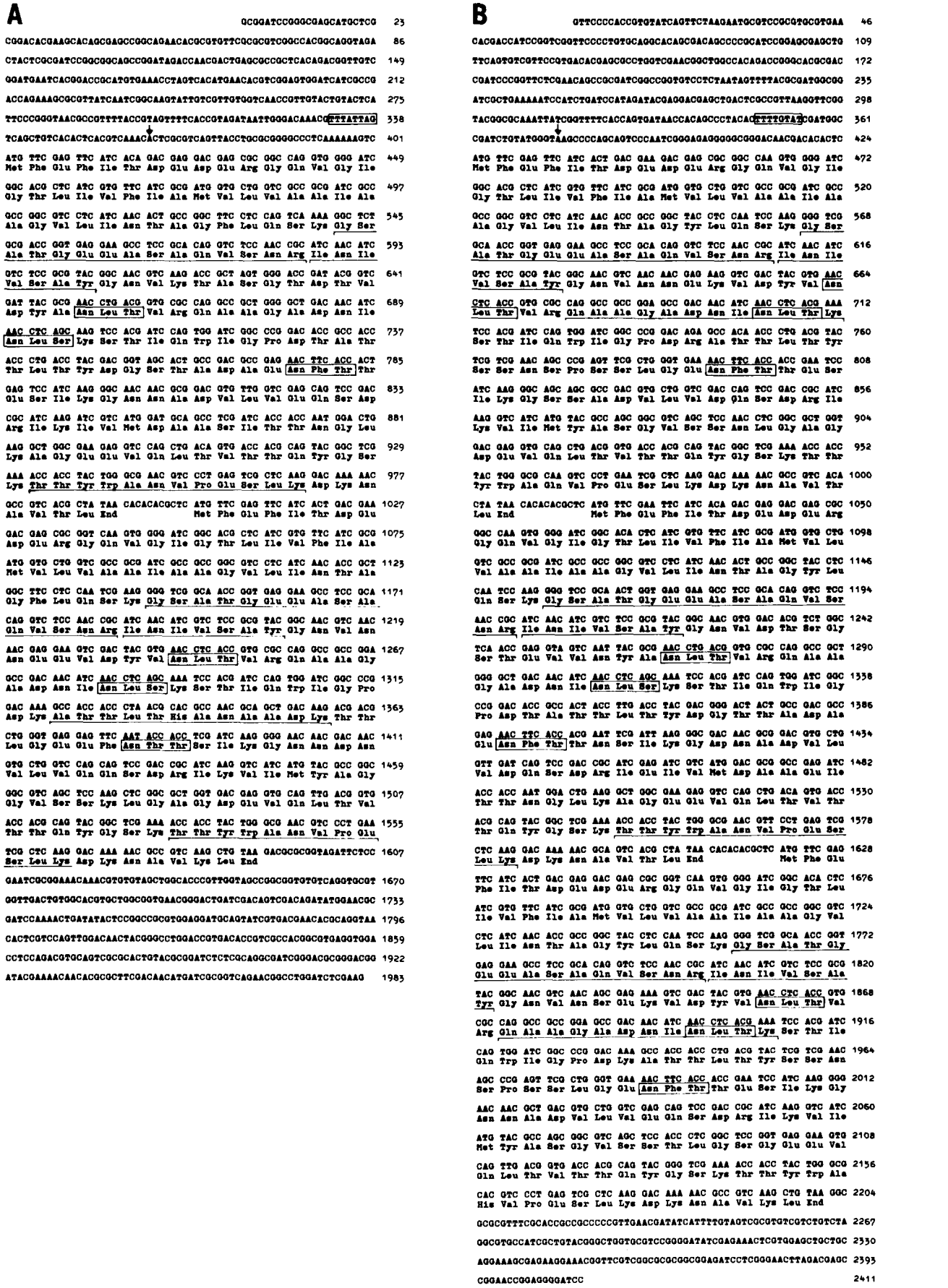

FIG. 2. **Nucleotide sequences and predicted amino acid sequences of genes flg A1 and flg A2 and of the genes flg B1, B2, and B3.** *Numbers* on the *right* indicate nucleotide positions. *A*, flg A locus. The first ORF

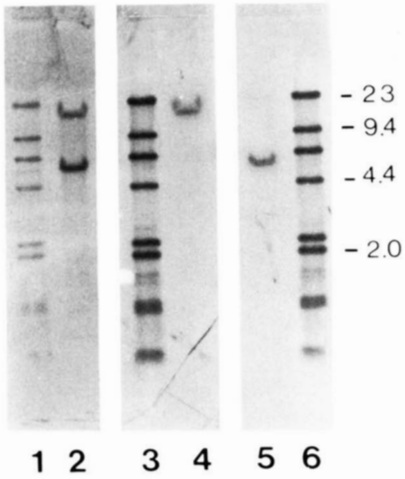

FIG. 3. **Southern blot analysis of genomic DNA.** Halobacterial DNA (2 μg/lane) was digested with *Bam*HI (*lanes 2, 4,* and *5*), separated by electrophoresis in 0.7% agarose, and transferred to a nitrocellulocse sheet. Radioactive size standards (λ *Hin*dIII) were applied to *lanes 1, 3,* and *6*. The filters were hybridized to three different probes. *Lane 2* was probed with the 350-bp *Eco*RI fragment of clone 129 which should recognize all related flagellin genes. In *lane 4*, the flg A1-specific oligonucleotide 5′ CGGTCCCACTAGCGGTC 3′ (complementary to positions 617–633 of flg A) was used as the probe. *Lane 5* was probed with the clone 129-specific oligonucleotide 5′ CGGTTGAGCCAGACGTG 3′ (complementary to positions 1233–1249 of flg B).

```
fla A1   MFEFITDEDERGQVGIGTLIVFIAMVLVAAIAAGVLINTAGFLQSKGSAT  50
fla A2   ..........................................F........  50
fla B1   ..........................................Y........  50
fla B2   ..........................................Y........  50
fla B3   ..........................................Y........  50

fla A1   GEEASAQVSNRINIVSAYGNVKTASGTDTVDYANLTVRQAAGADNINLSK  100
fla A2   ........................----NNEE.D.V..............S.  96
fla B1   ........................----NNEK.D.V..............T.  96
fla B2   ......................DTSGSTEV.N.A..............S.  100
fla B3   ........................----NSEK.D.V..............T.  96

fla A1   STIQWIGPDTATTLTY---DGSTADAENFTTESIKGNNADVLVEQSDRIK  147
fla A2   ..........K.....HANAADKTTLG.E.N.T....NNDN...Q.....K  146
fla B1   ..........R.....Y-SSNSPSSLG.N.T.E....SSAD...D.....K  145
fla B2   ..........T.....Y---DGTTADA.N.T.N....DNAD...D.....E  147
fla B3   ..........K.....Y-SSNSPSSLG.N.T.E....NNAD...E.....K  145

fla A1   IVMDAASITTNGLKAGEEVQLTVTTQYGSKTTYWANVPESLKDKNAVTL•  196
fla A2   -.IMY.GGVSSK.GA.D.....................N..........T.•  194
fla B1   -.IMY.SGVSSN.GA.D.....................Q..........T.•  193
fla B2   I.MDA.EITTNG.KA.E.....................N..........T.•  196
fla B3   -.IMY.SGVSST.GS.E.....................H..........K.•  193
```

FIG. 4. **Comparison of amino acid sequences of flagellin proteins.** The predicted amino acid sequences of all halobacterial flagellins are arranged for maximum sequence identity. Amino acid positions common to all five flagellins are indicated by *dots*. Amino acid deletions are indicated by *dashes*.

with the clone 129-specific oligonucleotide. Three of 150 transformants contained the desired flagellin gene. The restriction map and the sequencing strategy for this 5.5-kb fragment are shown in Fig. 1*B*. The results of the sequencing experiments are summarized in Fig. 2*B*. Surprisingly, the cloned 5.5-kb DNA fragment again contains a flagellin gene family consisting of a cluster of three closely related genes. These genes are called flg B1, B2, and B3. The spacing between each nonoverlapping open reading frame is again 11 nucleotides, *i.e.* each of the three flagellin genes lies within a different reading frame. Each spacer in the two gene families flg A and B has the conserved nucleotide sequence 5′ CACACACGCTC 3′.

Protein Sequence Comparison—In Fig. 4, the predicted amino acid sequences of all five flagellins are shown arranged for maximum homology. Large stretches of the amino acid sequences are highly conserved with the exception of three variable regions centered around positions 80, 120, and 155, respectively. These variable regions were used for the synthesis of either oligonucleotides or peptides which would discriminate between the different flagellins (see below). In their N-terminal region, all flagellins share identical amino acid sequences as long as 70 residues. Within this conserved region the sequence of 19 amino acid residues, Leu-Ile-Val-Phe-Ile-Ala-Met-Val-Leu-Val-Ala-Ala-Ile-Ala-Ala-Gly-Val-Leu-Ile (amino acids 19–37) is composed exclusively of hydrophobic amino acids. This hydrophobic stretch may represent a contact site for a flagellin-flagellin interaction.

Three possible *N*-glycosylation sites (Asn-*X*-Thr (or -Ser)) are located in each gene product (*boxed* amino acid sequences in Fig. 2). Glycosylations at the second glycosylation sites encoded by genes flg B1 or flg B3 were confirmed by amino acid sequencing of peptides derived from purified flagella. Remarkably, *N*-glycosylation sites are located immediately to the C-terminal sides of both the first and second variable regions of the polypeptide chains.

The calculated molecular masses for the gene products of flg A1 and A2 are 20,605 and 20,569 daltons, respectively. Those for the products of genes flg B1, B2, and B3 were calculated to be 20,437, 20,663, and 20,401 daltons, respectively. The N terminus of flagellin Fla I was resistant to Edman degradation. Therefore, a modification or processing at the N terminus may change the actual molecular masses of the mature flagellin protein.

The National Biomedical Research Foundation protein data base (17) was searched for sequence similarity between halobacterial flagellins and other proteins. No significant relationships were revealed, and in particular, no significant sequence similarity was detected with the flagellins of *Bacillus subtilis* (18), *Caulobacter crescentis* (19), *E. coli* (20) and *Salmonella typhimurium* (21).

Expression of Flagellin Genes—To test for *in vivo* expression of genes flg A1 and A2 and genes flg B1, B2, and B3, we hybridized a genomic flagellin sequence to RNAs isolated from exponentially growing halobacteria. When a flagellin gene fragment common to all five flagellin genes (*Eco*RI fragment of flg B2) was used as the probe, two mRNA species were detected with chain lengths of approximately 1300 and 1900 nucleotides, respectively (Fig. 5). However, when probed with oligonucleotides that represented only the nonhomologous regions of the flagellin genes (specific either for the A or B genes) we found that the smaller RNA is derived from the flg A genes and the larger RNA from the flg B gene cluster (Fig. 5).

The initiation sites for both flagellin mRNAs were determined by primer extension. To define the initiation site on the flg A gene, the synthetic oligonucleotide 5′ ACATGACTTTTTTTGAG 3′ (cDNA, nucleotide position 390–405) was incubated with total RNA and primer extension per-

is located between positions 402 and 992 (gene flg A1). Only 11 nucleotides downstream, the second ORF (gene flg A2) starts at position 1004 and ends at position 1588. *B*, flg B locus. Gene flg B1 is located between positions 425 and 1006, flg B2 between 1018 and 1608, and flg B3 between 1620 and 2201. The *boxed* nucleotide sequences may represent halobacterial promoters (22). The *arrows* in flg A and flg B indicate transcription start sites. The *boxed* amino acid sequences represent potential *N*-glycosylation sites. Amino acid sequences confirmed by sequencing of peptides are *underlined*.

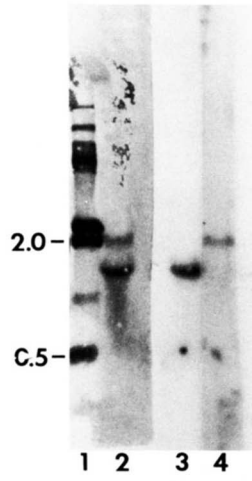

FIG. 5. **Northern blot analysis of halobacterial RNA.** Total RNA was extracted from *H. halobium*, denatured as described under "Materials and Methods," and separated by electrophoresis in 1.2% agarose. Each lane contained 30 μg of RNA, except for *lane 1* which contained denatured DNA size standards (λ *Hin*dIII). Northern transfers were prepared and challenged with a DNA probe hybridizing to all flg sequences (*lane 2*) or with oligonucleotide probes specific for individual flg gene sequences. *Lanes 3* and *4* are with the flg A1-specific probe and the flg B2-specific probe, respectively (see legend to Fig. 3).

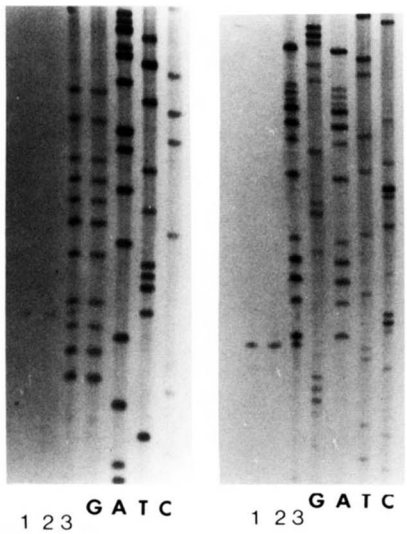

FIG. 6. **Mapping of the transcription initiation site of flg A and flg B.** Primer extension experiments were performed as described under "Materials and Methods." *Left*, the flg A1-specific oligonucleotide 5′ ACATGACTTTTTTGAG 3′ (complementary to positions 390–405) served as primer (*lanes 1* and *2*). The same oligonucleotide was used in a sequencing experiment (dideoxy chain termination) with the 4.5-kb DNA fragment carrying the flg A locus (*lanes G*, *A*, *T*, and *C*). In *lane 3*, the primer extension product was coelectrophoresed with the DNA fragments of *lane G*. *Right*, the same type of experiment was performed with the flg B1-specific oligonucleotide 5′ ACATGAGTGTGTCGTT 3′ (complementary to positions 413–428 of flg B). *Lanes 1*, *2*, *G*, *A*, *T*, and *C* are as described above. The template for the sequencing reaction was the 5.5-kb DNA fragment carrying the flg B locus. In *lane 3*, the primer extension product was coelectrophoresed with the DNA fragments of *lane A*. At both the flg loci, RNA synthesis is initiated with an adenine nucleotide.

formed with Molony murine leukemia virus reverse transcriptase. The same oligonucleotide was used as a primer in a Sanger sequencing experiment with the 4.5-kb DNA fragment as template. The analogous experiment was performed with an oligonucleotide complementary to nucleotides 413–428 of

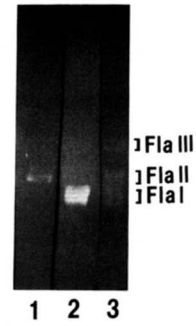

FIG. 7. **Immunological identification of the flg A1 and flg A2 gene products.** Purified flagella (9) containing the related protein populations Fla I, Fla II, and Fla III were fractionated on a 12% SDS-polyacrylamide gel, transferred to nitrocellulose, and probed with an IgG specific for flg A1 peptide (*lane 1*), flg A2 peptide (*lane 2*), or a polyclonal antiflagellin IgG (*lane 3*). The unique peptides used to raise these antibodies are described in the text. Antigen-antibody complexes were visualized by incubation with fluorescein isothiocyanate-conjugated anti-rabbit antibodies.

the flg B gene cluster. Messenger RNA synthesis is initiated on the flg A genes at nucleotide 363, 39 bp upstream of the first ATG codon, and on the flg B genes at nucleotide 376, 49 bp upstream of the first ATG codon (Fig. 6). The estimated chain lengths of both mRNAs indicate that they are polycistronic. This is further supported by specific recognition of the 1.9-kb mRNA by an oligonucleotide probe specific for gene flg B2 (Fig. 5).

A consensus sequence for archaebacterial promoters, based on RNA polymerase binding experiments is 5′ TTTA$\begin{matrix}A\\T\end{matrix}$ATA 3′ and is centered about 25 nucleotides upstream of the transcription initiation site (22). The upstream sequences of the transcription initiation sites of flg A and flg B loci support this hypothesis. At the proposed positions, the flg A and B loci exhibit the sequences TTTATTAG and TTTTGTAT, respectively. In addition, the recently cloned gene of the cell surface glycoprotein also exhibits a similar element, TTTACCAG, at the proposed promoter position (3). As reported for other halobacterial genes (bacteriorhodopsin (23) and halorhodopsin (24)), a potential ribosome binding site is located in all flagellins downstream from the initiation AUG codon. The nucleotide sequence 5′ GGGGATC 3′ complementary (with a single mismatch) to a sequence near the 3′-end of halobacterial 16 S rRNA (3′ CCACUAG 5′ (25)) is found 39 nucleotides downstream of the AUG initiation codon.

Immunological Identification of Gene Products—The flagellins have regions of variable amino acids around positions 80, 120, and 155. Thus, antibodies raised against these unique peptide sequences should distinguish between the flagellin species encoded by the individual flg genes. In particular, the chemically synthesized peptide Lys-Thr-Ala-Ser-Gly-Thr-Asp-Thr-Val-Asp (amino acids 72–81 of gene A1) and the peptide His-Ala-Asn-Ala-Ala-Asp-Lys-Thr-Thr-Leu-Gly-Glu-Glu (amino acids 112–124 of gene A2) are unique to flg A1 and A2, respectively. Therefore, these peptides were cross-linked independently to hemocyanin and used to raise antibodies in rabbits. The resulting antisera were analyzed by Western immunoblotting. Antibodies against the A1-specific peptide recognize only Fla II proteins (30 kDa), whereas those directed against the gene A2-specific peptide selectively bind to the Fla I (26 kDa) protein family (Fig. 7). This immunological approach does not allow us to distinguish the gene products of flg B1 and B3, because only minor variations occur in the amino acid sequences of these genes. The peptide Asp-Thr-Ser-Gly-Ser-Thr-Glu-Val-Val-Asn-Tyr (amino acids

72–82 of gene B2) is unique to the B2 gene product. For unknown reasons, attempts to raise antibodies against this peptide failed. Thus, identification of flg B-derived proteins is only possible after the purification of individual flagellins and subsequent collection of partial amino acid sequence data. This (protein-chemical) approach is currently under way.

DISCUSSION

In eubacteria, flagellin is usually encoded by a single gene. Exceptions are *Caulobacter* (19, 26) and *Bdellovibrio* (27), in which different types of flagellin are encoded by at least three and two genes, respectively. On SDS-polyacrylamide gels, purified halobacterial flagella, even after reversible dissociation, display a ladder-like pattern of bands with three distinct centers of intensity at positions corresponding to 26, 30, and 36 kDa. This pattern suggests three different flagellin polypeptides with variable extents of modification, *e.g.* glycosylations (4). The results described in this paper demonstrate that the structural organization of the halobacterial flagellum is even more complex. We cloned and determined the nucleotide sequences of five closely related flagellin genes, all of which appear to be expressed. Thus, the heterogenous appearance of halobacterial flagellins is explained by at least two different types of structural variations. The polypeptides appear in five different variants with each polypeptide possessing three *N*-glycosylation sites. The sulfated oligosaccharides linked to these glycosylation sites show structural variation with respect to glucuronic acid content and degree of sulfation (28).

Glycosylation of proteins in halobacteria is selectively inhibited *in vivo* by removal of Mg^{2+} ions in the growth medium (2, 6). Under these conditions, the molecular masses of newly synthesized flagellins are shifted toward lower values, the 26-kDa flagellin to 19 kDa, the 30-kDa flagellin to 23 kDa, and the 36-kDa flagellin to 29 kDa. The same shifts in apparent molecular masses are observed upon treatment of purified flagellins with anhydrous hydrogen fluoride (4). This treatment deglycosylates glycoproteins (29). Since all of the flg gene products code for polypeptides with molecular masses around 20 kDa, the appearance of deglycosylated flagellins with molecular masses of 23 and 29 kDa is an unexpected observation. Possibly, another modification of the flagellin molecules, insensitive to hydrogen fluoride treatment, remains to be discovered.

Glycosylation of halobacterial glycoproteins (including the flagellins) occurs at the extracellular surface of the cell membrane (2, 30). As a consequence, the flagellin polypeptides must be translocated across the cell membrane before glycosylation. If so, aggregation to a functional flagellum is likely to occur by a mechanism different from that proposed for the assembly of eubacterial flagella. This latter mechanism assumes transport of flagellin through the central channel of the hook to its tip (31, 32).

At this time, the stoichiometry of the five gene products in the halobacterial filament is unknown. Now that we know the primary structure of all flagellins, we plan to raise additional antisera against synthetic peptides. These antisera should allow by immune electron microscopy a detailed investigation of the structural organization of the halobacterial flagellum.

Acknowledgments—We thank A. Gissibl for excellent technical assistance. We wish to acknowledge the contributions of Drs. F. Lottspeich, G. Paul, and F. Wieland who provided us with amino acid sequences of peptides. We thank Drs. H. Koller and D. Oesterhelt for the preparation of synthetic peptides.

REFERENCES

1. Mescher, M. F., and Strominger, J. L. (1976) *J. Biol. Chem.* **251,** 2005–2014
2. Sumper, M. (1987) *Biochim. Biophys. Acta* **906,** 69–79
3. Lechner, J., and Sumper, M. (1987) *J. Biol. Chem.* **262,** 9724–9729
4. Wieland, F., Paul, G., and Sumper, M. (1985) *J. Biol. Chem.* **260,** 15180–15185
5. Alam, M., and Oesterhelt, D. (1984) *J. Mol. Biol.* **176,** 459–475
6. Sumper, M., and Herrmann, G. (1978) *Eur. J. Biochem.* **89,** 229–235
7. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
8. Hattori, M., and Sakaki, Y. (1986) *Anal. Biochem.* **152,** 232–238
9. Alam, M., and Oesterhelt, D. (1987) *J. Mol. Biol.* **194,** 495–499
10. Maniatis, T., Fritsch, E., and Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
11. Feinberg, A. P., and Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13
12. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517
13. Kris, R. M., Lax, J., Gullick, W., Waterfield, M. D., Ullrich, A., Fridkin, M., and Schlessinger, J. (1985) *Cell* **40,** 619–625
14. Steiner, A. L. (1974) *Methods Enzymol.* **38,** 96–105
15. Towbin, H., Staehlin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 4350–4354
16. Masui, Y., Mizumo, T., and Inouye, M. (1984) *Biotechniques* **2,** 81–85
17. Lipman, D. J., and Pearson, W. R. (1985) *Science* **227,** 1435–1441
18. Delange, R. J., Chang, J. Y., Shaper, J. H., and Glazer, A. N. (1976) *J. Biol. Chem.* **251,** 705–711
19. Gill, P. R., and Agabian, N. (1983) *J. Biol. Chem.* **258,** 7395–7401
20. Kuwajima, G., Asaka, J.-I., Fujiwara, T., Fujiwara, T., Node, K., and Kondo, E. (1986) *J. Bacteriol.* **168,** 1479–1483
21. Joys, T. M. (1985) *J. Biol. Chem.* **260,** 15758–15761
22. Thomm, M., and Wich, G. (1988) *Nucleic Acids Res.* **16,** 151–163
23. Dunn, R., McCoy, J., Simsek, M., Majumdar, A., Chang, S. H., RajBhandary, U. L., and Khorana, H. G. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6744–6748
24. Blanck, A., and Oesterhelt, D. (1987) *EMBO J.* **6,** 265–273
25. Kagramanova, A. S., Mankin, A. S., Baratova, L. A., and Bogdanov, A. A. (1982) *FEBS Lett.* **144,** 177–180
26. Loewy, Z. G., Bryan, R. A., Reuter, S. H., and Shapiro, L. (1987) *Genes Dev.* **1,** 626–635
27. Thomashow, L. S., and Rittenberg, S. C. (1985) *J. Bacteriol.* **163,** 1047–1054
28. Lechner, J., Wieland, F., and Sumper, M. (1985) *J. Biol. Chem.* **260,** 8984–8989
29. Mort, A. J., and Lamport, D. T. A. (1977) *Anal. Biochem.* **82,** 289–309
30. Lechner, J., Wieland, F., and Sumper, M. (1986) *Syst. Appl. Microbiol.* **7,** 286–292
31. Emerson, S. U., Tokuyasu, K., and Simon, M. I. (1970) *Science* **169,** 190–192
32. Homma, M., Fujita, H., Yamaguchi, S., and Iino, T. (1984) *J. Bacteriol.* **159,** 1056–1059