Volume 244, number 1, 137–140 FEB 06808 February 1989

Halobacterial flagellins are encoded by a multigene family

Identification of all five gene products

Lydia Gerl, Rainer Deutzmann and Manfred Sumper

Lehrstuhl Biochemie I, Universität Regensburg, Universitätsstrasse 31, D-8400 Regensburg, FRG

Received 19 December 1988

Flagellins of *Halobacterium halobium* are encoded in five different but homologous genes. Flagellins isolated from purified flagella were digested and the resulting peptides sequenced. The amino acid sequence data obtained prove that all five gene products are expressed and integrated into the flagellar bundle.

Flagellin; Primary structure; (*Halobacterium halobium*)

1. INTRODUCTION

Flagellar bundles of *Halobacterium halobium* consists of 5–10 filaments [1]. Halobacteria swim forward by clockwise and backward by counterclockwise rotation of their right-handed flagellar bundles [2]. These bundles do not fly apart when the sense of rotation changes.

On SDS gels purified flagella display a ladder-like pattern of different protein bands with three centers of intensity, corresponding to molecular masses around 26, 30 and 36 kDa, respectively. As shown recently, the halobacterial flagellins are encoded in five different but highly homologous genes [3]. In addition, all flagellins are sulphated glycoproteins [4]. Therefore, the heterogeneity of the flagellins may be caused by the presence of five different protein species and in addition by different degrees of glycosylation. Here, we prove from the amino acid sequence data the presence of all five gene products in the flagellar bundle.

2. MATERIALS AND METHODS

2.1. *Purification of flagellins*

Purification of flagella from *H. halobium* R_1M_1 was performed essentially as in [5] with the following modifications: Firstly, the flagella and superflagella in the culture fluid were precipitated with 4% polyethylene glycol 6000 and collected by centrifugation at 11000 × *g* for 10 min. The subsequent purification on a CsCl gradient usually displayed a single opaque band. Secondly, the flagellins were separated on 12% SDS-polyacrylamide gels. Areas of interest were cut out and the gel pieces squeezed through a steel net (200 mesh). The resulting gel slurry was then placed in water at 4°C for 4 h. The elution was repeated once. Finally, the eluted proteins were dialysed against water and lyophilized.

Correspondence address: L. Gerl, Lehrstuhl Biochemie I, Universität Regensburg, Universitätsstrasse 31, 8400 Regensburg, FRG

2.2. *Isolation of tryptic peptides*

Isolated flagellins were digested for 4 h at 37°C with TPCK-treated trypsin (enzyme/protein ratio ~1:50) in 0.1 M morpholinoacetate buffer (pH 7.5), 10 mM $CaCl_2$. The digestion was repeated once. The resulting peptide solution was lyophilized and dissolved in 6 M guanidinium chloride. The solution was directly applied to a reversed-phase C18 HPLC column (Lichrospher, Merck) and eluted with a linear gradient from 5 to 35% acetonitrile (90 min) in 0.1% trifluoroacetic acid.

2.3. *Amino acid sequencing*

Amino acid sequences were determined using a Pulsed Liquid Phase Sequencer with an on line PTH-analyzer (models 477A and 120A, Applied Biosystems). The PTH amino acids were separated by gradient elution on a 1.6 × 250 mm narrow-bore ODS-Hypersil column (Shandon, supplied by MZ-Analysentechnik, Mainz), essentially following the method described in Applied Biosystems User Bulletin 14 (1985).

3. RESULTS AND DISCUSSION

Five genes located at two different loci encode

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

the halobacterial flagellins [3]. Two flagellin genes are arranged tandemly at one locus (flg A1,A2), the remaining three being found in a tandem arrangement at a different locus (flg B1,B2,B3).

As predicted from the open reading frames, the polypeptide chains of the individual flagellins contain between 193 and 196 amino acid residues. All five polypeptides share common amino acid sequences with the exception of three variable regions centered around positions 80, 120 and 155, respectively. Peptides derived from these variable regions should allow the identification of all flagellin gene products. Therefore flagella were purified as described in section 2 and the flagellins separated on SDS-polyacrylamide gels, resulting in the typical protein pattern with three centers of intensity around 26, 30 and 36 kDa. Three stripes of the gel were cut out covering these main areas, as indicated in fig.1. Subsequently, the eluted proteins were digested with trypsin and the resulting peptides separated by HPLC on reversed-phase C18 columns. The elution patterns obtained are shown in fig.2 for the 26, 30 and 36 kDa protein population, respectively. The peptides from well-separated peaks were directly submitted to amino acid sequence analysis on an automated gas-phase sequencer. The amino acid sequence data obtained are summarized in table 1.

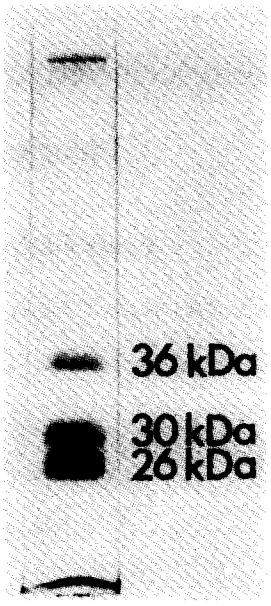

Fig.1. SDS-polyacrylamide gel electrophoresis (12%) of purified flagella, showing a ladder-like pattern of protein bands with centers of intensity around molecular masses of 26, 30 and 36 kDa, respectively.

A number of peptides are common to two or more flagellins. These peptides, summarized in column 4 of table 1 are not suitable for the une-

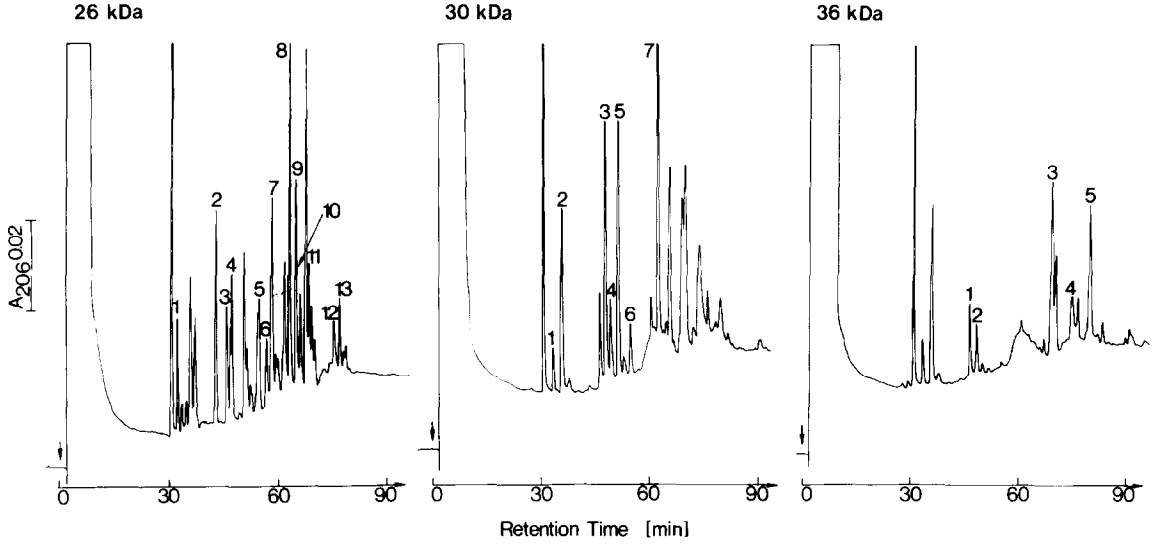

Fig.2. Tryptic maps of flagellins with molecular masses around 26, 30 and 36 kDa, respectively. Peptides were separated by HPLC on a reversed-phase C18 column as described in section 2. The material of numbered peaks was selected for amino acid sequencing and the corresponding sequences are listed in table 1.

Table 1

Amino acid sequences of tryptic peptides from 26, 30 and 36 kDa flagellins

Peptide	Amino acid sequence	Unique to gene	Common to genes
26 kDa/1	ATTLTHA. . .	A2	–
2	GNNDNVLVQQSDR	A2	–
	VIMYAGGVSSK	A2	–
3	NAVK	B3	–
4	GNNA. . .	–	A1,B3
5	TTLGEEF-TTSIK	A2	–
	VDYV-LTVR	–	B1,B3
6	VIMYASGVSS. . .	–	A2,B1,B3
7	ATTLTYSSNS. . .	–	B1,B3
8	STIQWIGPDK	–	A2,B3
9	STIQWIGPDR	B1	–
	LGAGDEVQLTVTTQY. . .	–	A2,B1
10	TTYWAHVPESLKDK	B3	–
11	TTYWAQVPESL. . .	B1	–
12	TTYWAHVPESLKDKNA. . .	B3	–
	INIVSAYGNVNNEKVD. . .	B1	–
13	TTYWANVP. . .	–	A1,A2,B2
30 kDa/1	QAAGADNI-LSK	–	A1,A2,B2
2	QAAGADNI-LSK	–	A1,A2,B2
3	GNNADVLVEQSDR	–	A1,B3
4	TASGTDTVDYA-LTVR	A1	–
	TYDGSTADAE-FTTE	A1	–
5	TASGTDTVDYA-LTVR	A1	–
6	IVMDAASITTNGLK	A1	–
7	INIVSAYGNVK	A1	–
36 kDa/1	NAVTL	–	A1,A2,B1,B2
2	GDNADVLVDQSDR	B2	–
3	TTYWANVPESLK	–	A1,A2,B2
4	STIQWIGPDTATTLTYDGTTADAE-FT. . .	B2	–
5	IEIVMDAAEI. . .	B2	–
	STIQWIGPDT. . .	–	A1,B2

quivocal identification of a given gene product. Only unique peptides are listed in column 3 and these data allow the following conclusions:

(i) Flagellins with apparent molecular masses around 26 kDa are encoded by the genes flg A2, flg B1 and flg B3.

(ii) Only the gene product of gene flg A1 is found to be present in the population of 30 kDa flagellins.

(iii) Only the gene product of gene flg B2 is found to be present in the population of 36 kDa flagellins.

Taken together, these data prove the expression of all five flagellin genes. Furthermore, all five gene products are shown to be intrinsic to the flagellar structure. At present, it is not known whether their distribution within the flagellar bundle is random or ordered. Immune electron microscopy using antibodies raised against unique peptides should allow us to resolve this question.

Peptides 30 kDa/1 and 30 kDa/2 (table 1) which are well separated by HPLC on reversed-phase C18 columns (fig.2) turned out to be identical with respect to amino acid sequence (table 1). Both display the sequence QAAGADNI-LSK. The expected asparagine at position 9 is not detectable, indicating *N*-glycosylation at this site. Therefore, the different chromatographic behaviour of these two peptides is most probably explained by structural variations in the oligosaccharides attached at

this site. Previous bisynthetic investigations had indeed shown structural heterogeneity of the oligosaccharides attached to halobacterial glycoproteins with respect to glucuronic acid content [6]. Thus, different oligosaccharide variants appear to be linked to one and the same glycosylation site of this halobacterial glycoprotein.

Acknowledgement: This work was supported by the Deutsche Forschungsgemeinschaft (SFB 43).

REFERENCES

[1] Houwink, A.L. (1956) J. Gen. Microbiol. 15, 146–150.
[2] Alam, M. and Oesterhelt, D. (1984) J. Mol. Biol. 176, 459–475.
[3] Gerl, L. and Sumper, M. (1988) J. Biol. Chem. 263, 13246–13251.
[4] Wieland, F., Paul, G. and Sumper, M. (1985) J. Biol. Chem. 260, 15180–15185.
[5] Alam, M. and Oesterhelt, D. (1987) J. Mol. Biol. 194, 495–499.
[6] Lechner, J., Wieland, F. and Sumper, M. (1985) J. Biol. Chem. 260, 8984–8989.