Analysis of the flagellin (hag) gene of alkalophilic Bacillus sp. C-125

YOH-ICHIRO SAKAMOTO,* KATHARINE J. SUTHERLAND, JIN TAMAOKA, TETSUO KOBAYASHI, TOSHIAKI KUDO and KOKI HORIKOSHI

The RIKEN Institute and JRDC, Hirosawa 2-1, Wako-shi, Saitama, 351-01, Japan

(Received 17 March 1992; accepted 23 June 1992)

Motility of the alkalophilic Bacillus sp. C-125, a flagellate bacterium, was demonstrated to be Na+- and pH-dependent. Flagellin protein from this strain was purified to homogeneity and the N-terminal sequence determined. Using the hag gene of Bacillus subtilis as a probe, the hag gene of Bacillus sp. C-125 was identified and cloned into Escherichia coli. Sequencing of this hag gene revealed that it encodes a protein of 272 amino acids (M, 29995). The predicted N terminal sequence of this protein was identical to that determined by N-terminal sequencing of the flagellin protein from strain C-125. The alkalophilic Bacillus sp. C-125 flagellin shares homology with other known flagellins in both the N- and C-terminal regions. The middle portion, however, shows considerable differences, even from that of flagellin from the related species, B. subtilis.

Introduction

The flagellar apparatus has been extensively studied in neutrophilic bacteria such as Escherichia coli and Bacillus subtilis and has been shown to be a complex structure composed of a basal body, a hook and a filament. The filament accounts for 98% of the mass of the organelle and is composed exclusively of flagellin, the product of the hag gene (Iino, 1977; Silverman & Simon, 1977).

Flagellin plays only a passive role in motility, the mechanical energy for flagellar rotation being generated by a motor in, or associated with, the basal body structure of the organelle (Ridgeway et al., 1977; Dean et al., 1984; Blair & Berg, 1990). Nevertheless, flagellin has proved a focus of interest due to its self-assembling properties (Asakura et al., 1964; Macnab & Aizawa, 1984), its method of transport across the cell membrane (Homma et al., 1987; LaVallie & Stahl, 1989) and other features such as its antigenicity.

The primary amino acid sequences of flagellins from a number of neutrophilic bacteria have been determined (Gill & Agabian, 1983; Joys, 1985; Wei & Joys, 1985; Kuwajima et al., 1986; Martin & Savage, 1988; LaVallie & Stahl, 1989; Mirel & Chamberlin, 1989). The flagellin of alkalophilic bacteria, being exposed to the alkaline environment in which these organisms live, is likely to show special adaptations to alkaline conditions. In addition, flagellin, unlike other extracellular proteins so far studied, as a polymer is a good model for the study of subunit-subunit interactions under alkaline conditions. However, hitherto, there have been no reports on the flagellin gene of alkalophiles.

We have selected the alkalophilic Bacillus sp. C-125 as a model alkalophile for molecular biological study for the following reasons: (i) it grows well on minimal medium, (ii) genes from this strain are expressed well in E. coli, and (iii) an efficient system for plasmid transformation of this strain is now available (Kudo et al., 1990). In this paper, we report the cloning and sequencing of the hag gene from the alkalophilic Bacillus sp. C-125.

Methods

Bacterial strains and media. Alkalophilic Bacillus sp. C-125 10596 (Trp- Ura- Cinr) was isolated in our laboratory (Honda et al., 1985; Kudo et al., 1990). It was grown on Horikoshi-II medium (pH 10.3) containing 1% (w/v) soluble starch, 0.5% yeast extract, 0.5% polypeptone, 0.1% KH2PO4, 0.02% MgSO4·7H2O, and 1% Na2CO3 (sterilized separately). The Bacillus DNA was cloned in E. coli MV1184 (ara Δlac–pro–AB) rpsL thy (φ80 lacZ ΔM15) Δ(srl–recA)306 Δ: Tn10 (Tet') F' [traD36 proAB lacF lacZ Δ(M15)]. E. coli was grown on 2× yeast/tryptone medium (Sambrook et al., 1989).

Plasmids, enzymes and chemicals. pUC vectors, restriction endonucleases and T4 DNA ligase were purchased from Takara Shuzo Co. The DIG (digoxigenin) DNA labelling and detection kit was obtained from
Boehringer Mannheim. Hybond N+ and Hybond N membranes were purchased from Amersham. Sequencing oligonucleotide primers were synthesized on an Applied Biosystems 380B DNA synthesizer. HindIII-digested \( \hat{H} \)aeIII-digested flx174 DNA standards (23130, 9416, 6557, 4361, 2322, 2027, 1353, 1078, 872 and 603 bp) were purchased from Toyobo. Protein standards, purchased from Pharmaex, consisted of \( \alpha \)-lactalbumin (14400), soybean trypsin inhibitor (20100), carbonic anhydrase (30000), ovalbumin (43000), bovine serum albumin (67000) and phosphorylase b (94400).

Electron microscopy. Flagellation of alkalophilic Bacillus sp. C-125 was observed by transmission electron microscopy with a JEOL model 200CX microscope at 100 kV after negative staining with 1\% (w/v) phosphotungstic acid.

Physical characterization of motility. Swimming speeds were measured using a television monitor attached to a microscope. Alkalophilic Bacillus sp. C-125 was grown aerobically to late exponential phase at 37 °C in Horikoshi-I1 medium (pH 10.3), collected (20 \( \mu \)g), carbonic anhydrase (30000), ovalbumin (43000), bovine serum albumin (67000) and phosphorylase b (94400).

Electron microscopy. Flagellation of alkalophilic Bacillus sp. C-125 was observed by transmission electron microscopy with a JEOL model 200CX microscope at 100 kV after negative staining with 1\% (w/v) phosphotungstic acid.

Physical characterization of motility. Swimming speeds were measured using a television monitor attached to a microscope. Alkalophilic Bacillus sp. C-125 was grown aerobically to late exponential phase at 37 °C in Horikoshi-I1 medium (pH 10.3), collected and washed twice. Cells were resuspended in a 25 mM-Tris/HCl buffer (pH 9.0), 5 mM-glucose solution containing 100 mM-NaCl unless indicated otherwise.

Purification of flagellin protein. Flagellin protein of alkalophilic Bacillus sp. C-125 was purified to homogeneity by the method of Martinez (1963). SDS-PAGE was carried out as described by Laemml (1970).

N-terminal sequence of the purified flagellin. The method of Edman & Henschen (1975) was used, with an Applied Biosystems 477A protein sequencer coupled to a PTH120A analyser.

Amino acid analysis. This was done on an Applied Biosystems PTH120A analyser.

Preparation of genomic DNA. The method of Saito & Miura (1963) was used.

Preparation of probe DNA. The B. subtilis hag gene used as probe was constructed by PCR using the oligonucleotides 5' ACGTGCCTTTA-CAACATATT 3' and 5' ATGAGGAATGATTAGGAGAT 3' as primers to regions upstream and downstream of the B. subtilis hag gene, respectively (LaVallie & Stahl, 1989; Mirel & Chamberlin, 1989).

Southern blot analysis. Genomic DNA (10 \( \mu \)g) of alkalophilic Bacillus sp. C-125 was digested with restriction endonucleases, separated on a 1\% (w/v) agarose gel and then blotted using a vacuum blotter (Pharmacia LKB) on to Hybond N+. Blots were probed with digoxigenin-labelled B. subtilis hag gene in a 20\% (v/v) formamide hybridization solution at 42 °C. Subsequent washing and detection steps were carried out according to instructions provided with the Boehringer Mannheim DIG-detection kit.

Constructing and screening a sub-genomic library. Genomic DNA (100 \( \mu \)g) of alkalophilic Bacillus sp. C-125 was digested with the appropriate restriction endonuclease and separated by electrophoresis on a 1\% (w/v) agarose gel. DNA of the desired fragment size was recovered from the agarose by electroelution, treated with phenol/chloroform and then precipitated with ethanol. The size-selected DNA was cloned into pUC vector in E. coli MV1184. Colony blots were prepared according to instructions provided with the Hybond N membrane and hybridized with digoxigenin-labelled B. subtilis hag gene as described above for Southern blot analysis.

DNA sequencing. This was done by the technique of Sanger et al. (1977) on a Dupont DNA Sequencer using single-stranded pUC as template and a Dupont sequencing kit.

Sequence alignment. Flagellin sequences were aligned using the GENETYX amino acid sequence homology program (Tokyo, Japan).

Results and Discussion

Characterization of the motility of alkalophilic Bacillus sp. C-125

Alkalophilic Bacillus sp. C-125 is a flagellate bacterium (Fig. 1). Whereas in neutrophilic bacteria flagellar rotation is powered by H⁺-driven motors, in alkalophiles, apparently because the protonmotive force is too low, rotation is powered by Na⁺-driven motors (Hirot et al., 1981). The effect of Na⁺ concentration on motility of alkalophilic Bacillus sp. C-125 cells is shown in Fig. 2. Maximum motility was observed at 50 mM-NaCl. When other cations (in the form of KCl, LiCl, NH₄Cl, RbCl, CsCl, CaCl₂ and MgCl₂) were used in place of Na⁺ no motility was observed. In addition, amiloride, a specific inhibitor of Na⁺-driven flagellar motors (Sugiyama et al., 1988) inhibited movement of alkalophilic Bacillus sp. C-125 cells at concentrations greater than 1 mM. We conclude, therefore, that as in other alkalophiles, the flagellar motor of alkalophilic Bacillus sp. C-125 is driven by Na⁺.

pH also affected the motility of alkalophilic Bacillus sp. C-125 cells, with no motility being observed at pH values less than 7.5. Microscopic observation of cells transferred from alkaline to neutral pH revealed that cells at the new lower pH were no longer flagellate (data not shown). This loss of flagella presumably accounts for the loss of motility at pH values below 7.5. Kamiya & Asakura (1976) showed that Salmonella flagella displayed a remarkable reversible polymorphism when transferred from neutral pH to pH 4-7 and lower. It is possible, therefore, that the loss of flagella by alkalophi-
The hag gene of alkalophilic Bacillus sp. C-125 at pH values below 7.5 is caused by a polymorphic transition of the flagellin filament which results in the flagella becoming unstable. As yet we have no data to support this speculation; however, it may be possible to test it by observing whether alkalophilic Bacillus sp. C-125 filaments reassociate around seeds of polymerized flagella when the pH is raised from neutral pH back to alkaline pH.

The flagellin protein of alkalophilic Bacillus sp. C-125

Flagellin purified from alkalophilic Bacillus sp. C-125 was pure as judged from the single band obtained on SDS-PAGE (Fig. 3). By comparison with standard proteins the polypeptide $M_r$ of the alkalophilic Bacillus sp. C-125 flagellin is 31000.

The N-terminal sequence of the alkalophilic Bacillus sp. C-125 flagellin was determined to be as follows: Met-Ile-Ile-Asn-His-Asn-Leu-Pro-Ala-Met-.

The amino acid content of the alkalophilic Bacillus sp. C-125 flagellin, as determined by amino acid analysis of the protein, is shown in column 2 of Table 1.

The hag gene of alkalophilic Bacillus sp. C-125

Fig. 4 shows the result of Southern blot analysis of Bacillus sp. C-125 genomic DNA cut with several endonucleases and probed with the B. subtilis hag gene. From this result it was apparent that the alkalophilic Bacillus sp. C-125 hag gene is not carried within a single EcoRI or HindIII fragment. The 1.5 kb HindIII fragment and the 1.2 kb EcoRI fragment (Fig. 4) were cloned separately, and subsequently joined at the common EcoRI site. The plasmid containing the complete alkalophilic Bacillus sp. C-125 hag gene was designated pC125Hg.

Sequencing of the inserted DNA fragment on pC125Hg resulted in the identification of an open reading frame containing the 10 amino acids determined by N-terminal sequencing of the protein (see Fig. 6). This open reading frame is immediately preceded by an ATG codon which we propose acts as the translation initiation
The amino acid composition of the alkalophilic Bacillus sp. C-125 flagellin, as determined from the sequence, is shown in Table 1, column 3. It compares favourably with the composition determined by direct amino acid analysis of the purified protein (column 2).

For comparison, the amino acid compositions of B. subtilis and Bacillus firmus RAB flagellins are also included in Table 1. B. firmus RAB flagellin contains few basic amino acids, a feature that has been proposed to make the flagellum more stable at high pH (Guffanti & Eisenstein, 1983). The alkalophilic Bacillus sp. C-125 flagellin composition, however, does not appear to be significantly different from the B. subtilis composition, the lysine content being only slightly lower.

Alignment with other known flagellin sequences

The amino acid sequence of the alkalophilic Bacillus sp. C-125 flagellin was compared to the sequences of E. coli (Kuwajima et al., 1986), S. typhimurium (Ji uys, 1985), B. subtilis

Table 1. Amino acid compositions of the flagellins of alkalophilic Bacillus sp. C-125 and related bacteria

<table>
<thead>
<tr>
<th>Residue</th>
<th>C-125*</th>
<th>C-125†</th>
<th>B. firmus RAB‡</th>
<th>B. subtilis§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>16 (5.8)</td>
<td>12 (4.4)</td>
<td>41 (11.7)</td>
<td>19 (6.3)</td>
</tr>
<tr>
<td>Thr</td>
<td>14 (5.1)</td>
<td>14 (5.2)</td>
<td>43 (12.3)</td>
<td>18 (5.9)</td>
</tr>
<tr>
<td>Ile</td>
<td>16 (5.8)</td>
<td>19 (7.0)</td>
<td>21 (6.0)</td>
<td>22 (7.2)</td>
</tr>
<tr>
<td>Met</td>
<td>10 (3.6)</td>
<td>11 (4.0)</td>
<td>3 (0.9)</td>
<td>8 (2.6)</td>
</tr>
<tr>
<td>Asp</td>
<td>20 (7.4)</td>
<td>-</td>
<td>-</td>
<td>22 (7.2)</td>
</tr>
<tr>
<td>Arg</td>
<td>16 (5.8)</td>
<td>15 (5.5)</td>
<td>8 (2.3)</td>
<td>14 (4.6)</td>
</tr>
<tr>
<td>Phe</td>
<td>4 (1.4)</td>
<td>3 (1.1)</td>
<td>2 (0.6)</td>
<td>5 (1.6)</td>
</tr>
<tr>
<td>Ala</td>
<td>33 (12.0)</td>
<td>33 (12.1)</td>
<td>84 (24.0)</td>
<td>39 (12.8)</td>
</tr>
<tr>
<td>Val</td>
<td>11 (4.0)</td>
<td>11 (4.0)</td>
<td>13 (3.7)</td>
<td>14 (4.6)</td>
</tr>
<tr>
<td>Pro</td>
<td>3 (1.1)</td>
<td>2 (0.7)</td>
<td>0 (0)</td>
<td>2 (0.7)</td>
</tr>
<tr>
<td>His</td>
<td>5 (1.6)</td>
<td>6 (2.2)</td>
<td>3 (0.9)</td>
<td>4 (1.3)</td>
</tr>
<tr>
<td>Glu</td>
<td>21 (7.7)</td>
<td>-</td>
<td>-</td>
<td>17 (5.6)</td>
</tr>
<tr>
<td>Asn</td>
<td>21 (7.7)</td>
<td>-</td>
<td>-</td>
<td>27 (8.9)</td>
</tr>
<tr>
<td>Trp</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Asx</td>
<td>40 (14.5)</td>
<td>-</td>
<td>66 (18.9)</td>
<td>-</td>
</tr>
<tr>
<td>Glx</td>
<td>48 (17.4)</td>
<td>-</td>
<td>34 (9.7)</td>
<td>-</td>
</tr>
<tr>
<td>Ser</td>
<td>17 (6.2)</td>
<td>17 (6.3)</td>
<td>0 (0)</td>
<td>24 (7.9)</td>
</tr>
<tr>
<td>Leu</td>
<td>30 (10.9)</td>
<td>31 (11.0)</td>
<td>23 (6.5)</td>
<td>29 (9.5)</td>
</tr>
<tr>
<td>Cys</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Tyr</td>
<td>2 (0.7)</td>
<td>1 (0.4)</td>
<td>2 (0.6)</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>Gln</td>
<td>-</td>
<td>27 (9.9)</td>
<td>-</td>
<td>24 (7.9)</td>
</tr>
<tr>
<td>Lys</td>
<td>11 (4.0)</td>
<td>9 (3.3)</td>
<td>7 (2.0)</td>
<td>15 (4.9)</td>
</tr>
</tbody>
</table>

* Obtained by direct amino acid analysis of the purified flagellin.
† Deduced from the hag gene sequence.
‡ Cited from Guffanti & Eisenstein (1983).
§ Derived from data reported by Mirel & Chamberlin (1989).
The *hag* gene of alkalophilic *Bacillus* sp. C-125

The alignment with the *B. subtilis* flagellin sequence is shown in Fig. 7. Whereas the N-terminal region (amino acids 1–96) and C-terminal region (amino acids 200–272) show 74 and 82% identity, respectively, to the comparable regions in the *B. subtilis* flagellin, the central region (amino acids 97–199) shows only 32% identity. Such a pattern of similarity in N- and C-terminal regions and dissimilarity in the central regions has been observed for...
Fig. 7. Alignment of the alkalophilic *Bacillus* sp. C-125 (lower) and the *R. subtilis* (upper) flagellin amino acid sequences. Identical amino acids and homologous amino acids are highlighted by asterisks and dots, respectively. Amino acids conserved in neutrophilic flagellins but altered in the alkalophilic flagellin are indicated by the vertical arrows. Domains 1, 2 and 3 [as proposed by Namba *et al.* (1989) for the *S. typhimurium* flagellin] are shown.

In addition to the apparent structural role of the N- and C-termini, these regions are believed to be involved in filament assembly (Joys, 1985; Wei & Joys, 1985; Trachtenberg & DeRosier, 1988). Moreover, the extreme C-terminal region of the *B. subtilis* flagellin has been demonstrated to be involved in its transport from the cytoplasm to the outer membrane (LaVallie & Stahl, 1989).

The central, variable region of the flagellin peptide has not been implicated with any structural or functional properties of the protein. In the structures proposed by Homma *et al.* (1987) and Namba *et al.* (1989) it is the central portion that forms the outer surface of the flagellin filament, and thus, this portion is exposed to the outer environment of the cell. For *Bacillus* sp. C-125 the environment is alkaline and so the central portion of the
flagellin, on the outer surface of the flagellar filament, is likely to require special adaptations to these alkaline conditions. This may explain the very low level of homology observed between the central regions of the alkalophilic *Bacillus* sp. C-125 and *B. subtilis* flagellins.

Decreased amounts of acidic and/or basic amino acids have been observed in some alkalophilic proteins (Guffanti & Eisenstein, 1983; Van der Laan et al., 1991). The overall amino acid content of the alkalophilic *Bacillus* sp. C-125 does not appear to be significantly different from that of the *B. subtilis* flagellin (Table 1). However, the alkalophilic *Bacillus* sp. C-125 domain 3 does appear to have fewer charged amino acids than the equivalent region in the *B. subtilis* flagellin (Fig. 7). Thus, it seems that, if we confine the comparison to regions that are presumed to be exposed to alkaline conditions, the low content of basic or acidic amino acids observed in other alkalophilic proteins (Guffanti & Eisenstein, 1983; Van der Laan et al., 1991) can also be observed in the alkalophilic *Bacillus* sp. C-125 flagellin.

Due to the high level of variation among the flagellin central regions it is difficult to pinpoint and investigate possible changes in the alkalophilic *Bacillus* sp. C-125 flagellin that may contribute to its adaptation to alkaline conditions. However, the conserved N- and C-terminal regions (domains 1 and 2) are also likely to show adaptations to alkalinity since these, although not situated on the outer surface of the flagellum, may, for example, be exposed during the assembly process. Moreover, Kanto et al. (1991) have shown that amino acids responsible for flagellin shape are located in the terminal regions of the *S. typhimurium* flagellin. A comparison of the N- and C-terminal regions revealed that only two amino acids, an Ala (Gln 22 of *Bacillus* sp. C-125 flagellin, domain 1) and a Val (Met 206 of *Bacillus* sp. C-125 flagellin, domain 2) (Fig. 7), conserved in all four neutrophilic bacterial flagellins are different in the alkalophilic *Bacillus* sp. C-125 flagellin. Site-directed mutagenesis work is currently in progress to investigate whether these changes contribute to the alkaline adaptation of alkalophilic *Bacillus* sp. C-125 flagellin.

We wish to thank Ms M. Yamamoto for carrying out the N-terminal sequence determination. This work was partially supported by a grant for the Biodesign Research Program from RIKEN to T. Kudo.

References


