Extremophiles (1997) 1:29–35

ORIGINAL PAPER

Hiromichi Wakai · Satoshi Nakamura
Hiroshi Kawasaki · Kei Takada · Shin Mizutani
Rikizo Aono · Koki Horikoshi

Cloning and sequencing of the gene encoding the cell surface glycoprotein of *Haloarcula japonica* strain TR-1

Received: September 20, 1996 / Accepted: October 9, 1996

Abstract The triangular disk-shaped halophilic archaeon *Haloarcula japonica* strain TR-1 has a glycoprotein on its cell surface. The complete gene encoding the cell surface glycoprotein (CSG) was cloned and sequenced. The gene has an open reading frame of 2586 bp, and a potential archaeal promoter sequence approximately 150 bp upstream of the ATG initiation codon. The mature CSG is composed of 828 amino acids and is preceded by a signal sequence of 34 amino acid residues. A hydropathy analysis showed a hydrophobic stretch at the C-terminus, that probably serves as a transmembrane domain. The amino acid sequence of the *Ha. japonica* CSG showed 52.1% and 43.2% identities to those from the *Halobacterium halobium* and *Haloferax volcanii* CSGs, respectively. Five potential N-glycosylation sites were found in the mature *Ha. japonica* CSG, sites that were distinctly different from those in *Hb. halobium* and *Hf. volcanii*. The *Ha. japonica* CSG gene was expressed in *Escherichia coli*.

Key words Halophilic archaea · Surface layer · Cell surface glycoprotein · *Haloarcula japonica* · Polymerase chain reaction · Gene cloning

Communicated by W.D. Grant

H. Wakai · S. Nakamura (✉) · H. Kawasaki · K. Takada · S. Mizutani · R. Aono · K. Horikoshi[1]

Department of Bioengineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama-shi, Kanagawa 226, Japan
Tel. +81-45-924-5765; Fax +81-45-924-5819
e-mail: snakamur@bio.titech.ac.jp

Present address:
[1] Department of Applied Chemistry, Toyo University, 2100 Nakanodai, Kujirai, Kawagoe-shi, Saitama 350, Japan

Introduction

Crystalline surface layers (S-layers) are frequently found on eubacterial and archaeal cells (Sleytr and Messner 1992). Being located at the cell surface, S-layers have been suggested to be important in the adaptation of microorganisms to environments. In most cases, the S-layer consists of a single protein or glycoprotein subunit ranging from 40 to 200 kDa in size, and the subunits assemble on the cell surface into a two-dimensional symmetric array (Sleytr and Messner 1992).

Mescher and Strominger (1976) were the first to show that the cell surface glycoprotein (CSG) of the halophilic archaeon *Halobacterium halobium* is essential for maintaining the rod shape of the cell. A similar glycoprotein is found also in a flat-disk-shaped halophilic archaeon, *Haloferax volcanii* (Kessel et al. 1988). Halophilic archaea (halobacteria) lack the peptidoglycan layer, and the CSG is the only wall component. Thus, the S-layer might act as a shape-maintaining corset in these organisms. A cell envelope profile of *Hb. halobium* has been described on the basis of X-ray analysis data (Blaurock et al. 1976). A three-dimensional structure model of the cell envelope of *Hf. volcanii* has been proposed from an electron microscopic study (Kessel et al. 1988). To date, genes encoding halobacterial CSGs have been cloned and sequenced from *Hb. halobium* (Lechner and Sumper 1987) and *Hf. volcanii* (Sumper et al. 1990). Furthermore, the chemical structures of the saccharide chains attached to these CSGs have been determined (Sumper 1987; Sumper et al. 1990; Mengele and Sumper 1992).

Haloarcula japonica strain TR-1 is a predominantly triangular disk-shaped halobacterium (Horikoshi et al. 1993). Although taxonomic characteristics (Takashina et al. 1990) and mode of cell division (Hamamoto et al. 1988) have been extensively studied, the mechanisms that maintain the triangular shape are still unclear. Previously, we have demonstrated the occurrence of a CSG that seems to be important in maintaining the characteristic shape (Nakamura et al. 1992). Furthermore, a recent microscopic study (Nishiyama et al. 1992) showed arrays of hexagonal structures arranged

on the cell surface of *Ha. japonica*. The primary structure of the *Ha. japonica* CSG, however, has not yet been reported.

In the present study, we have cloned and sequenced the gene encoding the *Ha. japonica* CSG. The primary structure of the *Ha. japonica* CSG, derived from the cloned gene, was then compared with those of other halobacterial CSGs. Expression of the *Ha. japonica* CSG gene in *Escherichia coli* has also been demonstrated.

Materials and methods

Strains, plasmids, and media. *Haloarcula japonica* strain TR-1 (JCM 7785, ATCC 49778) was grown at 37°C in a liquid medium as described previously (Nishiyama et al. 1995). *Escherichia coli* strains JM109 (for gene cloning) and MV1184 (for gene expression) were the hosts and were cultivated at 37°C in L broth (Sambrook et al. 1989). Plasmids pUC119 and pBluescript II SK+ were obtained from Takara Shuzo (Kyoto, Japan) and Toyobo (Osaka, Japan), respectively.

Isolation of chromosomal DNA. Chromosomal DNA was prepared according to the method of Lechner and Sumper (1987).

Amplification of DNA by PCR. The polymerase chain reaction (PCR) primers were synthesized with an Applied Biosystems model 391 (PCR-MATE) DNA synthesizer (CA, USA). The PCR mixture consisted of 1.5 µg of *Ha. japonica* chromosomal DNA, 100 pmol of primers, 0.25 mM each deoxynucleotide triphosphate, and 2.5 U of TET-Z polymerase (Amersham, England) in 100 µl of the reaction buffer recommended by the manufacturer. Thirty cycles of amplification (1.3 min at 97°C for denaturation, 2 min at 55°C for annealing, and 3 min at 72°C for extension) were carried out by using an Atto Zymoreacter AB-1800 (Tokyo, Japan).

Recombinant DNA techniques. Restriction endonucleases, T4 DNA ligase, and other DNA modifying enzymes were used as specified by the manufacturers (Takara Shuzo and Toyobo). Southern hybridization and colony hybridization were performed according to the standard protocols (Sambrook et al. 1989) using a DIG DNA labeling and detection kit (Boehringer-Mannheim, Germany). DNA sequencing was carried out by the dideoxy chain termination method of Sanger et al. (1977) with an Applied Biosystems model 373A and a Shimadzu model DSQ-500 DNA (Kyoto, Japan) sequencer.

Computer analysis. The DNA and the predicted protein sequences were analyzed with the GENETYX-MAC set of programs (Software Development, Tokyo, Japan).

DNA sequence accession number. The DNA sequence data reported in this paper will appear in the DNA Data Bank of Japan (DDBJ), European Molecular Biology Laboratory (EMBL), and GenBank nucleotide sequence databases under the accession number D87290.

SDS-PAGE. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was done on a 7.5% (w/v) polyacrylamide gel by the method of Laemmli (1970). Proteins were stained by Coomassie brilliant blue R-250 (CBB).

Western blotting analysis. The CSG was purified from the cell envelope fraction of *Ha. japonica* by preparative SDS-PAGE as described previously (Nakamura et al. 1995), and then used for immunization of a rabbit by standard procedures (Ausubel et al. 1995). Cell extracts of *E. coli* carrying the *Ha. japonica* CSG gene were prepared by sonication and applied to SDS-PAGE. Proteins in the gel were transferred to a nitrocellulose filter by the Western blotting method (Burnette 1981). The CSG on the filter was specifically stained using a rabbit antiserum against the *Ha. japonica* CSG and an Immunoblot kit (Bio-Rad, CA, USA).

Results and discussion

Generation of a DNA fragment by PCR

The CSGs from *Hb. halobium* and *Hf. volcanii* were reported to have considerable homology in amino acid sequence (Sumper et al. 1990). PCR primers, directed against the regions conserved between the CSGs from *Hb. halobium* and *Hf. volcanii*, were designed and synthesized. The sense and antisense primer sequences are 5′-ATI

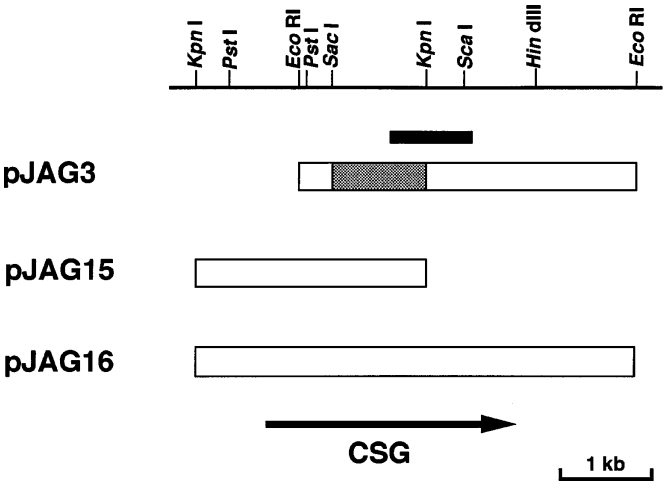

Fig. 1. Physical maps of chromosomal DNA fragments of *Haloarcula japonica*. Plasmid pJAG3 contains the *Eco*RI fragment of 3.5 kb; pJAG15 has the 2.5-kb *Kpn*I insert. *Open bars* show the cloned chromosomal DNA fragments of *Ha. japonica*. *Closed bar* corresponds to the polymerase chain reaction fragment. *Shaded region* indicates the probe for gene walking. *Arrow* is the position of the cell surface glycoprotein (CSG) gene determined from nucleotide sequencing

GA(A/G) ACI CA(A/G) TA(C/T) CTI GA(C/T) GA-3′ (corresponding to the amino acid sequence 329–336 of the mature *Hb. halobium* CSG) and 5′-GT GAC GAT (G/C)AG GTC GTC GCT-3′ (corresponding to the amino acid sequence 634–640 of the mature *Hb. halobium* CSG), respectively. A DNA fragment of about 890 bp was amplified from the chromosomal DNA of *Ha. japonica* by PCR. The PCR-generated DNA fragment was treated with T4 DNA polymerase, T4 polynucleotide kinase, and then ligated into the *Sma*I site of pUC119. Nucleotide sequencing analysis revealed that the deduced amino acid sequence encoded by the PCR fragment exhibited 54.5% and 49.2% identities to those of the *Hb. halobium* and *Hf. volcanii* CSGs, respectively. Therefore, the PCR fragment was used as a highly specific probe for the *Ha. japonica* CSG gene.

Cloning of the CSG gene

Chromosomal DNA of *Ha. japonica* was digested with several restriction enzymes and analyzed by Southern hybridization analysis, using the PCR fragment as a probe. Chromosomal DNA digested with *Eco*RI showed a single hybridization band at about 3.5 kb. DNA fragments of 3–4 kb were isolated from *Eco*RI-digested chromosomal DNA and ligated into the *Eco*RI site of pBluescript II SK+. Transformants were selected and screened by colony hybridization with the specific probe. Several positive clones were obtained and found to contain recombinant plasmids with an identical 3.5-kb *Eco*RI insert. One such plasmid was designated pJAG3 (Fig. 1). Partial nucleotide sequence analysis indicated that the 3.5-kb fragment lacked the 5′ end of the intact CSG gene. Therefore, another genomic library, digested with *Kpn*I, was constructed to obtain the missing region. A *Kpn* I fragment of 2.5 kb was cloned from the library using the *Sac* I–*Kpn* I fragment of pJAG3 as a probe for gene walking. Plasmid pJAG15 carries the 2.5-kb *Kpn* I fragment. The 3.5-kb *Eco*RI and 2.5-kb *Kpn* I fragments overlapped by 1.4 kb, and their organization is shown in Fig. 1.

DNA sequence

The results of the sequencing experiments are summarized in Fig. 2. The sequence contains an open reading frame (ORF) of 2586 bp, coding for a polypeptide of 862 amino acids. The ORF starts with an ATG codon at nucleotide position 250 and ends with a TAA codon at position 2836. The sequence 5′-TTTATA-3′, which resembles the eukaryotic TATA-like "box A" motif [consensus 5'-TTTA (A or T)A-3'] (Reiter et al. 1990), was observed about 150 bp upstream of this ATG.

The codon usage of the CSG gene of *Ha. japonica* was compared with those from the *Hb. halobium* and *Hf. volcanii* CSGs. There is a striking difference in the condon usage in the *Ha. japonica* CSG gene in the codons for aspartic acid, the most abundant amino acid in these CSGs. The proportions of GAU codons in the *Ha. japonica*, *Hb. halobium*, and *Hf. volcanii* CSGs are 43.9%, 11.4% and 2.2%, respectively.

Protein structure

Recently, the N-terminal amino acid sequence of the *Ha. japonica* CSG was determined to be: S-E-R-G-A-G-D-S-Y-T-T-G-P-T-D-G-N-Q (Nakamura et al. 1995). This sequence starts with the 35th amino acid of the ORF shown in Fig. 2, confirming that translation of the ORF indeed generates the primary structure of the *Ha. japonica* CSG. The N-terminally extended sequence of 34 amino acids is thought to be a signal peptide. The amino acid sequence around the potential cleavage site is A-A-A, consistent with the proposed recongnition sequence A-X-A (X: any amino acid) for signal peptidases (Perlman and Halvorson 1983) that has also been found in both the *Hb. halobium* (Lechner and Sumper 1987) and *Hf. volcanii* CSGs (Sumper et al. 1990).

The mature polypeptide contains 828 amino acids, with a calculated molecular mass of 87 166 Da. This value is much lower than that estimated from SDS-PAGE [170 kDa (Nakamura et al. 1992); see also Fig. 5a, lane 1], in spite of its high carbohydrate content (12% w/w; Nakamura et al. 1995). The amino acid composition of the mature CSG of *Ha. japonica* revealed that the CSG contained a large proportion of acidic amino acids (D, 14.7%; E, 10.4%). The *Ha. japonica* CSG may have a reduced capacity for SDS binding due to its unusual content of acidic residues. As a consequence, the reduced electrophoretic mobility would cause an overestimate of the molecular mass. Similar phenomena have been observed with the other halobacterial CSGs (Lechner and Sumper 1987; Sumper et al. 1990).

Hydropathy analysis (Kyte and Doolittle 1982) revealed that the mature CSG of *Ha. japonica* showed a single, highly hydrophobic stretch at the C-terminal end (Fig. 3). The other regions of the polypeptide chain mainly consisted of polar amino acids and showed a dominant negative charge. The C-terminal hydrophobic region probably serves as a membrane anchor. In the case of the *Hb. halobium* CSG, such a structural feature has been described by Lechner and Sumper (1987), and the *Ha. japonica* CSG may share this feature. Figure 4 shows the results of dot matrix homology plots. The amino acid sequence of the *Ha. japonica* CSG shows 52.1% and 43.2% identities to those from *Hb. halobium* and *Hf. volcanii*, respectively. The degree of homology drops somewhat towards the N-terminus, indicating the possibility of different architectures of the extracellular portions of the these CSGs. Five potential N-glycosylation sites [N-X-(S or T); X ≠ D, P] are found in the mature CSG of *Ha. japonica* (see Fig. 2). Fewer N-glycosylation sites are recognized in the *Ha. japonica* CSG than in the *Hb. halobium* and *Hf. volcanii* CSGs [12 and 7 glycosylation sites, respectively (Lechner and Sumper 1987; Sumper et al. 1990)]. Furthermore, the localization of the potential glycosylation sites in the *Ha. japonica* CSG – biased to the C-terminus – is quite different from that in the other halobacterial CSGs.

Expression of the CSG gene in *E. coli*

The 4.6-kb *Kpn*I–*Eco*RI Fragment, obtained by recombining the 3.5-kb *Eoc* RI and 2.5-kb *Kpn*I fragments at a com-

Fig. 2. Nucleotide sequence of the CSG gene from *Ha. japonica* and the deduced amino acid sequence. The nucleotide sequence was determined in both directions. Numbers on the *right* denote nucleotide positions. The putative box A sequence is *underlined*. The N-terminal amino acid sequence determined for the mature CSG from *Ha. japonica* is *double-underlined*. The *vertical arrowhead* indicates the cleavage site of signal peptidase. Five potential N glycosylation sites are *boxed*

```
CGTTCTTAACAGGTCTGGCTGATTCATTCACTAGTAGTTTGGATCACTTTCTTGGGAACT     60
GACGTATGGCAGACTCGTATTAGCTTGGAACACAAAGTATTTATAAACACCGGCGTTTGT    120
                                    boxA
TTTGGCTGTACCCCTACCATGCTATCAGCAGTGAGTAAATTAGGCGAAAAGTGTGTTGCA    180
CACGCCGATGATGTGCTGCACAACTATACGTAAATCAAAACTGCTGGTGGCGATTGCAAC    240
AATCAAACTATGACAGATACACAACAAAAAATCAAAGCGGTGCTCCTTACGGTGCTTATG    300
         M  T  D  T  Q  Q  K  I  K  A  V  L  L  T  V  L  M
GTGACTTCTGTATTCGCAGCCACCATAGCCTTCTCAGGAGCTGCCGCCGCCTCAGAACGA    360
V  T  S  V  F  A  A  T  I  A  F  S  G  A  A  A  A  S  E  R
                                                   ▲
GGCGCTGGCGATAGTTATACAACGGGACCGACTGACGGTAATCAAGACAATGTTGACAGC    420
G  A  G  D  S  Y  T  T  G  P  T  D  G  N  Q  D  N  V  D  S
GCAGGTAATGTTGGCGCTGGTGCAGTCGTCTTCCAAGGAGAAGACGATATTGAGGGCGAA    480
A  G  N  V  G  A  G  A  V  V  F  Q  G  E  D  D  I  E  G  E
TTCGCTGACGGAAACGGAGACACCGTTGGCATCGGTGAACTACAGAAGGTGTCCGGTGAC    540
F  A  D  G  N  G  D  T  V  G  I  G  E  L  Q  K  V  S  G  D
AACGAGGGCATCCTACTGGAGTCACCGATCCCACAGGACCAGCCAACTGGTCGATATACA    600
N  E  G  I  L  L  E  S  P  I  P  Q  D  Q  P  T  G  R  Y  T
GCAAATCCAGGCGTCGAAGGCACTGAAGGTGTGACCCTCCAAACGCCGCGCATCACTGAT    660
A  N  P  G  V  E  G  T  E  G  V  T  L  Q  T  P  R  I  T  D
CTCGAAATACAGAACAGTGACGAAGGCGACGTGACTGGGAGCATTCTGCAGGCCAACAAT    720
L  E  I  Q  N  S  D  E  G  D  V  T  G  S  I  L  Q  A  N  N
GATAACACTGCAGAGATCCTTGTCGATTACAACTACGATGAGGCCGAGGATCTTGAGCTC    780
D  N  T  A  E  I  L  V  D  Y  N  Y  D  E  A  E  D  L  E  L
ACCGTCGAGGACGAGGACGGTCTTGAAGTCACGGAGGAAATCCTCGCCGACGGTGCCTCG    840
T  V  E  D  E  D  G  L  E  V  T  E  E  I  L  A  D  G  A  S
GAAACGGTCAATACCAATGTCAACAATGATGATCATCCCAACCCCGCCGCTGATGGGGAC    900
E  T  V  N  T  N  V  N  N  D  D  H  P  N  P  A  A  D  G  D
CGTGATGATTCCTTCGATGCCGGGTTTACCATCAACCCGTCCAATGTTGACGAAGGCGAG    960
R  D  D  S  F  D  A  G  F  T  I  N  P  S  N  V  D  E  G  E
TACACGATTACCGTCGAAGGTGTCGAAGATCTCGACTTCGGTGATGCCTCGGAAACCGCA   1020
Y  T  I  T  V  E  G  V  E  D  L  D  F  G  D  A  S  E  T  A
ACAGTTGAGATCACGACCGACCAGACAGCTTCGCTCAGCCTCGACAGCGACGAGGTTACT   1080
T  V  E  I  T  T  D  Q  T  A  S  L  S  L  D  S  D  E  V  T
CAGGGTGAAGACCTTGGATTCGATATCGAAAACAGCCCAGAAGGGAACTTCCATGCGGTT   1140
Q  G  E  D  L  G  F  D  I  E  N  S  P  E  G  N  F  H  A  V
GTTATCGAAGAGAGTGAGTTCCGCGACAGCGCCTCCGCCAGCAATTACGCGAAGGTCTTC   1200
V  I  E  E  S  E  F  R  D  S  A  S  A  S  N  Y  A  K  V  F
CGGAACGTCGGTGACACTTCGGATCGCGGTCTAGTCGGTGAAGACGCCGATGGTAACACT   1260
R  N  V  G  D  T  S  D  R  G  L  V  G  E  D  A  D  G  N  T
GTGGCGGTCGCTCCGGGCGATGCTGATTCCCTAGAGAGCATTGACTACGCTTACGGCATC   1320
V  A  V  A  P  G  D  A  D  S  L  E  S  I  D  Y  A  Y  G  I
GTCGAGATTGATGGCGGTACAGGTGTTGGCTCGGCCGAAACTCAGTACCTCGACGACAGC   1380
V  E  I  D  G  G  T  G  V  G  S  A  E  T  Q  Y  L  D  D  S
TCGATTGATATCGATCTGTACGAGGCTGCTAACGGTGACTACACGGACAACAACGCACAC   1440
S  I  D  I  D  L  Y  E  A  A  N  G  D  Y  T  D  N  N  A  H
GTTAACGATATCAACCTAGTCACCGACGACACTTACGAAACGGATGACGAACAGGACTTC   1500
V  N  D  I  N  L  V  T  D  D  T  Y  E  T  D  D  E  Q  D  F
```

mon *Sac* I site, was inserted in the pUC119 vector to construct pJAG16 (see Fig. 1). The plasmid pJAG16, containing the full-length CSG gene of *Ha. japonica*, was introduced into *E. coli* strain MV1184. Cell extracts of the transformant were analyzed by SDS-PAGE. A band was detected at about 170 kDa by CBB staining; this band was not found in pUC119-transformed *E. coli* cells (Fig. 5a, lanes 1 and 2). Immunological detection of the *Ha. japonica* CSG in the call extract using a specific antiserum showed a positive band at about 170 kDa (Fig. 5b, lane 2). A vague band at 90 kDa was also observed in Fig. 5b, lanes 1 and 2, suggesting the existence of a host protein that has an affinity for the anti-CSG antiserum. These results seemed to show that the *Ha. japonica* CSG was produced in *E. coli*. In pJAG16, the CSG gene was cloned in the reverse orientation to that of the *lac* promoter of pUC119. This suggests that the internal promoter of *Ha. japonica* CSG gene can function in *E. coli*.

Fig. 2. *Continued*

```
GACGTTCTTGAGGGTGACCTCACCATCGACAGTCCGTCTGGGACGTACGTAACTGGGTCC  1560
D  V  L  E  G  D  L  T  I  D  S  P  S  G  T  Y  V  T  G  S

GAGGTGGATGTGAACGGGACCGCAAGCGAGGGCATTGATGATGTCGCAATCTACGCTCGC  1620
E  V  D  V  N  G  T  A  S  E  G  I  D  D  V  A  I  Y  A  R

GACAACAATGATTACGAACTCGTTGAAATCGATAGTGAAGAAACCATCAGCGTTGATGGT  1680
D  N  N  D  Y  E  L  V  E  I  D  S  E  E  T  I  S  V  D  G

GACGATACCTTCAGCGAGGAAGATATCAGCCTCAGTGGCGGTGACCTCGGTGGCAATGAC  1740
D  D  T  F  S  E  E  D  I  S  L  S  G  G  D  L  G  G  N  D

ATCCTCGGCCTTCCGGGTACCTACCGGATCGGTGTCGTTGATGTCGAGGATGCCGACTCG  1800
I  L  G  L  P  G  T  Y  R  I  G  V  V  D  V  E  D  A  D  S

AACTCCAACGGAACTGTCGATGACTCGCTTACGACGTCTGAGTTCAACAGCGGTGTCAGT  1860
N  S  N  G  T  V  D  D  S  L  T  T  S  E  F  N  S  G  V  S

TCCGCTGAGTCACTTCGAGTGACCGACACGGAGCTTAACGGTACGTTCATCACCTACAAC  1920
S  A  E  S  L  R  V  T  D  T  E  L  N  G  T  F  I  T  Y  N

GGCCAGATATCGAGTGACGATAATCAGATCGATGTCGAAGGTCAAGCGCCTGGCAAGGAC  1980
G  Q  I  S  S  D  D  N  Q  I  D  V  E  G  Q  A  P  G  K  D

AACCTTGTCATAGCATTCGTTGACTCGCGTGGCAACGCTGTTGCTACAGATATCTCCGTT  2040
N  L  V  I  A  F  V  D  S  R  G  N  A  V  A  T  D  I  S  V

GATGATGATGACACGTTCAGTGAAGACGACATCAGCATCTCAGCACTCAGCGAAGGGACT  2100
D  D  D  D  T  F  S  E  D  D  I  S  I  S  A  L  S  E  G  T

GTCACAGCACACATTATCTCCTCTGGCCGCGATAACCTGTTCGGAGATGGTGTCTCTGAT  2160
V  T  A  H  I  I  S  S  G  R  D  N  L  F  G  D  G  V  S  D

AGCAGTTCCGGGTTCGCAAGCCTAATCGAGGAAGAATACGCGAGCGGCAGCAGTACTGGT  2220
S  S  S  G  F  A  S  L  I  E  E  E  Y  A  S  G  S  S  T  G

GACCAGGTCCGATCGCGAATCCTGGAAAACAGTGTCGACGACACTGCCAGCGATGACCTC  2280
D  Q  V  R  S  R  I  L  E  N  S  V  D  D  T  A  S  D  D  L

ATCGTGAATGAACAGTTCCGTCTGGCAGATGGGCTAACCATCGTCGAATCGGTTAACAGC  2340
I  V  N  E  Q  F  R  L  A  D  G  L  T  I  V  E  S  V  N  S

CCAGTTGAAGCCAATGGAACGATTGAAATTGAGGGTACTACCAACCGTAAGCCGGACGAT  2400
P  V  E  A  N  G  T  I  E  I  E  G  T  T  N  R  K  P  D  D

AACACCATTACGGTTGAACTCCTCGACGACGAGGATGAATCAGTTACTGTCGACAGCACA  2460
N  T  I  T  V  E  L  L  D  D  E  D  E  S  V  T  V  D  S  T

GACGAGTGGAGCAGTGACGGTCAGTGGTCTGTCAGCCTCGATCTCTCGGATGAGAATGTC  2520
D  E  W  S  S  D  G  Q  W  S  V  S  L  D  L  S  D  E  N  V

GAGCCTGGTAACTTCACCGTTGAAGCTGACGACGGTGACAACACTGATCGCCAGAGTGTC  2580
E  P  G  N  F  T  V  E  A  D  D  G  D  N  T  D  R  Q  S  V

CAGATCGTTGAAGCCGGCTCGCTGGAGGAAGAGCAGCCGGCCACGGACACGCCGGAGCCG  2640
Q  I  V  E  A  G  S  L  E  E  E  Q  P  A  T  D  T  P  E  P

GACACGGACACGCCGGAGCCGGCCACGGACACGCCGGAGCCGGCCACGGACACGCCGGAG  2700
D  T  D  T  P  E  P  A  T  D  T  P  E  P  A  T  D  T  P  E

CCGGACACGGACACGCCGGAGCCGGATACGGAGACTGAAGAAGCTACGACCGAAGCAACT  2760
P  D  T  D  T  P  E  P  D  T  E  T  E  E  A  T  T  E  A  T

GGTCCTGGCTTCACAGCAGCTATCGCGCTCATCGCGCTCGTTGCTGCTGCACTCCTCGCC  2820
G  P  G  F  T  A  A  I  A  L  I  A  L  V  A  A  A  L  L  A

GTCCGCCGCGACAACTAACTGGATTCACTCCAGTTTTTGACGCTCCATTTTTTGTAGCCC  2880
V  R  R  D  N  *

TCGAAGCTTGGAACAACTTGTAGATAGTCACTGCCAAGTGATTAGTCGTCAAGCAATTCG  2940
```

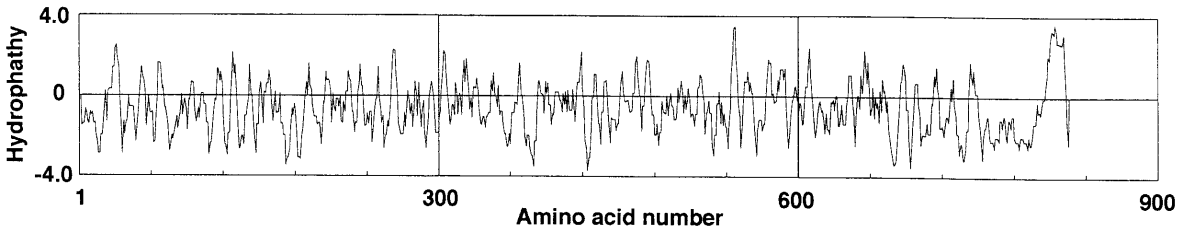

Fig. 3. Hydropathy analysis of the mature CSG of *Ha. japonica*. The curve plots the hydrophobicity index for each residue as the average hydrophobicity index for the window of 5 amino acids centered around that residue

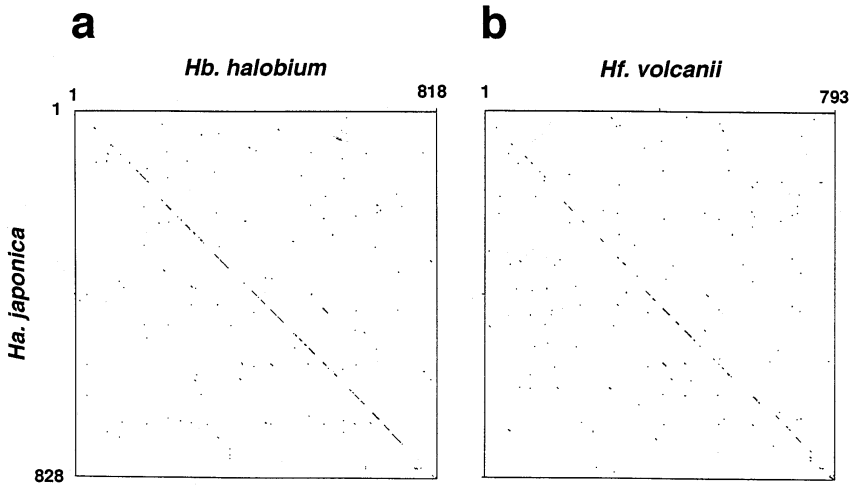

Fig. 4. Amino acid sequence homologies between the *Halobacterium halobium* and *Ha. japonica* CSGs (**a**), and between the *Haloferax volcanii* and *Ha. japonica* CSGs (**b**). The mature amino acid sequence of the *Ha. japonica* CSG is represented on the *vertical axis*, and the *horizontal axes* show the other CSGs. The dot matrix plots were prepared using the GENETYX-MAC program. Segments of 5 amino acid residues from the horizontal axis were compared with segments from the vertical axis. A dot was placed whenever the number of matching residues was ≥4

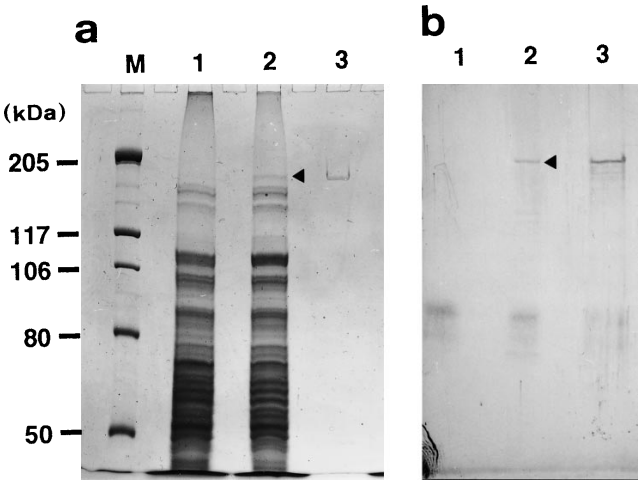

Fig. 5. Sodium dodecyl sulfate (SDS) polyacrylamide gel electraphoresis of the *Ha. japonica* CSG and *Escherichia coli* cell extracts. An SDS–7.5%-polyacrylamide gel was used. After electrophoresis, proteins in the gel were stained with Coomassie brilliant blue R-250 (**a**) or applied to the Western blotting analysis (**b**). *Lanes 1*, the cell extract of *E. coli* carrying pUC119; *lanes 2*, the cell extract of *E. coli* carrying pJAG16; *lanes 3*, the purified CSG of *Ha. japonica*. *M* indicates the molecular mass markers. The *arrowhead* shows the possible CSG produced by *E. coli*

Acknowledgments We thank Dr. Chiaki Kato (Japan Marine Science and Technology Center) for nucleotide sequencing and Dr. Toshiaki Kudo (Institute of Physical and Chemical Research) for the preparation of anti-CSG antiserum. This work was partially supported by a grant from the Japan Securities Scholarship Foundation to S. N.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1995) Short protocols in molecular biology (3rd edn). Wiley, New York

Blaurock AE, Stoeckenius W, Oesterhelt D, Scherphof GL (1976) Structure of the cell envelope of *Halobacterium halobium*. J Cell Biol 71:1–22

Burnette WW (1981) "Western blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate – polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal Biochem 112:195–203

Hamamoto T, Takashina T, Grant WD, Horikoshi K (1988) Asymmetric cell division of a triangular halophilic archaebacterium. FEMS Microbiol Lett 56:221–224

Horikoshi K, Aono R, Nakamura S (1993) The triangular halophilic archaebacterium *Haloarcula japonica* strain TR-1. Experientia 49:497–502

Kessel K, Wildhaber I, Cohen S, Baumeister W (1988) Three-dimensional structure of the regular surface glycoprotein layer of *Halobacterium volcanii* from the Dead Sea. EMBO J 7:1549–1554

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680–685

Lechner J, Sumper M (1987) The primary structure of a procaryotic glycoprotein. J Biol Chem 262:9724–9729

Mengele R, Sumper M (1992) Drastic differences in glycosylation of related S-layer glycoproteins from moderate and extreme halophiles. J Biol Chem 267:8182–8185

Mescher MF, Strominger JL (1976) Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarium*. Proc Natl Acad Sci USA 73:2687–2691

Nakamura S, Aono R, Mizutani S, Takashina T, Grant WD, Horikoshi K (1992) The cell surface glycoprotein of *Haloarcula japonica* TR-1. Biosci Biotechnol Biochem 56:996–998

Nakamura S, Mizutani S, Wakai H, Kawasaki H, Aono R, Horikoshi K (1995) Purification and partial characterization of cell surface glycoprotein from extremely halophilic archaeon *Haloarcula japonica* strain TR-1. Biotechnol Lett 17:705–706

Nishiyama Y, Takashina T, Grant WD, Horikoshi K (1992) Ultrastructure of the cell wall of the triangular halophilic archaebacterium *Haloarcula japonica* strain TR-1. FEMS Microbiol Lett 99:43–48

Nishiyama Y, Nakamura S, Aono R, Horikoshi K (1995) Electron microscopy of halobacteria. In: DasSarma S, Fleischman EM (eds) Archaea, a laboratory manual, halophiles. Cold Spring Harbor Laboratory Press, New York

Perlman D, Halvorson HO (1983) A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J Mol Biol 167:391–409

Reiter W-D, Hüdepohl U, Zillig W (1990) Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. Proc Natl Acad Sci USA 87:9509–9513

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, New York

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Sleytr UB, Messner P (1992) Crystalline bacterial cell-surface layers. Adv Microb Physiol 33:213–275

Sumper M (1987) Halobacterial glycoprotein biosynthesis. Biochim Biophys Acta 906:69–79

Sumper M, Berg E, Mengele R, Strobel I (1990) Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*. J Bacteriol 172:7111–7118

Takashina T, Hamamoto T, Otozai K, Grant WD, Horikoshi K (1990) *Haloarcula japonica* sp. nov., a new triangular halophilic archaebacterium. Syst Appl Microbiol 13:177–181