ELSEVIER

Biochimica et Biophysica Acta 1281 (1996) 1–4

BBA Biochimica et Biophysica Acta

Short sequence-paper

Sequence of a Na^+/glucose symporter gene and its flanking regions of *Vibrio parahaemolyticus* [1]

Rafiquel I. Sarker [a], Yoshie Okabe [a], Masaaki Tsuda [a], Tomofusa Tsuchiya [a,b,*]

[a] *Department of Microbiology, Faculty of Pharmaceutical Sciences, Okayama University, Tsushima, Okayama 700, Japan*
[b] *Gene Research Center, Okayama University, Tsushima, Okayama 700, Japan*

Received 20 November 1995; revised 31 January 1996; accepted 5 February 1996

Abstract

The nucleotide sequence of an approximately 6 kbp segment of chromosomal DNA of *Vibrio parahaemolyticus* was determined. The nucleotide sequence revealed four open reading frames (ORFs) in this region. Hydropathy profiles of the deduced amino acid sequence of the ORFs indicate that ORF1 encodes a hydrophobic polypeptide with typical characteristics of a membrane transport protein. All other ORFs encode hydrophilic polypeptides. ORF1 showed significant amino acid sequence similarity to proteins of the SGLT (Na^+/glucose symporter) family, and the amino acid sequence of ORF4 showed very high similarity to several bacterial transcriptional repressor proteins (GalR-LacI family). We observed elevated glucose transport activity in cells harboring a plasmid carrying the DNA region corresponding to ORF1, and the glucose transport was greatly stimulated by Na^+. Thus, we believe that ORF1 encodes a Na^+/glucose symporter.

Keywords: Sodium ion/glucose symporter; Glucose transport; Sequence analysis; Marine bacterium; (*V. parahaemolyticus*)

Vibrio parahaemolyticus is a marine bacterium that requires Na^+ for its growth [1]. This organism shows extremely fast growth under favorable conditions. This property in addition to its diarrheic toxin is closely related to the fact that this bacterium causes food poisoning at high frequency. Cells of *V. parahaemolyticus* can grow more than twice as fast as cells of *Escherichia coli*, which is the best characterized bacterium. It seemed to us that this organism possessed very efficient or very unique energy transducing systems to support such fast growth. Thus, we have been interested in membrane-related energy transducing processes in *V. parahaemolyticus* cells. We found that Na^+ was involved in many energy transducing processes in membranes of this organism [2–5]. So far, we have found a respiratory Na^+ pump [2], a Na^+-stimulated ATP synthesizing system [3], a Na^+/H^+/adenosine symport system [4] and a Na^+/glucose symport system [5] in this organism. During the course of our studies on the membrane transport systems, we cloned DNA fragments from the *V. parahaemolyticus* chromosome and sequenced them. We found one ORF (open reading frame) which specifies the glucose transport protein.

V. parahaemolyticus AQ3334 [2] was used as the source of chromosomal DNA for cloning. Cells of AQ3334 were grown in Luria–Bertani (LB) medium at 37°C under aerobic conditions. *E. coli* strain Sϕ1660 (*ΔnupC*, *ΔnupG*) [6], which lacks adenosine transport activity and cannot grow on adenosine as a sole source of carbon, and JM83 were used as hosts for cloning and subcloning, respectively. *E. coli* cells were grown in a minimal medium [7] supplemented with 10 mM adenosine and solidified with 1.5% agar or in L broth [8] in the presence of ampicillin (50 μg/ml) at 37°C under aerobic conditions. Plasmid pBR322 was used for the construction of various hybrid plasmids. Plasmid pBluescript IIKS(−) was used for subcloning and DNA sequencing. The nucleotide sequence was determined by the dideoxy chain termination method [9]. Double stranded plasmid DNA was used for the sequencing reaction as specified by Applied Biosystems for use in the 373A Automated Sequencer. The DNA sequences of both strands were determined. Sequences were analyzed with the sequence analysis program, GENETYX.

* Corresponding author (at address a). Fax: +81 86 2517957; e-mail: tsuchiya@ph2ews1.okayama-u.ac.jp.

[1] The nucleotide sequence data reported in this paper have been submitted to the DDBJ, EMBL and Genbank nucleotide sequence databases under the accession number D78137.

PII S0005-2736(96)00025-9

E. coli PPA172 [10], which lacks a major glucose transport system (phospho*enol*pyruvate:carbohydrate phosphotransferase system), was used to evaluate the effect of the cloned DNA region on glucose transport activity in the cells. Transport assays [5] and protein assays [11] were performed as reported.

A 6.4 kbp DNA fragment of *V. parahaemolyticus* AQ3334 was cloned by the following procedure. Chromosomal DNA of *V. parahaemolyticus* AQ3334 was partially digested with *Sau*3AI and fractionated by sucrose density gradient centrifugation, and DNA fragments of 4 to 10 kbp were ligated to the *Bam*HI site of plasmid pBR322 that was previously treated with bacterial alkaline phosphatase. Competent cells of *E. coli* Sϕ1660 were transformed with

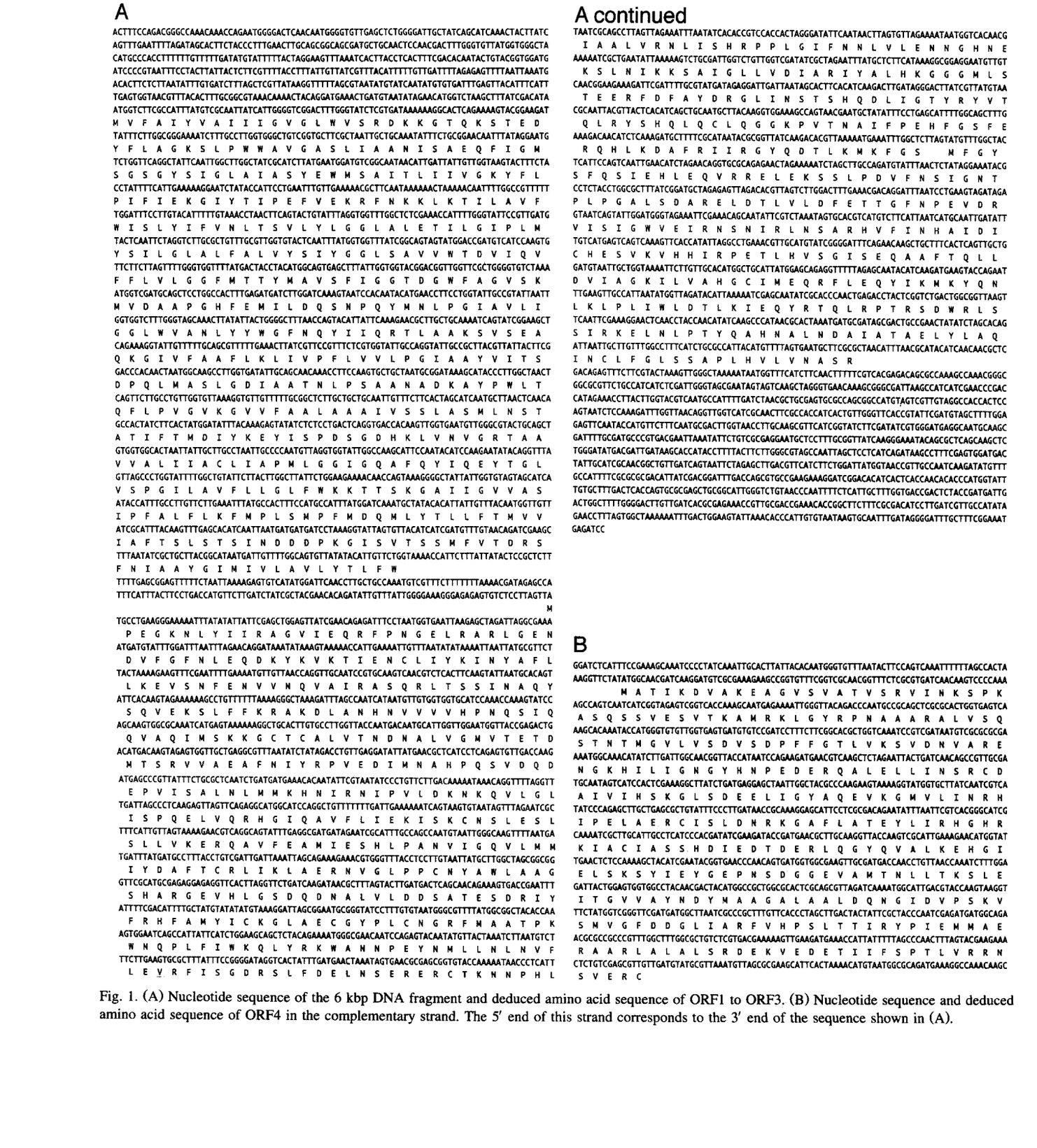

Fig. 1. (A) Nucleotide sequence of the 6 kbp DNA fragment and deduced amino acid sequence of ORF1 to ORF3. (B) Nucleotide sequence and deduced amino acid sequence of ORF4 in the complementary strand. The 5′ end of this strand corresponds to the 3′ end of the sequence shown in (A).

the hybrid plasmids, spread on agar plates containing adenosine as the sole source of carbon. One hybrid plasmid could complement the growth of Sϕ1660 cells on adenosine. The recombinant plasmid was designated as pYAT26. An approximately 2.2 kbp *Dra*I-*Dra*I fragment containing ORF1 (see below) was ligated at the *Eco*RV site of pBR322. The constructed plasmid was named pYAT271A.

The nucleotide sequence of 5857 bp in the pYAT26 was determined (Fig. 1A). Computer analysis of the sequence data predicted four potential ORFs (ORF1 to ORF4). Location and directions of the ORFs in the pYAT26 are shown in Fig. 2. The ORFs are not overlapped and the direction of ORF4 is opposite with respect to the other ORFs.

ORF1 is 1590 nucleotides long with a possible ATG start codon at position 541 of the sequenced DNA and ending with a TAA stop codon at position 2130 (Fig. 1A). The termination codon TAA is followed by a rho-independent transcriptional terminator-like sequence at position 2150–2172. ORF1 encodes a peptide of 530 amino acid residues with a predicted molecular weight of 57.4 kDa. The most abundant amino acid residue of this protein was Leu (61 residues) followed by Ala (57 residues) and Ile (54 residues). The least abundant amino acids were Cys (1 residue) and His (2 residues). The protein specified by ORF1 is highly nonpolar (71% nonpolar residues). Thus, this protein is highly hydrophobic.

ORF2 is 1695 nucleotides long, beginning with an ATG start codon at position 2340 of the sequenced DNA fragment and ending with a TAA at position 4034 (Fig. 1A). ORF2 encodes a polypeptide of 565 amino acid residues with a predicted molecular weight of 64.2 kDa. ORF2 contains 54% nonpolar and 46% polar amino acid residues.

ORF3 is 702 nucleotides long, beginning with a possible ATG start codon at position 4039 and ending with a TAA stop codon at position 4740 (Fig. 1A). ORF3 could encode a peptide of 234 amino acid residues with a predicted molecular weight of 26.7 kDa. ORF3 contains 52% nonpolar residues.

ORF4 is 996 nucleotides long beginning with an ATG start codon at position 101 of reverse complementary sequences and ending with a TGA stop codon at position

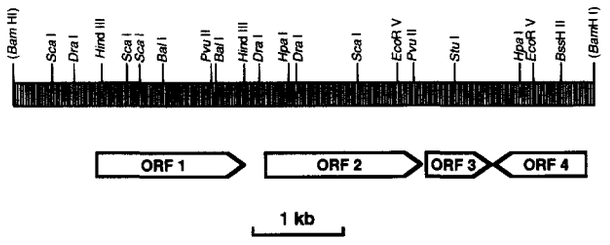

Fig. 2. Restriction map of the 6 kbp DNA fragment derived from the *V. parahaemolyticus* chromosome and carried in the plasmid pYAT26, and location and directions of ORFs found in this region. A plasmid pYAT271A carries *Dra*I-*Dra*I fragment containing ORF1.

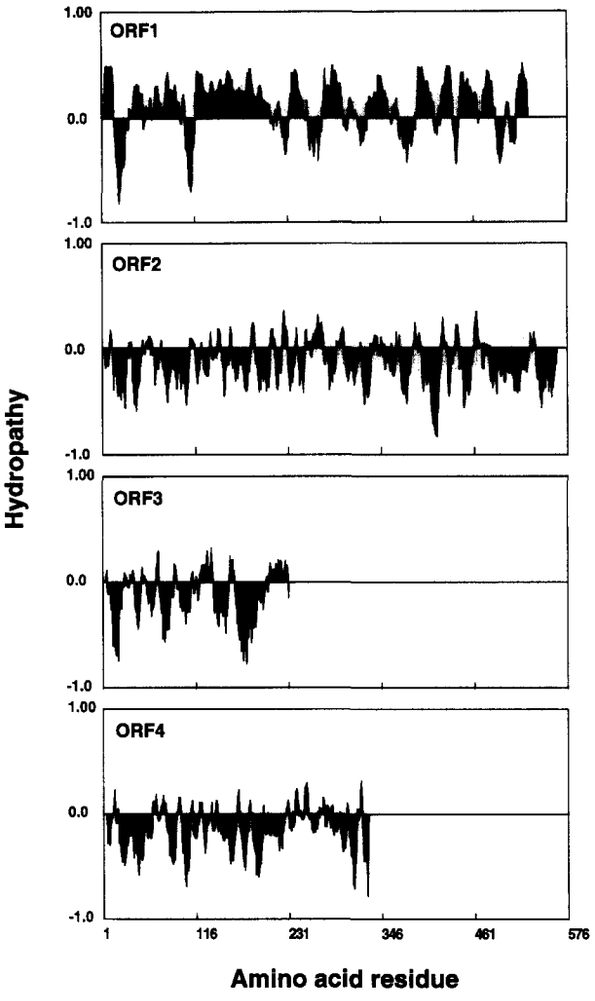

Fig. 3. Hydropathy profiles of ORF1 to ORF4. The hydrophobicity and hydrophilicity of the four ORFs were calculated by the method of Eisenberg et al. [13] with a window of eight residues, and the profiles were plotted from the amino terminus (left) to the carboxyl terminus (right). The portions above the zero lines indicate hydrophobic regions.

1096 of reverse complementary sequence (Fig. 1B). ORF4 encodes a polypeptide of 332 amino acid residues with a predicted molecular weight of 36.2 kDa. The protein derived from ORF4 contains 53% nonpolar residues.

All of the four ORFs are preceded by possible Shine–Dalgarno sequences [12]. Promoter-like sequences are present upstream from ORF1 and ORF2.

Hydropathy analysis of the deduced amino acid sequence of the ORFs was carried out by the method of Eisenberg [13], and the profiles are shown in Fig. 3. The patterns suggest that the protein encoded by ORF1 is very hydrophobic and the proteins encoded by ORF2, ORF3 and ORF4 are hydrophilic. The hydropathy pattern of ORF1 indicates that it is a typical integral membrane protein with 10 to 14 transmembrane regions.

The deduced amino acid sequences of the four ORFs were searched for homology to known protein sequences in the SWISSPROT protein sequence database. Among the four proteins potentially encoded by the clone, only ORF1

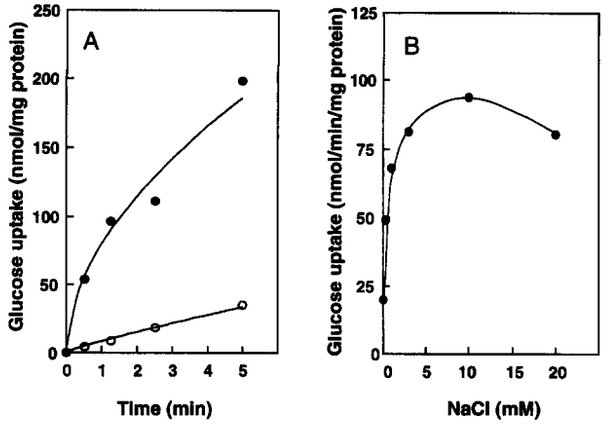

Fig. 4. Glucose transport in cells harboring a plasmid carrying ORF1. (A) Glucose transport in cells of *E. coli* PPA172 (○) and PPA172/pYAT271A (●) was measured in the presence of 10 mM NaCl as described previously [5]. (B) Initial velocity of glucose transport in cells of *E. coli* PPA172/pYAT271A was measured in the presence of indicated concentrations of NaCl.

and ORF4 showed significant homology to known protein sequences. The protein sequence comparisons revealed significant similarity of ORF1 to members of the SGLT (Na^+/glucose symporter) family of mammalian and bacterial origin. The deduced amino acid sequence of ORF1 showed 31% identity and 75% similarity with the Na^+/glucose symporter (SGLT1) of the human intestine [14]. About 66% and 63% similarity was found with the Na^+/proline [15] and Na^+/pantothenate [16] symporters of *E. coli*, respectively. On the other and, the protein encoded by ORF4 showed very high similarity to the GalR-LacI family, the transcriptional regulatory proteins in bacteria. The protein encoded by ORF4 showed 56% identity and 88% similarity with the galactose operon repressor, GalR, of *E. coli* [17]. Other members of the GalR-LacI family showed 68% to 86% similarities with the protein specified by the ORF4.

From the homology data it seemed that the protein encoded by ORF1 is a Na^+/solute symporter, probably a Na^+/glucose symporter which was reported previously [5]. Thus, we investigated whether the protein specified by ORF1 can transport glucose. Cells of *E. coli* PPA172 whose glucose transport system is defective were transformed with plasmid pYAT271A which carries ORF1, and transport of glucose was measured. Cells of PPA172/pYAT271A showed very high activity of glucose transport (Fig. 4A). On the other hand, the host cells (PPA172) showed very low activity. Furthermore, the glucose transport activity observed in cells of PPA172/pYAT271A was greatly stimulated by Na^+ with maximum stimulation at 10 mM (Fig. 4B). These results support the idea that ORF1 specifies a Na^+/glucose symporter.

We observed no significant increase in adenosine transport activity in cells harboring the plasmid pYAT271A which carries ORF1. Further studies should clarify why plasmid pYAT26 which carries the four ORFs complemented the growth defect of *E. coli* Sϕ1660 on adenosine.

References

[1] Baumann, P. and Schubert, H.W. (1984) Bergey's Manual of Systematic Bacteriology (Vol. 1), pp. 516–550, Williams and Wilkins, Baltimore.
[2] Tsuchiya, T. and Shinoda, S. (1985) J. Bacteriol. 162, 794–798.
[3] Sakai, Y., Moritani, C., Tsuda, M. and Tsuchiya, T. (1989) Biochim. Biophys. Acta 973, 450–456.
[4] Okabe, Y., Sakai-Tomita, Y., Mitani, Y., Tsuda, M. and Tsuchiya. T. (1991) Biochim. Biophys. Acta 1059, 332–338.
[5] Sarker, R.I., Ogawa, W., Tsuda, M., Tanaka, S. and Tsuchiya, T. (1994) J. Bacteriol. 176, 7378–7382.
[6] Hansen, S.E.W., Jensen, N. and Munch-Petersen, A. (1987) Eur. J. Biochem. 168, 385–393.
[7] Tanaka, S., Lerner, S.A. and Lin, E.C.C. (1967) J. Bacteriol. 93, 642–648.
[8] Lennox, E.S. (1955) Virology 1, 190–206.
[9] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[10] Kuroda, M., De Waard, S., Mizushima, K., Tsuda, M., Postma, P. and Tsuchiya, T. (1992) J. Biol. Chem. 267, 18336–18341.
[11] Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) J. Biol. Chem. 193, 265–275.
[12] Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342–1346.
[13] Eisenberg, D., Schwarz, E., Komaromy, M. and Wall, R. (1984) J. Mol. Biol. 179, 125–142.
[14] Hediger, M.A., Turk, E. and Wright, E.M. (1989) Proc. Natl. Acad. Sci. USA 86, 5748–5752.
[15] Nakao, T., Yamato, I. and Anraku, Y. (1987) Mol. Gen. Genet. 208, 70–75.
[16] Jackowski, S. and Alix. J.-H. (1990) J. Bacteriol. 172, 3842–3848.
[17] Von Wilcken-Bergmann, B. and Muller-Hill, B. (1982) Proc. Natl. Acad. Sci. USA 79, 2427–2431.