Journal of Microbiology and Biotechnology
Vol. 5, No. 3, 125-131, 1995

Nucleotide Sequence on Upstream of the *cdd* Locus in *Bacillus subtilis*

KIM, KYE-WON, SEON-KAP HWANG, JOO-WON SUH[1], BANG-HO SONG[2] SOON-DUCK HONG AND JONG-GUK KIM*

Department of Microbiology, [2]Department of Biology, Kyungpook Nat'l University, Taegu, 702-701, [1]Department of Biology, Myung-Ji University, Anyang, 449-728, Korea

A 3,346 bp of the *cdd* upstream region in *Bacillus subtilis* was sequenced from the pSO1 (Song BH and J Neuhard. 1989. *Mol. Gen. Genet.* **216**: 462-468) and sequence homology was searched to the known genes in Genbank and European Molecular Biology Laboratory databanks. Five complete and one truncated putative coding sequences deduced from the nucleotide sequence were found through the ORF searching by Genetyx and Macvector software, and one of them was identified as the *dgk* (diacylglycerol kinase) gene and another, a truncated one, as the *pho*H (phosphate starvation-inducible gene) gene. The *B. subtilis dgk* gene, having a role for response to several environmental stress signals, revealed an open reading frame of 134 amino acids with 43.1% of sequence identity to the *Streptococcus mutans dgk* gene. The carboxy terminal 59 residues of the truncated *phoH* gene showed 52.7% and 34.5% of sequence identity in amino acids with the corresponding genes of *Mycobacterium leprae* and *Escherichia coli*. The four remaining coding sequences consisting of 115, 421, 91, and 91 residues were thought to be unknown ORFs because they have no significant similarity to known genes.

Bacillus subtilis genome analysis was started by European (16) and Japanese groups (24) about five years ago. Compared to the *E. coli* genome which has already been arranged by the ordered clones (14) and sequenced for about half of the whole genome (5), *B. subtilis* genome analysis proved much more difficult because some parts of the *Bacillus* gene fragments could not be recovered by cloning in the *E. coli* host (8). With the eukaryotic genome analysis of *Saccharomyces cerevisiae* (25), *Schizosaccharomyces pombe* (23), *Caenorhabditis elegans* (32), Drosophila (30), mouse (15), rice (22), *Arabidopsis thaliana* (26), and human (33), systematic genome analysis becomes a new frontier for developing useful and disease inducing/supressing genes and for understanding gene density, distribution, and organization in biological terms.

In order to improve and resolve any confusion over the *Bacillus* genome, genome analysis was carried out to the upstream boundary region of the *cdd* gene located at 223° on the genetic map. According to the *B. subtilis* genetic map (3), the *cccA*, *grpE*, *dnaJ*, *sigA* and *dnaG* genes were located in the downstream area of the *cdd* gene which was transcribed by clockwise orientation. However, the upstream region of the gene was not yet known. Based on the *B. subtilis cdd* gene encoding cytidine deaminase which catalyzes the conversion of cytidine/deoxycytidine to uridine/deoxyuridine, originally isolated and sequenced by Song and Neuhard (31), a further upstream region of the *cdd* gene was sequenced and analyzed. As a result, the location of the *pho*H and *dgk* genes was identified through sequence homology search as being in the upstream 3.3 kb stretch along with four unknown ORFs.

*Corresponding Author
Key words: *phoH, dgk, cdd,* diacylglycerol kinase, cytidine deaminase, *Bacillus subtilis*

MATERIALS AND METHODS

Bacterial Strains and Plasmids

The bacterial strains used in this work are *Bacillus subtilis* ED40 (*pyr-2 lys cdd-1*; 31), *Escherichia coli* JM109 [*recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi*△(*lac-proAB*)/F'(*traD36 proAB*$^+$ *lacI*^{q} *lacZ* △M15) ; 36], XL-1-Blue [*supE hsdR* lac$^-$/F'(*proAB*$^+$ *lacI*^{q} *lacZ* △M15) ; 6], and JF611 (*pyrE60 cdd thi-1 argE3 his-4 proA2 thr-1 leu-6 mtl-1 xyl-5 ara-14 galK2 lacY1 rpsL supE44* ; 31). Plasmid vectors used for subcloning and DNA sequencing are pSO1 (Ap^r Tet^r ; 7.3 kb *cdd*/pBR322 ; 31), pUC18/19 (Ap^r *lacZ lacI* ; 20), pBluescript II (Ap^r *lacZ lacI* ; 1), and M13 mp18/19 (*lacZ* ; 36).

Media and Cultural Conditions

E. coli strain was routinely cultured in Luria-Bertani (LB) medium (21). Plasmid-harboring *E. coli* cells were cultured in LB medium supplemented with ampicillin (50 μg/ml). M9 minimal medium (28) plates were used for maintenance of the JM109 episome. X-Gal agar plates (LB agar, 50 μg/ml ampicillin, 0.1 mM IPTG and 40 μg/ml X-Gal) were used for subcloning with the pUC series of vectors in *E. coli* JM109. *E. coli* JM109 harboring M13 phage vectors were grown in 2 x yeast-tryptone (YT) medium (28). The bottom and top agar were employed for subcloning with M13 phage vector.

DNA Manipulation

Plasmids were isolated by the method of SDS/alkaline lysis procedure (4). Restriction and exonuclease III digestion, and DNA ligation were performed by accordance with the manufacturers specification. After precise restriction mapping on the 11.5 kb insert of the pSO1, fragmentation with restriction enzyme and serial deletion with exonuclease III were employed for subcloning and deletion for rescue of the *cdd* upstream region. Nucleotide sequence was determined by using Sanger's dideoxy chain termination methods (29). The structural organization and identification of the genes were performed by searching the sequence homology using Genetyx software and BLAST program (2) on the databases of the GenBank (Genetic Sequence Data Bank), EMBL (European Molecular Biology Laboratory), NBRF (National Biomedical Research Foundation), and PIR (Protein Identification Resource).

Miscellaneous

Most restriction enzymes, RNase A, and T4 DNA ligase were obtained from Korea Steel Chemical Co., Promega Co. or Boerhinger-Mannheim. Exonuclease III and S1 nuclease were obtained from Pharmacia Co. Enzymes were used according to the recommendations of the manufacturers. 5-Bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) and isopropyl-β-D-thiogalactoside (IPTG) were obtained from Sigma Chemical Co.. The α-[^{35}S]-dATP and autoradiography film were obtained from Amersham and the DNA sequencing kits (Sequenase™ version 2.0 kit) were from United States Biochemical Co. and Taq DNA sequencing kit from Boerhinger-Mannheim. Geneclean kit was used for recovery of DNA fragment from agarose gel. All other chemicals and enzymes were obtained from commercial sources in reagent grade.

RESULT

Distribution of *B. subtilis* Chromosome in the pSO1

A 7.3 kb of insert in the pSO1 contains the *B. subtilis* chromosomal DNA fragment and λ-phage DNA because the pSO1 was originally cloned from λ-library of *B. subtilis* chromosome (31). To escape the λ-phage DNA segment in the insert of pSO1, Southern blotting was performed by using λ-phage as a probe.

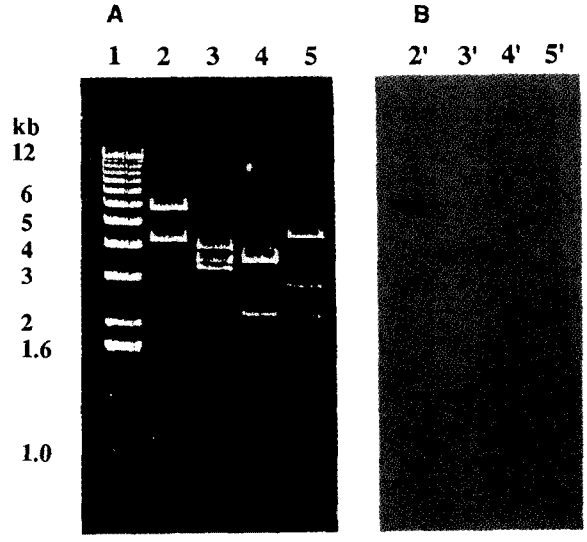

Fig. 1. Localization of the *B. subtilis* chromosomal DNA in the insert of pSO1 by Southern blotting.
EcoRI/HindIII fragments of *lambda* DNA were used as a probe after labelling with ^{32}P-dCTP. DNA electophotogram (A) and autoradiogram (B) represent as follows lane 1, 1 kb ladder as a size marker; 2, *EcoRI/HindIII* fragment; 3, *EcoRI/NruI* fragment; 4, *NruI/HindIII* fragment ; 5, *ClaI/NruI* fragment of pSO1 insert. Blotted fragments containing *lambda* DNA indicate arrows.

For identifying the λ-phage DNA segment in the insert of the pSO1, various double digested fragments of the stretch with *EcoRI* and *HindIII* were blotted with the corresponding restriction fragments of λ-phage DNA as a probe after labelling with [^{32}P]-dCTP. Fig. 1 illustrates that 5.7 kb fragments of $EcoRI_1/HindIII_1$ in lane 2, 3.4 kb of $NruI_1/NruI_2$ in lane 3 and 4, 2.2 kb of $NruI/ClaI_1$ in lane 5, were blotted by the λ-DNA. This result indicated that the $EcoRI_1/NruI_2$ segment of the insert originated from the λ-DNA as diagrammed in Fig. 2. Considering cloning process of the pSO1 by using λ-library of *Sau3AI* fragments of *B. subtilis* chromosome (31), the *Sau3AI/BamHI* site located just downstream of the $NruI_2$ thought to be a fusion site between *B. subtilis* chromosome and λ-DNA. The 4.5 kb stretch of the Sau3AI/BamHI and EcoRI edged segment was assumed to originate from the *B. subtilis* chromosome, and this segment was subjected to sequencing. Because of rare restriction sites in the $NruI_2/HindIII_1$ region, the fragmentation was performed with serial deletion by exonuclease III combined with primer extension. Combined explanations with restriction analysis, sequencing strategy, and analyzed genetic organization to the sequenced data after homology searching is diagrammed in Fig. 2.

Nucleotide Sequencing and Identification of the Coding Sequences.

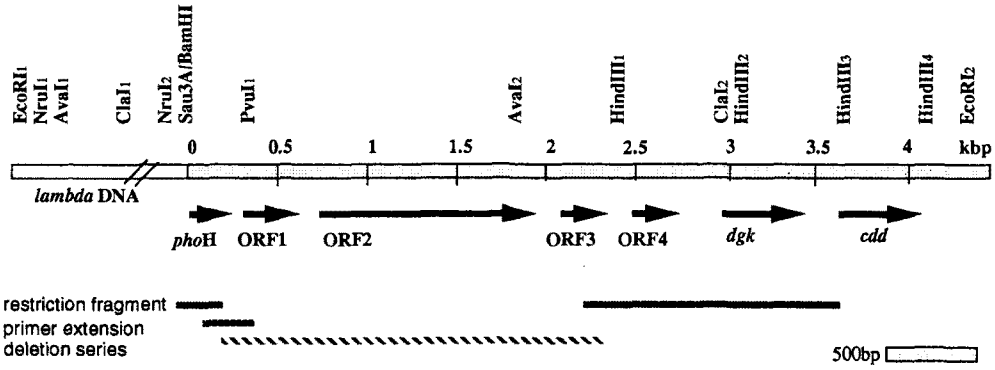

Fig. 2. Genetic organizaton, Sequencing Strategy and Restriction map of upstream stretch of *B. subtilis cdd* gene.
Main bar with scale and restriction map indicate the exogenote harbored in pSO1. After defining the *lambda* DNA (open bar) and B. *subtilis* chromosome (shadowed bar), about 3.3 kb stretch was sequenced by dideoxy termination method. After restriction mapping to the stretch, *Hind*III-edged fragments and *Nru*I$_2$ boundary region (dark shadowed bar) were subcloned to the M13 series vector and rest of the stretch was applied to the vector after serial fragmentation with exonuclease III (slant bar). One part of the stretch was rescued by using primer for extension of the stretch (dark slant bar). Defined four complete and one truncated coding sequences by homology searching expressed with arrows and gene marker together. Direction of the arrow indicates transcription orientation.

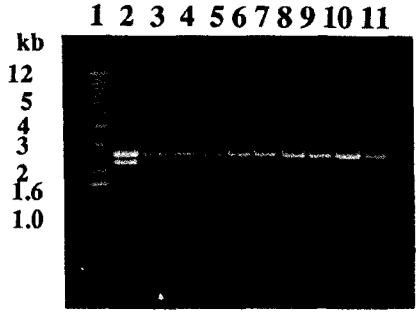

Fig. 3. Serial fragmentation with exonuclease III to the *Nru*I$_2$/*Hind*III$_1$ segment.
The segment digested serially with the exonuclease III by time interval and subjected to the agarose gel. Each lane represents various digested fragments from 2 to 11 with different size. Lane 1 indicates size marker of 1 kb ladder. A 2.7 kb fragment of vector after deletion of *Nru*I$_2$/*Hind*III$_1$ segment arranged at whole lanes with the same pattern in the middle part of the photogram.

Based on a previously determined 1,170 bp of *Cla*I$_2$/*Eco*RI$_2$ fragment harboring the *cdd* gene in pSO1, the remaining 3.3 kb of the *Sau*3AI/*Bam*HI-*Hin*dIII$_2$ stretch on the further upstream region of the *cdd* gene was sequenced and analyzed.

To escape the lambda DNA fragment from the sequencing, the 5.8 kb of *Eco*RI$_1$/*Hin*dIII$_1$ stretch was divided into 0.2 kb of *Eco*RI$_1$/*Nru*I$_1$, 3.3 kb of *Nru*I$_1$/*Nru*I$_2$ and 2.3 kb of *Nru*I$_2$/*Hin*dIII$_1$ by *Nru*I cutting, and then the latter fragment only subjected to the sequencing. About 600 bp of *Hin*dIII$_1$/*Hin*dIII$_2$ and 480 bp of *Hin*dIII$_2$/*Hin*dIII$_3$ fragments in the upstream of the *cdd* gene were subcloned into M13 series vectors. However, a further upstream region from *Hin*dIII$_1$ comprised rare restriction sites, therefore, this part was subcloned after serial cutting with the exonucleaseIII and the contiguous connection of each fragments was performed. But on the fragments cutted by serial digestion, one assumed junction part between the fragment shown in lane 2 and that in lane 3 was observed as shown in Fig. 3. To connecting this junction part of the sequence, the primer 5'-CAAGAGGAGGTTCTTGT-3' which complement in the positions from 245 to 261 in the sequence data of Fig. 4 was hybridized and extended.

The sequenced 3,346 bp of *Sau*3AI/*Bam*HI-*Hin*dIII$_3$ stretch as shown in Fig. 4 was analyzed with the Genetyx and Macvector software. Five complete and one truncated coding sequences were found in the same orientation with the *cdd* gene which read in the clockwise direction. No other reading frames were found in the reverse direction. Translated polypeptides deduced from these coding sequences were searched to homology with the known sequences in the protein databanks. As a result, one complete coding sequences in positions 2769 to 3194 and one truncated one in positions 2 to 179 were identified as the *dgk* (diacylglycerol kinase) and *phoH* (phosphate starvation-inducible gene) genes, respectively. However, other four coding sequences designated as ORF 1, 2, 3, and 4, were unknown.

The *B. subtilis dgk* gene was thought to be a result of sequence homology searching which revealed 43.1 % of identity index and 84.3 % of amino acid residues conservation to the *Streptococcus mutans dgk* gene. Another truncated coding sequence encoding carboxy terminal 59 residues, showed 52.7 % and 34.5 % of the sequence identity to the *phoH* genes of *Mycobacterium leprae* and *E. coli*. About 85.5 % of the *B. subtilis* PhoH residues was conserved in both *M. leprae* and *E. coli* corresponding enzymes (Fig. 5). The Kyte-Doolittle hydropathy profile (17) for deduced polypeptide of *B. subtilis* diacylglycerol kinase revealed almost the same

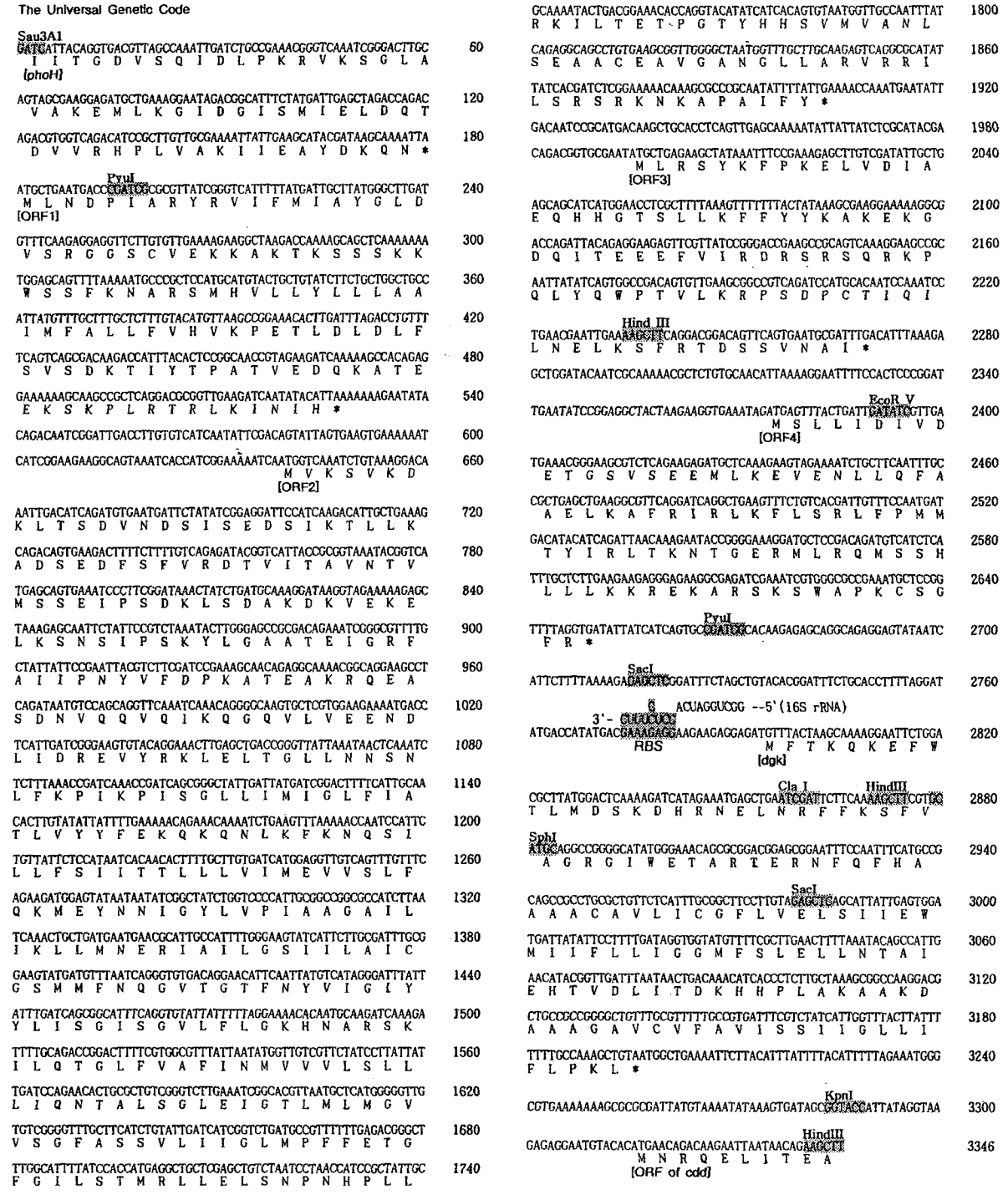

Fig. 4. Nucleotide sequence of upstream region of the *B. subtilis cdd* gene.
Four complete and one truncated coding sequences were defined from ORF searching by Genetyx software. Genetic organization of the six ORFs with start and end points of the reading frames were diagrammed. Restriction sites, ORFs with the single letters of amino acids residues, and SD sequence of the *dgk* gene were represented. Accession number is GenBank U29177.

pattem with that of *E. coli* (27) and *S. mutans* (35) enzymes (Fig. 6).

The remaining four coding sequences ORF1 in positions 181 to 525 (13.1 kDa, 115 amino acids), ORF2 in positions 639 to 1901 (46.4 kDa, 421 amino acids), ORF3 in positions 1995 to 2267 (10.8 kDa, 91 amino

acids), and ORF4 in positions 2375 to 2647 (10.7 kDa, 91 amino acids), revealed no significant similarity to the known polypeptides after homology searching.

Analysis of the Coding Sequences.

Each coding sequence with the endpoints at the start and stop positions, including Shine-Dalgarno (SD) sequences as a ribosome binding site (RBS) and the putative translation start codons (ATG), are listed in Table 1. The 5'-gAGGAaG-3' sequence as a SD sequence of the *dgk* gene was located in 9 bases upstream from the ATG codon. Unusual distance between SD sequence and ATG codon was observed in ORF2, but most of the other ORFs revealed homologous to the consensus sequences. Identifying the precise promoter regions and transcription initiation sites in each coding sequence need more experiments.

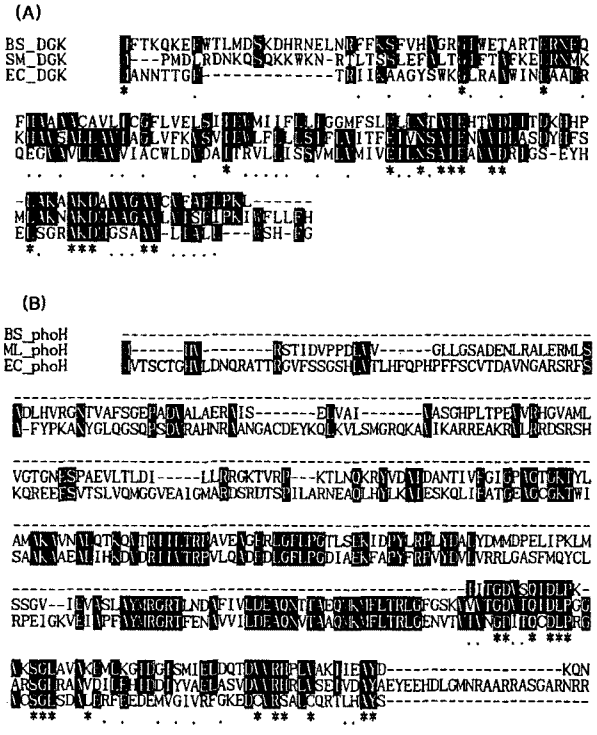

Fig. 5. Sequence homology of the *B. subtilis* Dgk (A) and PhoH (B) polypeptides to the corresponding known enzymes.
Multiple sequence alignments using Clustal V program were carried out. The polypeptides of one complete (134 residues) and one truncated (59 residues) aequences encoded by the *B. subtilis dgk* and *phoH* genes were compared with their corresponding polypeptides encoded by *S. mutans* and *E. coli dgk* genes, and by *M. leprae* and *E. coli phoH* genes, respectively. Homologous amino acids represent in the black boxes. The symbols asterisk and dot below the sequences denote invariant and conservative replacement, respectively.

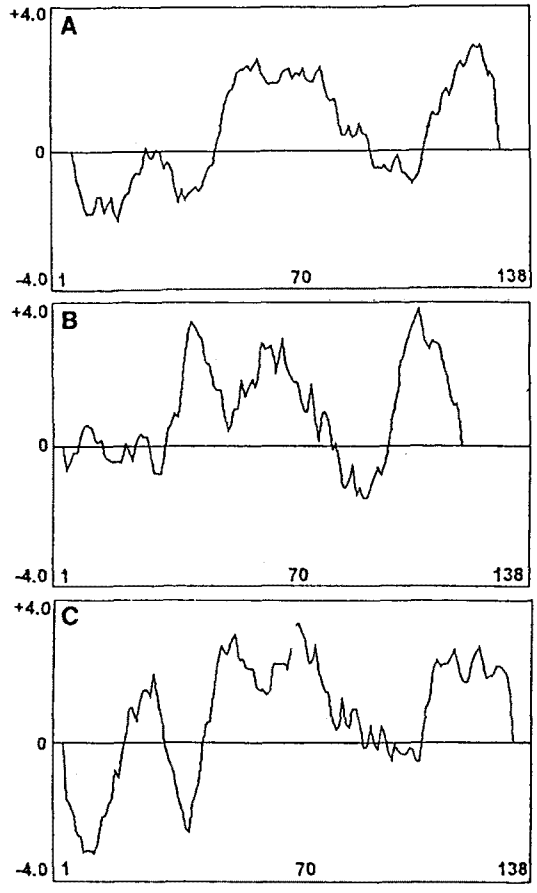

Fig. 6. Comparison of hydropathy profiles for the poly peptides encoded by the *dgk* genes from different bacterial origin.
Hydropathy was analyzed by sequence algorithm of Kyte and Doolittle (17) with Genetyx software. The polypeptides encoded by *B. subtilis* (A), *S. mutans* (B), and *E. coli* (C) were diagrammed. Horizontal scale represents the number of residues from amino terminal initiation and vertical scale represents hydrophillicity (+) and hydrophobicity (−).

Table 1. Genetic organization of the stretch with SD sequence and translational initiation codon[a].

Gene names	Endpoints	Molecular size aa/kilo-dalton	Tanslational start[b]
phoH	2-›179	59 /	-
ORF1	81-›525	115 / 13.1	AAGcAaaaattaATGCTGAAT
ORF2	639-›1901	421 / 46.4	AGGcaGTaaatcaccatcggaaaatcaATGGTCAAA
ORF3	1995-›2267	91 / 10.8	cAGacGGTGcgaatATGCTGAGA
ORF4	2375-›2647	91 / 10.7	AAGaAGGTGaaatagATGAGTTTA
dgk	2769-›3194	134 / 15.1	gAGGAaGaagaggagATGTTTACT
cdd	3315-›371		

[a] Five complete and one truncated coding sequences, [b] SD sequence is underlined.

DISCUSSION

Genome analysis for *B. subtilis* and *E. coli* was performed systematically. The Gram positive bacterium *B. subtilis* which diverged from the Gram negative strain *E. coli* about 2 billion years ago (34), was used as a useful target for spore formation and extracellular enzyme secretion. The *B. subtilis* chromosome is composed of 4,175 kb and about 20~30% of the whole genome has been progressed to sequencing. The upstream region of the *cdd* gene which appears at about 223 degrees on the genetic map was analyzed from the harbored fragment on pSO1 in which the fragment was cloned by complementation of the *cdd* gene from the lambda library of *B. subtilis* chromosomal DNA. Five complete and one truncated coding sequences were deduced from the sequenced 3.3 kb stretch by ORF searching. Two of them were identified as homologous to the known sequences such as the *dgk* and *phoH* genes, however the remained four coding sequences did not reveal any significant identity.

The various *dgk* genes encoding diacylglycerol kinase from different sources, such as *E. coli* (18), *Drosophila* (11, 19), human (7, 12), and rat (9, 10), were characterized. This enzyme catalyzes the phosphorylation of diacylglycerol and has a role in adaptive response against environmental stress such as low pH, and high osmolarity or temperature. In the aspect of signal transduction, this *Bacillus dgk* gene might illustrate more precisely the phosphorylation of diacylglycerol combined with the function of a second messenger with the activation of protein kinase C. Simultaneous Kyte-Doolittle hydropathy profile of *B. subtilis* diacylglycosyl kinase to that of *E. coli* one (27) suggested the similar pattern of transmembrane topology of both polypeptides.

When searching sequence homology to the reported genes stocked in the databanks, the truncated polypeptide at the 5′ terminal of the sequenced stretch was shown to be homologous to the PhoH of *E. coli* (13) and *Mycobacterium leprae* (U00016). Because the full length of these PhoH polypeptides is 334 and 354 residues in each of these strains, this comparison on truncated 59 residues in carboxy terminal could have some inconsistency. However, considering the reasonable high percentage of sequence identity, the truncated peptide might react as a PhoH funtionally. Of course, it needs more study on its enzymatic function and the truncated coding sequence supposed to be the *B. subtilis phoH* gene. When this gene product has a function of ATP-binding activity and has a role for transport of the combined or free phosphates like *E. coli* PhoH, the gene might be identified as the *phoH* gene. Combined data for structural analysis in protein level and primer extension to each coding sequences including four unknown ORFs will illustrate the precise structure and genetic organization in this stretch. Anyway this may be the first report for localizing and mapping the *B. subtilis dgk* and *phoH* locus, and for nucleotide sequencing of the *phoH-cdd* stretch located in the 223 degrees of the genetic map.

Acknowledgement

We are grateful to Prof. J. Neuhard in the Copenhagen University, Denmark for the gift of pSO1 and valuable advice. Database searches were performed using the BLAST network service. This work was supported by the '93 Genetic Engineering Research Grant from the Korea Ministry of Education.

REFERENCES

1. Alting-Mees, M. A., J. A. Sorge, and J. M. Short. 1992. pBluescriptII: multifunctional cloning and mapping vectors. *Methods Enzymol.* **216**: 483-495.
2. Altschul, S. F., W. Gish , W. Miller, E. W. Myers, and, D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.
3. Anagnostopoulos, C. P. J. Piggot, and J. A. Hoch. 1993. The genetic map of *Bacillus subtilis*, p. 425-461. *In* A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis and other gram positive Bacteria; Biochemistry, Physiology and Molecular Genetics*, ASM, Washington, D.C.
4. Bimboim, H. C. and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**: 1513-1523.
5. Blattner, F. R., V. Burland, G. Plunkett, H. J. Sofia, and D. L. Daniels. 1993. Analysis of the *Escherichia coli* genome IV DNA sequence of the region from 89.2 to 92.8 minutes. *Nucleic Acids Res.* **21**: 5408-5417.
6. Bullock, W. O., J. M. Fernandez, and J. M. Short. 1987. A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Bio Techniques.* **5**: 376.
7. Fujikawa, K., S. Imai, S. F., and H. Kanoh. 1993. Isolation and characterization of the human diacylglycerol kinase gene. *Biochem. J.* **294**: 443-449.
8. Glaser, P., F. Kunst, M. Arnaud, M. P. Coudart, W. Gonzales, M-F. Hullo, M. Ionescu, B. Lubochinsky, L. Marcelino, I. Moszer, E. Presecan, M. Santana, E. Schneider, J. Schweizer, A. Vertes, G. Rapoport, and A. Danchin. 1993. *Bacillus subtilis* genome project: cloning and sequencing of the 97 kilo bases region from 325° to 333°. *J. Mol. Microbiol.* **10**: 371-384.
9. Goto, K., H. Watanabe, H. Kondo, H. Yuasa, F. Sakane, and H. Kanoh. 1992. Gene cloning, sequence, expression and *in situ* localization of 80 kDa diacylglycerol kinase specific to oligodendrocyte of rat brain. *Brain Res. Mol. Brain Res.* **16**: 75-87.
10. Goto, K. and H. Kondo. 1993. Molecular cloning and

expression of a 90-kDa diacylglycerol kinase that predominantly localizes in neurons. *Proc. Natl. Acad. Sci. USA* **90**: 7598-7602.

11. Inoue, H., T. Yoshioka, and Y. Hotta. 1992. Partial Purification and Chracterization of membrane-associated diacylglycerol kinase of Drosophila heads. *Biochim. Biophys. Acta.* **1122**: 219-224.
12. Kai, M., Sakane, F. Sakane, S. Imai, I. Wada, and H. Kanoh. 1994. Molecular cloning of a diacylglycerol kinase isozyme predominantly expressed in human retina with a truncated and inactive enzyme expression in most other human cells. *J. Biol. Chem.* **269**: 18492-18498.
13. Kim, S. K., K. Makino, M. Amemura, H. Shinagawa, and A. Nakata. 1993. Molecular Analysis of the *phoH* gene, Beloning to the phosphate regulon in *Escherichia coli. J. Bateriol.* **175**: 1316-1324.
14. Kohara, Y., K. Akiyama, and K. Isono. 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* **50**: 495-508.
15. Kominami, R. *et al.* 1993. Genome analysis of mice. *Protein Nucleic acid and Enzymes* **38**: 696-703.
16. Kunst, F. and K. Devine. 1991. The project of sequencing the entire *Bacillus subtilis* genome. *Res. Microbiol.* **142**: 905-912.
17. Kyte, J. and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.
18. Lightner, V. A., R. M. Bell, and P. Modrich. 1983. The DNA sequences encoding *plsB* and *dgk* loci of *Escherichia coli. J. Biol. Chem.* **258**: 10856-10861.
19. Masai, I., A. Okazaki, T. Hosoya, and Y. Hotta. 1993. Drosophila retinal degeneration A gene encodes an eye-specific diacylglycerol kinase with cysteine-rich zinc-finger motifs and ankyrin repeats. *Proc. Natl. Acad. Sci. USA* **90**: 11157-11161.
20. Messing, J. and J. Vieira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* **19**: 269-276.
21. Miller, J. H. 1972. *Experiments in molecular genetics.* p. 352-355. Cold Spring Harbor, New York.
22. Minobe, Y. 1993. Analysis of rice genome. *Protein Nucleic acid and Enzyme* **38**: 704-712.
23. Mizukami, T., I. Garkavtsev, D. B. Marr, O. Niwa, and M. Yanagida. 1993. Genome mapping of *Schizosaccharomyces pombe. Protein Nucleic acid and Enzymes* **38**: 677-684.
24. Ogasawara, N., S. Nakai, and H. Yoshikawa. 1994. Systematic sequencing of the 180 kilobase region of the *B. subtilis* chromosome containing the replication origin. *DNA research* **1**: 14-20.
25. Oliver, S. G. *et al.* 1992. The complete DNA sequence of yeast chromosome III. *Nature* **357**: 38-46.
26. Regad, F., M. Lebas, and B. Lescure. 1994. Interstitial telomeric repeats within the *Arabidopsis thaliana* genome. *J. Mol. Biol.* **239**: 163-169.
27. Ronald, L. S., J. F. O'toole, M. E. Maguire, and C. R. Sanders II. 1994. Membrane topology of *Escherichia coli* diacylglycerol kinase. *J. Bacteriol.* **176**: 5459-5465.
28. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual,* second edition. Cold Spring Harbor Lab., New York.
29. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.
30. Shinomiya, T. and S. Ina. 1994. Mapping an initiation region of DNA replication at a single-copy chromosomal locus in Drosophila melanogaster cells by two-dimensional gel methods and PCR-mediated nascent-strand analysis: multiple replication origins in a broad zone. *Mol. Cell Biol.* **14**: 7394-7403.
31. Song, B. H. and J. Neuhard. 1989. Chromosomal location, cloning and nucleotide sequence of the *Bacillus subtilis cdd* gene encoding cytidine/deoxycytidine deaminase. *Mol. Gen. Genet.* **216**: 462-468.
32. Sulston, J., Z. Du, K. Thomas, R. Wilson, L. Hillier, R. Staden, N. Halloran, P. Green, J. Thierry-Mieg, L. Qiu, S. Dear, A. Coulson, M. Craxton, R. Durbin, M. Berks, M. Metzstein, T. Hawkins, R. Ainscough, and R. Waterston. 1992. The *C. elegans* genome sequencing project: a beginning. *Nature* **356**: 37-41.
33. Wienberg, J., A. Jauch, H. J. Ludecke, G. Senger, B. Horsthemke, U. Claussen, T. Cremer, N. Arnold, and C. Lengauer. 1994. The origin of human chromosome 2 analyzed by comparative chromosome mapping with a DNA microlibrary. *Chromosome Res.* **2**: 405-410.
34. Woese, C. R. 1987. Bacterial evolution. *Microbiol Rev.* **51**: 221-271.
35. Yamashita, Y., T. Takehara, and H. K. Kuramitsu. 1993. Molecular characterization of a *Streptococcus mutans* mutant altered in environmental stress responses. *J. Bacteriol.* **175**: 6220-6228.
36. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

(Received March 4, 1995)