JOURNAL OF BACTERIOLOGY, May 2000, p. 2387–2392
0021-9193/00/$04.00+0
Copyright © 2000, American Society for Microbiology. All Rights Reserved.
Vol. 182, No. 9

A New IS*4* Family Insertion Sequence, IS*4Bsu*1, Responsible for Genetic Instability of Poly-γ-Glutamic Acid Production in *Bacillus subtilis*

TOSHIRO NAGAI,† LAM-SON PHAN TRAN, YASUHIRO INATSU, AND YOSHIFUMI ITOH*

National Food Research Institute, Ministry of Agriculture, Forestry and Fisheries, Tsukuba 305-8642, Japan

Received 22 October 1999/Accepted 29 January 2000

Certain *Bacillus subtilis* strains, such as *B. subtilis* (*natto*) starter strains for the manufacture of natto (fermented soybeans), produce capsular poly-γ-glutamate (γPGA). In *B. subtilis* (*natto*), γPGA synthesis is controlled by the ComP-ComA two-component regulatory system and thereby induced at the beginning of the stationary growth phase. We have found a new insertion sequence (IS), designated IS*4Bsu1*, in the *comP* gene of a spontaneous γPGA-negative mutant of *B. subtilis* (*natto*) NAF4. IS*4Bsu1* (1,406 bp), the first IS discovered in *B. subtilis*, encodes a putative transposase (Tpase) with a predicted M_r of 34,895 (374 residues) which displays similarity to the Tpases of IS*4* family members. Southern blot analyses have identified 6 to 11 copies of IS*4Bsu1*, among which 6 copies were at the same loci, in the chromosomes of *B. subtilis* (*natto*) strains, including NAF4, three commercial starters, and another three γPGA-producing *B. subtilis* (*natto*) strains. All of the eight spontaneous γPGA$^-$ mutants, which were derived from five independent NAF4 cultures, had a new additional IS*4Bsu1* copy in *comP* at six different positions within 600 bp of the 5′-terminal region. The target sites of IS*4Bsu1* were determined to be AT-rich 9-bp sequences by sequencing the flanking regions of IS*4Bsu1* in mutant *comP* genes. These results indicate that IS*4Bsu1* transposes by the replicative mechanism, in contrast to other IS*4* members that use the conservative mechanism, and that most, if not all, of spontaneous γPGA$^-$ mutants appear to have resulted from the insertion of IS*4Bsu1* exclusively into *comP*. The presence of insertion hot spots in *comP*, which is essential for γPGA synthesis, as well as high transposition activity, would account for the high frequency of spontaneous γPGA$^-$ mutation by IS*4Bsu1* in *B. subtilis* (*natto*).

Insertion sequences (ISs) are small mobile units of DNA consisting of, in general, a unique gene (*tnp*) for transposase (Tpase) and terminal inverted repeats (IRs) that serve as the sites for recognition and cleavage by Tpases in transposition reactions (9, 10, 30, 37). Some ISs form complex transposable elements, i.e., transposons, by flanking a DNA region encoding antibiotic resistance or containing catabolic or pathogenic genes (3, 30, 40, 44). It is well recognized that transposition of such mobile elements results in insertional mutation or activation of a downstream gene (6, 8, 13, 19, 22, 23, 27, 43). As ISs and transposons are often associated with transmissible plasmids and bacteriophages, they distribute in a wide range of bacteria by horizontal transmission. To date a large number of ISs have been identified in many organisms, and these have been classified into 17 families based principally on the amino acid sequence similarities of their Tpases (30). Many bacteria possess multiple ISs of different families in their genome. However, no IS has been reported for *Bacillus subtilis* strains, including *B. subtilis* 168, whose whole genome sequence has been determined (24).

B. subtilis (*natto*), a starter strain for production of natto (fermented soybeans) (47), produces a unique capsular polymer of glutamate with a γ-peptide linkage, poly-γ-glutamate (γPGA) (35). γPGA synthesis in *B. subtilis* appears to occur in the stationary growth phase (46). Such growth phase-dependent production of γPGA has been well recognized in natto factories. In a conventional natto fermentation process (at around 40°C for 20 h), the γPGA production starts after ca. 16 h and rapidly reaches the levels appropriate for natto products, at 20 h. Cell density-dependent phenotypes of *B. subtilis* are regulated by a quorum-sensing mechanism involving the ComP-ComA two-component signal transduction system. The *comP* gene specifies a sensor protein kinase of this two-component regulatory system and is included in the *comQXPA* quorum-sensing operon (11, 25, 29, 39, 49). The sensor domain of ComP has eight transmembrane helices and four extracytosolic loops which are most likely to interact with the extracellular ComX pheromone (36). The ComX pheromone, a 10-amino-acid (aa) peptide with a modified tryptophan residue, is processed from the C terminus of ComX, perhaps by ComQ (25, 29), and accumulates in the medium as cells grow to high density to act as a quorum-sensing signal. By analogy with other two-component systems (42), the interaction between the external pheromone and the sensor domain of ComP induces the activation of the ComP-ComA signal transduction system, resulting in phosphorylation of the cognate response regulator ComA to give ComA-PO_4. ComA-PO_4, an active transcription regulator, then mediates the expression of several genes, including *srfA* and *degQ* (11, 25, 32, 33, 39). The *srfA* operon includes the surfactin synthetase genes and *comS* (11, 16). ComS is required for upregulation of *comK*, encoding the regulator protein of late competence genes (11, 12, 25, 39), while DegQ controls the expression of degradative enzymes. Thus, the *comQXPA* operon plays a key role in the production of degradative enzymes and surfactin, in the development of genetic competence, and in the adaptation of cells to a high cell population density (12, 25, 33). We have found that insertional inactivation of *comP* by Tn*917*-LTV1 abolishes the γPGA production of *B. subtilis* (*natto*) NAF4 (Y. Itoh, Y. Inatsu, T. Nishijo, and T. Nagai, Abstr. Int. Conf. Bacilli, abstr. 133,

* Corresponding author. Mailing address: National Food Research Institute, Kannondai 2-1-2, Tsukuba 305-8642, Japan. Phone: 81-298-38-8075. Fax: 81-298-38-7996. E-mail: yosifumi@nfri.affrc.go.jp.

† Present address: Genetic Resource Center, National Institute of Agrobiological Resources, MAFF, Tsukuba 305-8602, Japan.

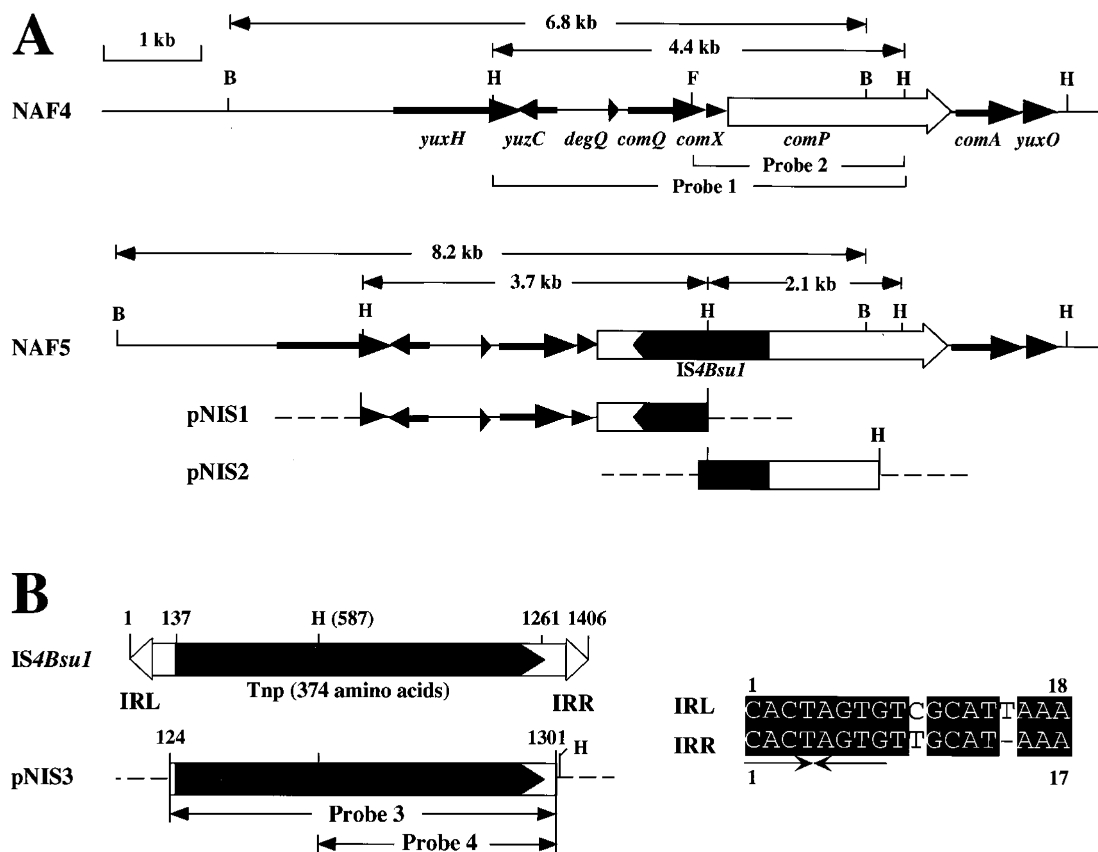

FIG. 1. Structures of the *comP* gene of PGA⁻ mutant *B. subtilis* (*natto*) NAF5, IS*4Bsu1*, and plasmids containing part of the IS*4Bsu1* sequence. (A) The *yuxH*, *yuzC*, *degQ*, *comQXPA*, and *yuxO* genes of *B. subtilis* (*natto*) NAF4 are assigned by sequence similarity to the corresponding genes of *B. subtilis* 168 (accession no. Z99120). Broken lines present the pUC118 sequence (48). Abbreviations for restriction sites: B, *Bgl*II; H, *Hin*dIII; F, *Fsp*I. (B) The numbers refer to the nucleotide coordinates of IS*4Bsu1* (accession no. AB031551). pNIS3 contains the PCR-amplified IS*4Bsu1* sequence (nt 124 to 1301) that was used in the Southern blot experiments as a probe. IRL and IRR, IRs at the upstream (left) and downstream (right) termini, respectively. The broken line represents the vector pUC118. The 8-bp palindrome structure within the terminal IRs is indicated by arrows. H, *Hin*dIII site.

1998), indicating that the γPGA production in *B. subtilis* (*natto*) is also controlled by the ComP-ComA system. The structure of the *comQXPA* operon of *B. subtilis* (*natto*) NAF4 and properties of the *comP*::Tn*917*-LTV1 mutant will be described elsewhere.

It has been noted that spontaneous mutants of *B. subtilis* (*natto*) that are defective in γPGA production (γPGA⁻) arise at a high frequency (ca. 1% in 20 generations grown under normal conditions) by an as-yet-unknown mechanism (34, 35). When we analyzed the *comP* gene of a spontaneous γPGA⁻ mutant by Southern blotting, we found an insertion of a 1.4-kb DNA fragment in this gene. Nucleotide sequencing of the insert revealed the first IS element in *B. subtilis*, designated IS*4Bsu1*. In the present work we describe the structure, transposition mechanism, and target site of IS*4Bsu1*.

MATERIALS AND METHODS

Bacterial strains, plasmids, and media. The following γPGA-producing *B. subtilis* (*natto*) strains were used in this study. *B. subtilis* (*natto*) NAF4 (Rif^r γPGA⁺) is derived from *B. subtilis* (*natto*) Asahikawa (34, 35), and *B. subtilis* (*natto*) NAF5 (Rif^r γPGA⁻) (34) is a spontaneous mutant of NAF4 that is defective in γPGA production (γPGA⁻). *B. subtilis* (*natto*) Naruse, Miura, and Takahashi were obtained from commercial source, and other γPGA-producing *B. subtilis* (*natto*) AR strains were from our laboratory collection (35). *B. subtilis* IFO3335MU5 carrying pLS20 (21) was obtained from S. Bron, and pUC118 (48) was used in gene cloning experiments with *Escherichia coli* DH5α [F⁻/*endA1 hsdR17* (r_K^-, m_K^+) *supE44 thi-1 recA1 gyrA relA1* Δ(*lacIZYA-argF*)*U169 deoR* (φ80d*lac*Δ(*lacZ*)*M15*)] (Bethesda Research Laboratories, Bethesda, Md.) as the host. Luria-Bertani medium (41) and GSP (35) medium were used to grow cells from which the chromosomal DNA and plasmid DNA were prepared and to assay γPGA production, respectively. Ampicillin (100 μg/ml) was added to Luria-Bertani medium when appropriate.

DNA techniques. Chromosomal DNA was purified from *B. subtilis* (*natto*) cells by CsCl-ethidium bromide density gradient centrifugation (7) or by the method of Bron (5). Plasmid DNA of *E. coli* was purified using a Midi kit (Qiagen, Chatsworth, Calif.) or a Flex-Prep Kit (Pharmacia LKB, Uppsala, Sweden), and pLS20 of *B. subtilis* IFO3335MU5 was isolated according to Bron (5) and purified by CsCl-ethidium bromide density gradient centrifugation. DNA ligation was done using a DNA Ligation Kit Version 1 (Takara Shuzo, Kyoto, Japan), and DNA was transformed into *E. coli* DH5α as described previously (18). Restriction endonuclease digestion and agarose gel electrophoresis were performed as described previously (41). PCR was performed using DNA polymerase KOD (Toyobo Biochemicals, Osaka, Japan) under the optimal conditions recommended by the supplier. Oligonucleotide primers were purchased from Hokkaido Service (Sapporo, Japan).

Cloning of IS*4Bsu1*. The chromosomal DNA of *B. subtilis* (*natto*) NAF5, which carries IS*4Bsu1* in *comP*, was digested with restriction endonuclease *Hin*dIII, ligated into pUC118 (48) at the *Hin*dIII site, and subsequently transformed into *E. coli* DH5α. A transformant that harbors recombinant plasmid pNIS1 containing the 3′ part of IS*4Bsu1* was identified by colony hybridization using the 4.4-kb *Hin*dIII fragment containing the *yuxH′-yuzC-degQ-comQXP′* genes (probe 1) (Fig. 1A) as a probe. Two oligonucleotide primers, 5′-GTTCATAATCCAAGT AACCCCG-3′ (complementary to nucleotide [nt] 725 to 746 of IS*4Bsu1*; accession no. AB031551) and 5′-ATGATCTGTCTCCTCGTTCACC-3′ (complementary to nt 1474 to 1495 of *comP*; accession no. AB031552), were used to amplify the 5′ half of IS*4Bsu1* from the NAF5 chromosome by PCR. The amplified DNA fragment was cloned into the *Hin*cII site of pUC118, giving pNIS2 (Fig. 1A). Then, the sequence from nt 124 to 1301 of IS*4Bsu1* was amplified by PCR using the primers 5′-AAGGACAATAAGCATGGATAAG-3′ (nt 124 to 145) and 5′-ACTATAATCTTTGACGAGTGCA-3′ (complementary

to nt 1280 to 1301) and cloned into pUC118 as described above to generate pNIS3 (Fig. 1B).

Southern blotting and colony hybridization. Restricted DNA fragments were separated by electrophoresis on 1% agarose HS (Nippon Gene, Toyama, Japan) and transferred onto a Hybond-N$^+$ nylon membrane (Amersham Pharmacia Biotec, Amersham, United Kingdom) with a VacGene blotter (Pharmacia LKB). After fixation at 80°C for 120 min, hybridization was performed using DNA probes prepared by means of a Random Prime Labeling Kit (Amersham Pharmacia Biotec), followed by detection of hybridized DNA with ECL Detection Systems (version II; Amersham Pharmacia Biotec). Colony hybridization was also performed with ECL Detection Systems, as described above, after transferring colonies onto a Hybond-N$^+$ nylon membrane, and cells were lysed by a standard method (41).

Nucleotide sequencing and DNA analysis. The nucleotide sequences of both strands were determined using an ABI310 DNA sequencer and Big-Dye primer and terminator sequencing kits (Perkin-Elmer ABI, Foster City, Calif.). DNA and amino acid sequences were analyzed using the programs Blast (1), Clustal W (45), and GCG (Genetics Computer Group, Madison, Wis.).

Determination of insertion sites in *comP*. The *comP* sequences carrying IS*4Bsu1* were PCR amplified from the mutant chromosomes with primers 5′-AGTCGGGTTCTCTGGTAACATTGCCCAG-3′ (nt −636 to −663) and 5′-CACCTCTTCACGGCACGGATTATCACCC-3′ (nt 1800 to 1827). The *comP* sequences contiguous to the IRs were determined using the PCR-amplified DNA fragments as templates and the sequencing primers 5′-GTGTAAACTTATCCATGCTTATTGTC-3′ and 5′-GATGTTAACTACCTCTATTCAAATGTC-3′, corresponding to nt 127 to 152 and 1311 to 1337 (complementary strand) of the IS*4Bsu1* sequence, respectively.

Nucleotide sequence accession number. The nucleotide sequence of IS*4Bsu1* is available in the DNA databases under accession no. AB031551.

RESULTS AND DISCUSSION

Spontaneous PGA$^-$ mutation. γPGA$^-$ mutants, which are defective in γPGA production, of *B. subtilis* (*natto*) strains (including commercial starter strains) occur spontaneously at rates of as high as 1 to 5% after 20 generations under normal growth conditions (34). γPGA$^-$ mutants emerge at much higher frequencies in nutrient agar cultures stored at room temperature. When colonies from five independent cultures of NAF4, which had been maintained for approximately 6 months, were tested for PGA production, γPGA$^-$ mutants occurred at a rate of 18%, on average.

Identification and structure of IS*4Bsu1*. The spontaneous γPGA$^-$ mutant *B. subtilis* (*natto*) NAF5 was derived from *B. subtilis* (*natto*) NAF4 (35). Since NAF5 produced a reduced level of proteases, as did *B. subtilis* (*natto*) NAF12 (*comP*::Tn*917*-LTV1), which is defective in γPGA synthesis (Itoh et al., Abstr. Int. Conf. Bacilli, 1998), we assumed that NAF5 would have a structural alternation in *comP* or its cognate *comQXA* genes. To investigate this assumption, the chromosomal DNA of NAF5, after digestion with restriction endonuclease *Hin*dIII, was probed with the 4.4-kb *Hin*dIII fragment (probe 1) containing the *yuzC′-degQ-comQXP′* region (Fig. 1A). Southern blotting revealed two *Hin*dIII fragments, 3.7 and 2.1 kb in length, in the NAF5 chromosomal DNA (Fig. 2), indicating that the relevant chromosomal region of NAF5 has an extra DNA sequence of at least 1.4 kb containing the *Hin*dIII site. Another Southern blot analysis with the *Bgl*II-digested chromosomal DNAs of parent NAF4 and mutant NAF5 gave 6.8- and 8.2-kb fragments, respectively (Fig. 2), demonstrating 1.4 kb of the insert. To determine the DNA sequence of the insert, the DNA regions covering the whole insert were cloned in three plasmids, pNIS1, pNIS2, and pNIS3 (Fig. 1), as described in Materials and Methods. The nucleotide sequence data for the cloned DNA fragments revealed that insertion had taken place at nt 406 of the *comP* gene (Fig. 1A) and that the insert was 1,406 bp in size (Fig. 1B). The insert encodes a putative Tpase with a predicted M_r of 34,895 (374 residues). A Blast search (1) revealed low but significant similarity (similarity scores of from 32 to 136) of this protein to the IS*4* Tpases, but no Tpase of other IS families that had a similarity score higher than 30 was detected. The putative

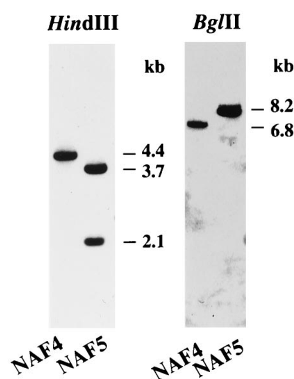

FIG. 2. Identification of an insertion in the *yuzC-comP* region of *B. subtilis* (*natto*) NAF5. The *Hin*dIII and *Bgl*II fragments containing the *yuzC-comP* regions of *B. subtilis* (*natto*) NAF4 and NAF5 were analyzed by Southern blotting using probe 1 (Fig. 1A).

Tpase of this IS element shares 28% identical amino acids with the Tpases of IS*5377* of *Bacillus stearothermophilus* CU21 (accession no. X67862) and other IS*4* family members (17 to 26% identity), including IS*4* (accession no. J01733). Moreover, the DDE motif, which is conserved in most Tpases and other enzymes catalyzing cleavage of DNA or RNA strands (10, 30, 37, 38), could be found in the Tpase, i.e., D (aa 127), D (aa 196), E (aa 296), and K (aa 303). The distance, ca. 110 aa, between the DE motifs in the Tpases of IS*4* family members is longer than that in other Tpases (ca. 35 aa) (10, 30, 37). This characteristic distance is also conserved in the putative Tpase (100 aa). Most ISs have short terminal IRs of between 10 and 40 bp. The insert has IRs of 18 bp at the ends (Fig. 1B). The distal 9-bp sequences match perfectly and have a palindrome structure of 8 bp (Fig. 1B). This insertion DNA is flanked by a 9-bp duplication of the target site (see below), probably as a consequence of transposition (10, 15, 17). These results support the idea that the IS, designated IS*4Bsu1*, belongs to the IS*4* family.

Copy number of IS*4Bsu1* in NAF4. The copy number of IS*4Bsu1* in the *B. subtilis* (*natto*) NAF4 chromosome was determined by Southern blotting. The NAF4 chromosomal DNA was digested with restriction endonuclease *Eco*RV or *Pst*I (neither cut the IS sequence) and hybridized with the IS*4Bsu1* DNA (probe 3) (Fig. 1B). Nine positive fragments were identified with both *Eco*RV- and *Pst*I-digested DNAs (Fig. 3), indicating the presence of nine copies or isoforms of IS*4Bsu1* on the chromosome. When the NAF4 chromosomal DNA was digested with *Spe*I, which cuts IS*4Bsu1* at nt 2 and 1400 (Fig. 1B), only a band of 1.4 kb was detected (Fig. 4). Furthermore, double digestion with *Spe*I and *Hin*dIII, which cuts at nt 587 (Fig. 1B), resulted in two bands of 0.82 and 0.58 kb (Fig. 4). Thus, these nine detected DNA fragments appear to share a sequence very similar, and perhaps identical, to that of IS*4Bsu1*.

Distribution of IS*4Bsu1* in *B. subtilis* strains. There have been no reports of ISs in *B. subtilis*, and no IS is present in the *B. subtilis* 168 chromosome, with the exception of very short (ca. 200-bp) noncoding sequences that are considered to be remnants of ISs, although they have no homology to known IS elements (20). As revealed by sequencing of bacterial genomes, multiple ISs of different families occur in many bacteria, including both eubacteria and archaebacteria. For instance, *E. coli* K-12 has one to seven copies or isoforms of 10 ISs that belong to six distinct families (4). Therefore, *B. subtilis*

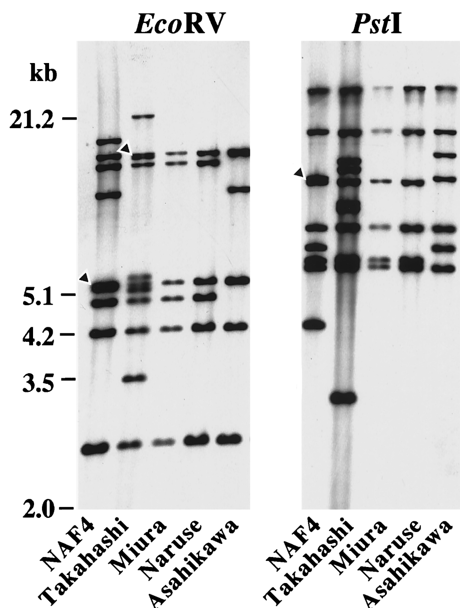

FIG. 3. Southern blot analyses of the IS*4Bsu1* sequences in *B. subtilis* (*natto*) strains. The indicated chromosomal DNAs of the *B. subtilis* strains were digested with *Eco*RV or *Pst*I, and the IS*4Bsu1* sequences were detected with probe 3 (Fig. 1B). The duplicate bands are indicated by triangles to the left.

seems to be an exception with respect to the inheritance of ISs. *B. subtilis* (*natto*) NAF4, which contains IS*4Bsu1* and is known to produce γPGA, is a derivative of a *B. subtilis* (*natto*) strain (called Asahikawa) originally isolated from fermented natto (35). Since ISs are often associated with genes that confer particular traits to cells, such as antibiotic resistance, catabolism of certain compounds, or pathogenesis (3, 28, 40, 44), it seemed likely that IS*4Bsu1* might be associated with the genes involved in γPGA synthesis. We therefore investigated three commercial starter strains, Takahashi, Miura, and Naruse, and 16 γPGA-producing AR strains from our collection (35) for inheritance of IS*4Bsu1*. As shown in Fig. 3, the Takahashi strain produced 11 IS*4Bsu1* copies, and the Miura and Naruse strains have 6 copies. These strains appear to share the six copies at the same loci, as demonstrated by the six common positive *Eco*RV and *Pst*I fragments (Fig. 3). NAF4 has an additional three copies of IS*4Bsu1* that are not present in its parent Asahikawa, while one copy in Asahikawa is absent from NAF4 (Fig. 3). Again, digestion with *Spe*I and with both *Spe*I and *Hin*dIII gave the bands of 1.4 kb and of 0.82 and 0.58 kb, respectively, (Fig. 4). We then analyzed 16 other γPGA-producing *B. subtilis* AR strains from our laboratory collection (35). Three natto isolates, AR6, AR39, and AR74, were found to harbor IS*4Bsu1*. Seven positive *Eco*RV fragments were identified in the AR6 strain and six fragments were identified in the AR39 and AR74 strains (data not shown), and the six positive bands in the AR6, AR39, and AR74 strains correspond to the six common bands found in NAF4 and the starter strains (Fig. 3). On the other hand, no positive band was detected in the other 13 γPGA-producing AR strains that were isolated from sources other than natto (data not shown). These results indicate that the *B. subtilis* (*natto*) starters share the same origin and that IS*4Bsu1* is not always associated with γPGA production.

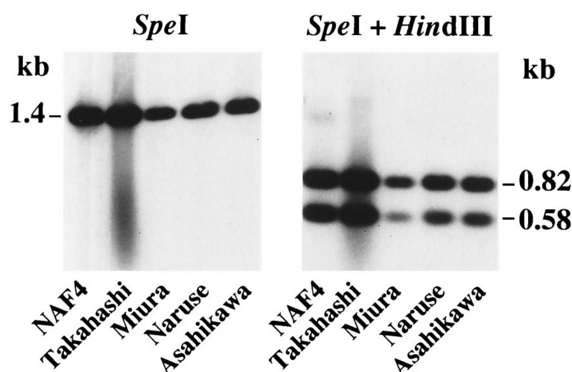

FIG. 4. Digestion of the IS*4Bsu1* sequences with restriction enzymes *Spe*I and *Hin*dIII. The indicated chromosomal DNAs of the *B. subtilis* (*natto*) strains were digested with *Spe*I and doubly with *Spe*I and *Hin*dIII, and the IS*4Bsu1* sequences were detected with probe 3 (Fig. 1B).

It should be noted that *B. subtilis* (*natto*) strains harbor two plasmids, pLS20 (pNAGL1) (21, 35) and pTA1015 (pUH1) (31, 35). As evident from the nucleotide sequence data, pTA1015 has no IS (31). Accordingly, we examined pLS20 by Southern blotting, but no IS*4Bsu1* was detected in this plasmid.

Spontaneous γPGA$^-$ mutation accompanies transposition of IS*4Bsu1*. IS*4Bsu1* appears to be a cause of spontaneous γPGA$^-$ mutations in *B. subtilis* (*natto*) strains. We next addressed the question whether the γPGA$^-$ phenotype is associated with transposition of IS*4Bsu1* but not with transposition of other unknown ISs. To obtain an answer to this question, we carried out Southern blot experiments with the *Pst*I-digested chromosomal DNAs from eight γPGA$^-$ mutants, M1 to M8, isolated from the aforementioned five independent cultures. All eight γPGA$^-$ mutants had a new positive band corresponding to a 20-kb *Pst*I fragment that contained the *comQXPA* operon, as in the case of NAF5 (data not shown).

***comP* has hot spots of IS*4Bsu1*.** Because all spontaneous γPGA$^-$ mutants had a new IS*4Bsu1* copy in the same *Pst*I fragment carrying the *comQXPA* operon, we next examined whether IS*4Bsu1* translocates into other genes, such as *comQ*, *comX*, and *comA*, which would also be required for γPGA synthesis. Southern blotting was done with the *Hin*dIII-digested mutant chromosomal DNAs, using the *Fsp*I-*Hin*dIII fragment containing the 5′ part of *comP* (probe 2) (Fig. 1A) as a probe. All mutant chromosomal DNA fragments produced two *Hin*dIII fragments (data not shown), in contrast to the parent NAF4 chromosomal DNA, which gave a single *Hin*dIII fragment of 4.4 kb as detected by probe 1 (Fig. 2). These results indicate that IS*4Bsu1* had transposed exclusively into *comP* in all mutants. Based on the sizes of the *Hin*dIII fragments as determined by probe 2 (Fig. 1A) and probe 4 (Fig. 1B), the location and direction of IS*4Bsu1* in the *comP* gene were established (Fig. 5A). In all cases, IS*4Bsu1* resided at one of six different positions in *comP*: nt 70, 270, 360, 400, 440, 580, or 590 (Fig. 5A). Finally, DNA sequencing was used to determine the exact location and the target sequences of IS*4Bsu1*. Alignment of the seven determined target sequences revealed the possible 9-bp consensus sequence 5′-ATNTWWWWW-3′ (Fig. 5B), where W indicates an A or a T. This consensus sequence apparently lacks the palindrome structure that exists in the target sites of IS*10* (5′-NGCTNAGCN-3′) (2) and IS*231A* [5′-GGG(N)$_5$CCC-3′] (14). In spite of the fact that the target sequences of IS*4Bsu1* are not well conserved, IS*4Bsu1* has hot spots in the region between nt 65 and 600 of the *comP* gene. In addition to the target sequence itself, the neighboring 6-bp sequences, which comprise two turns of B-DNA (21 bp) with the target sequence or local DNA configuration of the target site, also make an important contribution to the target

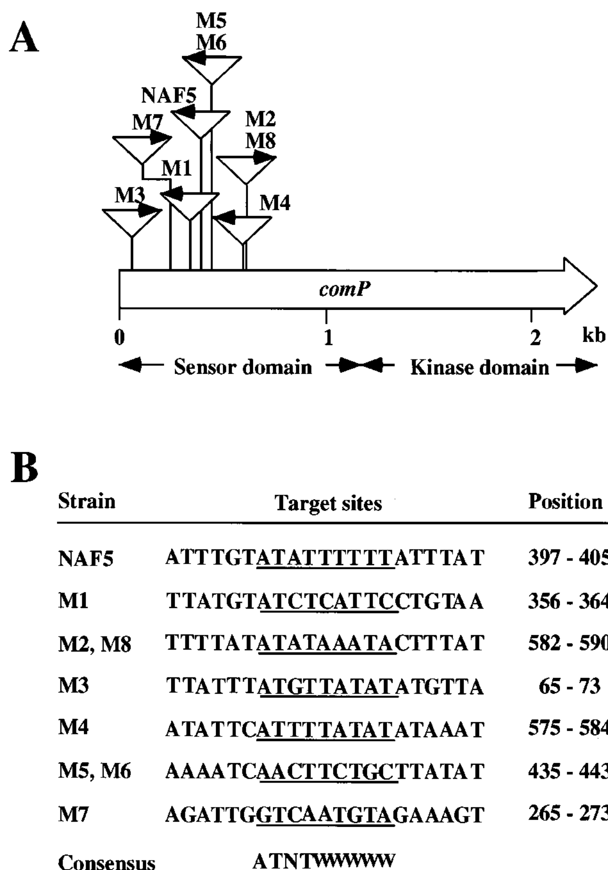

FIG. 5. Location and orientation of IS*4Bsu1* in *comP* of the spontaneous γPGA-negative mutants (A) and alignment of the IS*4Bsu1* target sequences. (B). (A) Arrows indicate the direction of *tnp* transcription. The names of strains having the relevant insertion are indicated above the arrows. (B) Nine-base-pair target sites are underlined. Positions refer to the *comP* nucleotide coordinates (accession no. AB031552). In the consensus sequence, N is A, C, G, or T, and W is A or T.

site selection (2, 14). Although the significant features of these flanking sequences for target site selection remain to be determined, we can point out that two turns of B-DNA including the IS*4Bsu1* target sites have a high AT content (72 to 95%) (Fig. 5B). Such a highly AT-rich stretch would facilitate unwinding of the target strands in a reaction joining donor and recipient strands.

Replicative transposition of IS*4Bsu1*. ISs transpose by either the conservative or replicative pathway, depending on the manner of the terminal strand cleavage mediated by Tpases (9, 10, 15, 17, 30). The replicative pathway involves single-strand transfer of ISs to target DNA to allow cointegration by subsequent replication or insertion of a second strand. IS*6* family members are known to use the replicative pathway (30). The conservative, or cut-and-paste, pathway involves excision of double or single strands of IS, which can be directly inserted into target DNA by Tpases (9, 10, 15, 17, 30). Members of the IS*4* family, including IS*4*, IS*10*, IS*50*, and IS*321A*, as well as members of the IS*3* family, have been demonstrated to preferentially or exclusively use a conservative mechanism of transposition (2, 14, 26, 30). Our present results clearly demonstrate the replicative mechanism of IS*4Bsu1*, which leaves the parent sequence at the original site to accumulate its copies in the chromosome of *B. subtilis* (*natto*) strains. IS*4Bsu1* thus represents an example of an IS*4* family member that uses the replicative mechanism, although the mechanism by which IS*4Bsu1* transposes into another replicon remains to be investigated.

ACKNOWLEDGMENTS

We thank S. Bron for his provision of pLS20 and T. Nishijo for his help with DNA sequencing.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Bender, J., and N. Kleckner.** 1992. Tn*10* insertion specificity is strongly dependent upon sequences immediately adjacent to the target-site consensus sequence. Proc. Natl. Acad. Sci. USA **89:**7996–8000.
3. **Berg, D. E., and M. M. Howe (ed.).** 1989. Mobile DNA. American Society for Microbiology, Washington, D.C.
4. **Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1474.
5. **Bron, S.** 1990. Plasmids, p. 75–174. *In* C. R. Harwood and S. M. Cutting (ed.), Molecular biological methods for *Bacillus*. John Wiley & Sons, West Sussex, United Kingdom.
6. **Camarena, L., S. Poggio, A. Campos, F. Bastarrachea, and A. Osorio.** 1998. An IS*4* insertion at the *glnA* control region of *Escherichia coli* creates a new promoter by providing the −35 region of its 3′-end. Plasmid **39:**41–47.
7. **Chesney, R. H., J. R. Scott, and D. Vapnek.** 1979. Integration of the plasmid prophages P1 and P7 into the chromosome of *Escherichia coli*. J. Mol. Biol. **130:**161–173.
8. **Coucheron, D. H.** 1991. An *Acetobacter xylinum* insertion sequence element associated with inactivation of cellulose production. J. Bacteriol. **173:**5723–5731.
9. **Craig, N. L.** 1995. Unity in transposition reactions. Science **270:**253–254.
10. **Craig, N. L.** 1997. Target site selection in transposition. Annu. Rev. Biochem. **66:**437–474.
11. **D'Souza, C., M. M. Nakano, and P. Zuber.** 1994. Identification of *comS*, a gene of the *srfA* operon that regulates the establishment of genetic competence in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **91:**9397–9401.
12. **Dubnau, D.** 1993. Genetic exchange and homologous recombination, p. 555–584. *In* A. L. Sonenshein, J. A. Hoch, and R. Losick (ed.), *Bacillus subtilis* and other gram-positive bacteria. American Society for Microbiology, Washington, D.C.
13. **Hall, B. G.** 1998. Activation of the *bgl* operon by adaptive mutation. Mol. Biol. Evol. **15:**1–5.
14. **Hallet, B., R. Rezsöhazy, J. Mahillon, and J. Delcour.** 1994. IS*231A* insertion specificity: consensus sequence and DNA bending at the target site. Mol. Microbiol. **14:**131–139.
15. **Hallet, B., and D. J. Sherratt.** 1997. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. FEMS Microbiol. Rev. **21:**157–178.
16. **Hamoen, L. W., H. Eshuis, J. Jongbloed, G. Venema, and D. Van Sinderen.** 1995. A small gene, designated *comS*, locates within the coding region of the fourth amino-acid-activation domain of *srfA*, required for competence development in *Bacillus subtilis*. Mol. Microbiol. **15:**55–63.
17. **Haniford, D. B., and G. Chaconas.** 1992. Mechanistic aspects of DNA transposition. Curr. Opin. Genet. Dev. **2:**698–704.
18. **Inoue, H., H. Nojima, and H. Okayama.** 1990. High efficiency transformation of *Escherichia coli* with plasmids. Gene **96:**23–28.
19. **Kallastu, A., R. Horak, and M. Kivisaar.** 1998. Identification and characterization of IS*1411*, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*. J. Bacteriol. **180:**5306–5312.
20. **Kasahara, Y., S. Nakai, and N. Ogasawara.** 1997. Sequence analysis of the 36-kb region between *gntZ* and *trnY* genes of *Bacillus subtilis* genome. DNA Res. **4:**155–159.
21. **Koehler, T. M., and C. B. Thorne.** 1987. *Bacillus subtilis* (*natto*) plasmid pLS20 mediates interspecies plasmid transfer. J. Bacteriol. **169:**5271–5278.
22. **Kondo, K., and S. Horinouchi.** 1997. Characterization of an insertion sequence, IS*12528*, from *Gluconobacter suboxydans*. Appl. Environ. Microbiol. **63:**1139–1142.
23. **Kondo, K., and S. Horinouchi.** 1997. A new insertion sequence IS*1452* from *Acetobacter pasteurianus*. Microbiology **143:**539–546.
24. **Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, and A. Danchin.** 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature **390:**249–256.
25. **Lazazzera, B. A., T. Palmer, J. Quisel, and A. D. Grossman.** 1999. Cell

density control of gene expression and development, p. 27–46. *In* G. M. Dunny and S. C. Winans (ed.), Cell-cell signaling in bacteria. ASM Press, Washington, D.C.

26. **Léonard, C., and J. Mahillon.** 1998. IS*231A* transposition: conservative versus replicative pathway. Res. Microbiol. **149:**549–555.
27. **Lewis, L. A., D. Lewis, V. Persaud, S. Gopaul, and B. Turner.** 1994. Transposition of IS*2* into the *hemB* gene of *Escherichia coli* K-12. J. Bacteriol. **176:**2114–2120.
28. **Lindsay, J. A., A. Ruzin, H. F. Ross, N. Kurepina, and R. P. Novick.** 1998. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. Mol. Microbiol. **29:**527–543.
29. **Magnuson, R., J. Solomon, and A. D. Grossman.** 1994. Biochemical and genetic characterization of a competence pheromone from *B. subtilis*. Cell **77:**207–216.
30. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725–774.
31. **Meijer, W. J., G. B. Wisman, P. Terpstra, P. B. Thorsted, C. M. Thomas, S. Holsappel, G. Venema, and S. Bron.** 1998. Rolling-circle plasmids from *Bacillus subtilis*: complete nucleotide sequences and analyses of genes of pTA1015, pTA1040, pTA1050 and pTA1060, and comparisons with related plasmids from gram-positive bacteria. FEMS Microbiol. Rev. **21:**337–368.
32. **Msadek, T., F. Kunst, A. Klier, and G. Rapoport.** 1991. DegS-DegU and ComP-ComA modulator-effector pairs control expression of *Bacillus subtilis* pleiotropic regulatory gene *degQ*. J. Bacteriol. **173:**2366–2377.
33. **Mueller, J. P., G. Bukusoglu, and A. Sonenshein.** 1992. Transcriptional regulation of *Bacillus subtilis* glucose starvation-inducible genes: control of *gsiA* by the ComP-ComA signal transduction system. J. Bacteriol. **174:**4361–4373.
34. **Nagai, T., and Y. Itoh.** 1997. Characterization of a generalized transducing phage of poly-γ-glutamic acid-producing *Bacillus subtilis* and its application for analysis of Tn*917*-LTV1 insertional mutants defective in poly-γ-glutamic acid-production. Appl. Environ. Microbiol. **63:**4087–4089.
35. **Nagai, T., K. Koguchi, and Y. Itoh.** 1997. Chemical analysis of poly-γ-glutamic acid produced by plasmid-free *Bacillus subtilis* (*natto*): evidence that plasmids are not involved in poly-γ-glutamic acid production. J. Gen. Appl. Microbiol. **43:**139–143.
36. **Piazza, F., P. Tortosa, and D. Dabnau.** 1999. Mutation analysis and membrane topology of ComP, a quorum-sensing histidine kinase of *Bacillus subtilis* controlling competence development. J. Bacteriol. **181:**4540–4548.
37. **Plasterk, R. H.** 1993. Molecular mechanisms of transposition and its control. Cell **74:**781–786.
38. **Rezsöhazy, R., B. Hallet, J. Delcour, and J. Mahillon.** 1993. The IS*4* family of insertion sequences: evidence for a conserved transposase motif. Mol. Microbiol. **9:**1283–1295.
39. **Roggiani, M., and D. Dabunau.** 1993. ComA, a phosphorylated response regulator protein of *Bacillus subtilis*, binds to the promoter regions of *srfA*. J. Bacteriol. **175:**3182–3187.
40. **Salyers, A. A., N. B. Shoemaker, A. M. Stevens, and L. Y. Li.** 1995. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol. Rev. **59:**579–590.
41. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. **Stock, J. B., M. G. Surette, M. Levit, and P. Park.** 1995. Two-component signal transduction systems: structure-function relationships and mechanisms of catalysis, p. 25–51. *In* J. A. Hock and T. J. Silhavy (ed.). Two-component signal transduction. ASM Press, Washington, D.C.
43. **Takemura, H., S. Horinouchi, and T. Beppu.** 1991. Novel insertion sequence IS*1380* from *Acetobacter pasteurianus* is involved in loss of ethanol-oxidizing ability. J. Bacteriol. **173:**7070–7076.
44. **Tan, H. M.** 1999. Bacterial catabolic transposon. Appl. Microbiol. Biotechnol. **51:**1–12.
45. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4674–4680.
46. **Thorne, C. B., C. G. Gómez, H. E. Noyes, and R. D. Housewright.** 1954. Production of glutamyl polypeptide by *Bacillus subtilis*. J. Bacteriol. **68:**307–315.
47. **Ueda, S.** 1989. Industrial application of *B. subtilis*, p. 143–161. *In* B. Maruo and H. Yoshikawa (ed.), *Bacillus subtilis*: molecular biology and industrial application. Kodansha Ltd., Tokyo, Japan.
48. **Vieira, J., and J. Messing.** 1987. Production of a single-strand plasmid DNA. Methods Enzymol. **153:**3–11.
49. **Weinrauch, Y., R. Penchev, E. Dubnau, I. Smith, and D. Dubnau.** 1990. A *Bacillus subtilis* regulatory gene product for genetic competence and sporulation resembles sensor protein members of the bacterial two-component signal-transduction system. Genes Dev. **4:**860–872.