APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 1998, p. 53–61 Vol. 64, No. 1
0099-2240/98/$04.00+0

IS*D1*, an Insertion Element from the Sulfate-Reducing Bacterium *Desulfovibrio vulgaris* Hildenborough: Structure, Transposition, and Distribution

RONGDIAN FU† AND GERRIT VOORDOUW*

Department of Biological Sciences, The University of Calgary, Calgary, Alberta T2N 1N4, Canada

Received 8 July 1997/Accepted 12 October 1997

Insertion element IS*D1*, discovered when its transposition caused the insertional inactivation of an introduced *sacB* gene, is present in two copies in the genome of *Desulfovibrio vulgaris* Hildenborough. Southern blot analysis indicated at least two insertion sites in the *sacB* gene. Cloning and sequencing of a transposed copy of IS*D1* indicated a length of 1,200 bp with a pair of 44-bp imperfect inverted repeats at the ends, flanked by a direct repeat of the 4-bp target sequence. AAGG and AATT were found to function as target sequences. IS*D1* encodes a transposase from two overlapping open reading frames by programmed translational frameshifting at an A_6G shifty codon motif. Sequence comparison showed that IS*D1* belongs to the IS*3* family. Isolation and analysis of the chromosomal copies, IS*D1*-A and IS*D1*-B, by PCR and sequencing indicated that these are not flanked by direct repeats. IS*D1*-A is inserted in a region of the chromosome containing the *gapdh-pgk* genes (encoding glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase). Active transposition to other loci in the genome was demonstrated, offering the potential of a new tool for gene cloning and mutagenesis. IS*D1* is the first transposable element described for the sulfate reducers, a large and environmentally important group of bacteria. The distribution of IS*D1* in genomes of sulfate-reducing bacteria is limited. A single copy is present in the genome of *D. desulfuricans* Norway.

Bacterial insertion sequences (ISs) are mobile genetic elements of 0.7 to 2 kb that code only for functions necessary for their transposition (11, 14). The majority contains imperfect inverted repeats (IRs) of up to 46 bp at the ends and produce target site duplication upon insertion. The transposition of an IS element can have different genetic consequences, including insertional mutation of a gene and activation or inactivation of nearby genes (1, 14). IS elements are found either alone or at the ends of composite transposable elements (transposons). In spite of their diversity, sequence analysis has revealed the existence of several large families, e.g., IS*3* (9, 32) and IS*4* (26). The IS*3* family includes members isolated from both gram-negative and gram-positive bacteria. All are organized similarly, with two overlapping open reading frames (*orfA* and *orfB*) and a string of purines (A_6G or similar motif containing a run of adenines) in the overlap region, usually followed by a potential stem-loop structure (3, 24). This unusual arrangement has been shown to produce an OrfA-OrfB fusion protein with transposase activity by a programmed −1 translational frameshift at the A_6G-like site in IS*911* (21), IS*150* (39), and IS*3* (34). OrfA, which is not highly conserved, contains a potential helix-turn-helix motif, possibly involved in binding to the terminal IRs of the cognate IS element (24). OrfB is more conserved among the family and contains a D-(1)-G-(33)-E or D-(35)-E motif which is also shared by the retroviral/retrotransposon integrases (7, 9, 18).

Bacteria of the genus *Desulfovibrio* are gram-negative sulfate-reducing anaerobes, for which the genetics and molecular biology have been relatively well studied (29, 41–44). In the course of a gene replacement mutagenesis study of *dcrA* of *Desulfovibrio vulgaris* Hildenborough, encoding an oxygen-sensing protein, using the *Bacillus subtilis sacB* gene as a counterselection marker, we obtained mutants which were sucrose resistant by insertion of a 1.2-kb DNA element into *sacB* (10). The cloning and characterization of this element, which we named IS*D1*, its homology to the IS*3* family, and its distribution among *Desulfovibrio* spp. are reported here.

MATERIALS AND METHODS

Bacterial strains, phages, plasmids, and growth conditions. Bacteria, phages, and plasmids used in this study are listed in Table 1. *D. vulgaris* Hildenborough and its derivative strains and *Escherichia coli* TG2 were grown in medium C and TY medium, respectively, as described previously (10). Chromosomal DNA samples from the following bacteria were available in the laboratory for screening of the distribution of IS*D1* by Southern blot analysis: *D. vulgaris* NCIMB8399, *D. vulgaris* Miyazaki, *D. vulgaris* subsp. *oxamicus*, *D. desulfuricans* Norway, *D. desulfuricans* NCIMB8407, *D. gigas*, *D. africanus*, *D. salexigens* NCIMB8365, *D. baarsii* DSM2075, *Desulfobotulus sapovorans*, *Desulfobulbus propionicus*, *Desulfococcus multivorans*, *Desulfosarcina variabilis*, *Desulfotomaculum ruminis* NCIMB8452, *Desulfotomaculum ruminis* ATCC 23192, *Shewanella putrefaciens*, and *Clostridium* sp. These are not listed in Table 1.

Southern blot analysis of *D. vulgaris* F1SR strains. Chromosomal DNAs of *D. vulgaris* F1SR strains were isolated from 5-ml cultures by a minipreparation protocol (10). The DNAs were restricted with *Bam*HI and *Hin*dIII, and the digests were fractionated on an agarose gel and blotted onto a Hybond-N nylon membrane (Amersham). The blot was probed with ^{32}P-labeled *sacB* DNA, obtained as a 2.4-kb *Xba*I fragment from pMOB2 (33) to characterize the insertion mutation.

Cloning of IS*D1*. Chromosomal DNA from *D. vulgaris* F1SR12, containing a putative IS element inserted into the *sacB* gene, was restricted with *Pst*I. Restriction fragments of 3 to 4 kb were isolated from a low-melting-temperature agarose gel and ligated into the *Pst*I site of plasmid pNOT19. The ligation mix was used to transform *E. coli* TG2. Ampicillin-resistant colonies were screened by hybridization with the *sacB* probe. Plasmids isolated from positive clones were mapped to verify the presence of the putative insertion element. One of these was named pHIS1. Fragments of the *sacB* gene containing the insertion element were gel isolated, ^{32}P labeled, and used to probe a Southern blot of chromosomal DNAs to verify the *D. vulgaris* origin of the insertion element. The same probe

* Corresponding author. Mailing address: Department of Biological Sciences, The University of Calgary, 2500 University Dr. NW, Calgary, Alberta T2N 1N4, Canada. Phone: (403) 220-6388. Fax: (403) 289-9311. E-mail: voordouw@acs.ucalgary.ca.

† Present address: The Molecular Sciences Institute, La Jolla, CA 92037.

TABLE 1. Bacterial strains, plasmids, and phages used for IS*D1* characterization

Bacterial strain, plasmid, or phage	Genotype, comments	Reference or source
D. vulgaris		
Hildenborough	NCIMB 8303; wild type, source of the IS*D1* element	23
F1	Derivative of Hildenborough with pΔDcrA2CTB integrated into the chromosome; Km^r Cm^r Suc^s	10
F1SR	All sucrose-resistant derivatives of F1 with *sacB* gene mutated; Km^r Cm^r Suc^r	This study
F1SR12	One of the F1SR strains containing an IS*D1* insertion within the *sacB* gene	This study
F1SR11	One of the F1SR strains containing an IS*D1* insertion within the *sacB* gene and two additional newly transposed copies on the chromosome	This study
E. coli TG2	Δ(*lac-pro*) *supE thi hsdM hsdR recA* F′ (*traD36 proAB*$^+$ *lacZ*ΔM15 I^q)	30
Plasmids and phages		
pNOT19	Cloning vector derived from pUC19; Ap^r	33
pNEB193	Cloning vector derived from pUC19; Ap^r	New England Biolabs
pHIS1	pNOT19 with a 3.8-kb *Pst*I fragment from *D. vulgaris* F1SR12 containing the *sacB* gene with an IS*D1* insertion; Ap^r	This study
*Dv*H λ library	Ordered genomic λ library of *D. vulgaris* Hildenborough	5, 40
λ-G724	Clone from the *D. vulgaris* λ library, with a 18-kb insert containing IS*D1*	

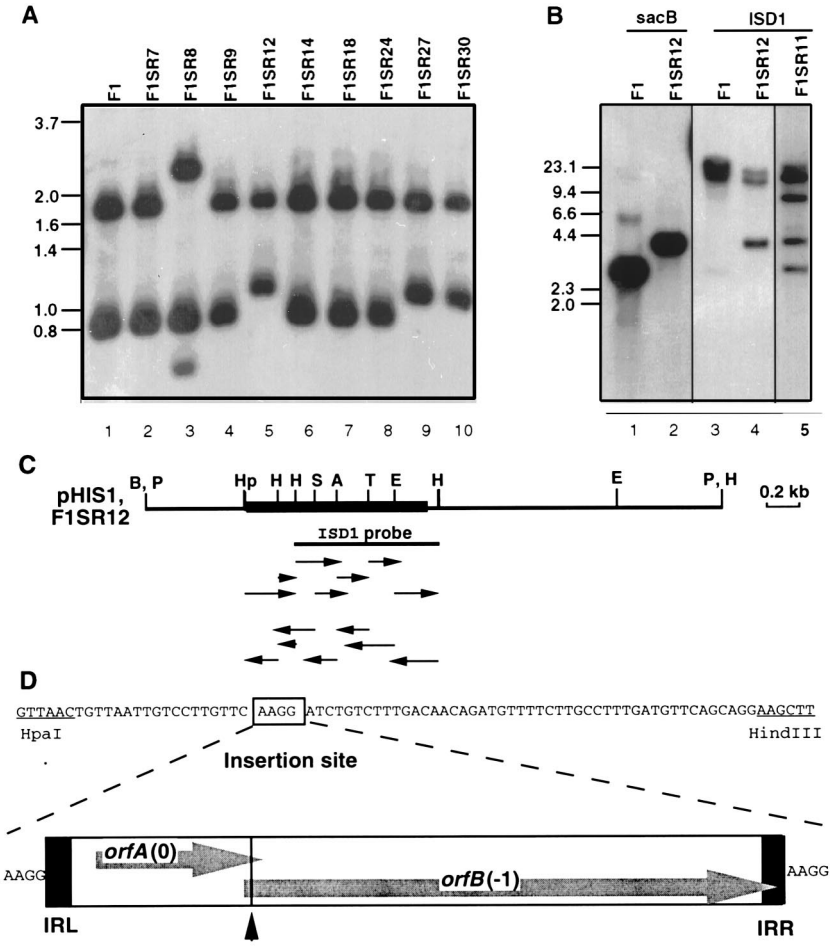


FIG. 1. Cloning and sequencing of IS*D1* from a *D. vulgaris* F1SR strain. (A) Southern blot analysis of *D. vulgaris* FISR strains. Chromosomal DNAs were restricted with *Bam*HI and *Hin*dIII, and the blot was hybridized with the *sacB* probe. (B) Southern hybridization of F1, F1SR11, and F1SR12 strains restricted with *Pst*I, using the *sacB* probe (lanes 1 and 2) and the IS*D1* probe (lanes 3 to 5). In both panels, the size markers are denoted in kilobases on the left. (C) Restriction map of a 3.8-kb fragment of the *sacBR* genes (thin line) with the IS*D1* insertion (thick line) as cloned in plasmid pHIS1. Restriction sites: A, *Acc*I; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I; P, *Pst*I; S, *Sac*I; T, *Taq*I. The 0.7-kb *Hin*dIII fragment used as the IS*D1* probe in panel B is indicated by the solid line below the map. The arrows indicate the sequencing strategy. (D) Nucleotide sequence of the *Hpa*I-*Hin*dIII region of the *sacB* gene from strain F1SR12 indicating the site of IS*D1* insertion. A survey of the structure of IS*D1*, which is flanked by a duplicated target sequence (AAGG) and contains IRL and IRR, is also shown. OrfA and OrfB in reading frames 0 and −1, respectively, are separated by a shifty codon motif (▲) as discussed in the text.

was also used to probe an ordered *D. vulgaris* Hildenborough genomic λ library (5, 40) to identify λ clones containing the native IS*D1* element.

Nucleotide sequence determination. The insertion element DNA cloned in plasmid pHIS1 was mapped and subcloned to facilitate nucleotide sequence determination. The recombinant plasmids were used directly for double-stranded DNA sequencing by the dideoxy-chain termination method using a T7 sequencing kit (Pharmacia). A cycling sequencing kit (Pharmacia) was used for sequencing the flanking regions of the native IS*D1* elements obtained by PCR amplification of wild-type *D. vulgaris* DNA, using the PCR primers described below. The sequencing gel autoradiograms were read manually, and the sequence data were assembled into a contiguous sequence by using the Fragment Assembly system of the Genetics Computer Group (GCG) package (version 8.0.1-UNIX).

Cloning of the flanking regions of native IS*D1* by PCR. After determination of the complete nucleotide sequence of the cloned insertion element, two outward-pointing oligonucleotide primers, GCACTCCATGAGGCAATC (P101) and AGTACAACGAGGAACGAC (P102), complementary to sequences near the two ends of the element were synthesized. Chromosomal DNA from wild-type *D. vulgaris* was digested with *Hin*cII or *Cla*I, and the fragments produced were circularized with T4 DNA ligase and then used as templates for PCRs to amplify the flanking sequences. These PCRs were conducted in 100 μl of buffer (50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl [pH 9.0]) containing 2 mM deoxynucleoside triphosphates 0.4 μM primers, and 2.5 U of *Taq* DNA polymerase, with 5 min of preheating at 94°C followed by 30 cycles of 45 s at 94°C, 30 s at 50°C, and 4 min at 72°C. The last polymerization at 72°C was extended to 10 min. The PCR product was cloned into the *Pme*I site of plasmid pNEB193, or used directly, for sequence determination with a cycling sequencing kit (Pharmacia) as described above.

Nucleotide sequence analysis. The nucleotide sequence of IS*D1* and amino acid sequences deduced from open reading frames and longer than 50 amino acids (aa) were used to search for homologous sequences in the GenBank and protein databases, using the BLAST program in the network server at blast @ncbi.nlm.nih.gov. Homologous sequences were aligned by using the multiple-sequence alignment program PILEUP of the GCG package by varying parameters for gap creation and extension penalties until a satisfactory alignment was achieved. The output file of aligned multiple sequences was used to deduce a matrix of the pairwise evolutionary distances with the DISTANCES program, which uses the Kimura protein distance correction method. Phylogenetic trees were constructed from the distance matrix by using the GROWTREE program of the GCG package (12).

Distribution of IS*D1*-like elements. Chromosomal DNAs from 18 different bacteria were restricted with *Pvu*II, which has no site in the cloned IS*D1*, and blotted onto Hybond-N membranes. The blots were probed with ^{32}P-labeled IS*D1* DNA under highly stringent conditions. The final posthybridization washing step was in 1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate at 68°C for 30 min.

Nucleotide sequence accession number. The complete nucleotide sequence of IS*D1* described in this report has been deposited in GenBank under accession no. AF034211.

RESULTS

Cloning of IS*D1* from an insertionally inactivated *sacB* gene. In earlier work, Southern analysis using the *sacB* gene as a probe had shown that 6 of 16 *D. vulgaris* F1SR strains contain a 1.2-kb insertion in the *sacB* gene (10). We found two insertion patterns, one represented by F1SR8 (Fig. 1A, lane 3) and one represented by F1SR12, F1SR27, and F1SR30 (Fig. 1A, lanes 5, 9, and 10). Other sucrose-resistant mutants of *D. vulgaris* F1 (Fig. 1A, F1SR7, F1SR9, F1SR14, F1SR18, and F1SR24) had a hybridization pattern unchanged from that of *D. vulgaris* F1, indicating that sucrose resistance was not caused by *sacB* inactivation through IS element insertion.

A 3.8-kb *Pst*I fragment containing the 1.2-kb IS element was identified in *D. vulgaris* F1SR12 (Fig. 1B, lane 2) and was cloned into the *Pst*I site of pNOT19 to give plasmid pHIS1 (see Materials and Methods). A 0.7-kb *Hin*dIII fragment located mostly within the element (Fig. 1C, IS*D1* probe) was isolated. Use in probing a Southern blot of *Pst*I-restricted chromosomal DNA showed a 3.8-kb fragment, corresponding to the IS element-inactivated *sacB* gene, as well as two additional *Pst*I-fragments of 15 and 23 kb for *D. vulgaris* F1SR12 (Fig. 1B, lane 4). The latter two were also found in *D. vulgaris* F1 (Fig. 1B, lane 3) and in wild-type *D. vulgaris* Hildenborough (not shown), indicating that IS*D1* is present in two identical or very similar copies in the *D. vulgaris* genome.

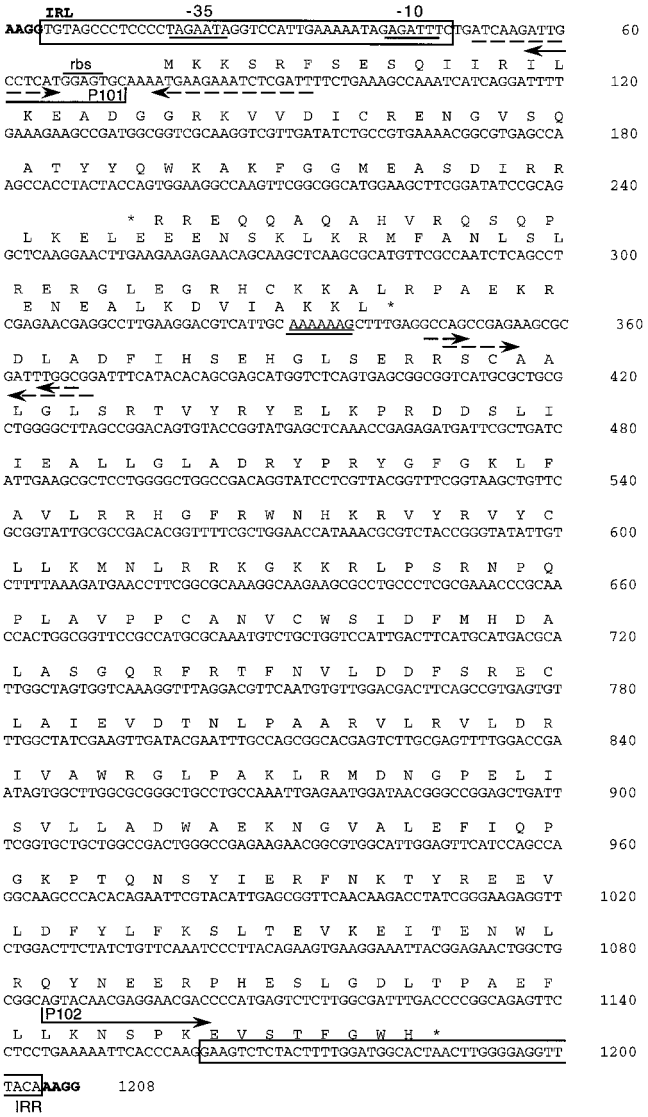


FIG. 2. Nucleotide sequence of IS*D1*. The flanking duplicated target site (AAGG) is indicated in bold. IRL and IRR are boxed. The two open reading frames that encode the cognate transposase (*orfA* [nt 77 to 343] and *orfB* [nt 256 to 1188]) are indicated by translation into protein. *orfA* gene is preceded by a ribosome-binding site (rbs). As in other IS elements, the exact translational start codon of *orfB* is unclear. A shifty codon motif (nt 331 to 337), a putative promoter (35 [TAGAAT, nt 18 to 23] and −10 [GAGATT, nt 41 to 46]), several IRs (dashed arrows), and the primers P101 and P102 are indicated.

Nucleotide sequence of IS*D1*. The nucleotide sequence of the 3.8-kb *Pst*I insert of pHIS1 was determined by the strategy shown in Fig. 1C. Parts of the sequence, determined for both strands, are shown in Fig. 1D and 2. The actual IS*D1* element is 1,200 bp long (Fig. 2, nucleotides [nt] 5 to 1204) with a GC content of 52.2%, considerably lower than that of the *D. vulgaris* genome (62 to 65%). The sequence has a pair of 44-bp imperfect IRs with 11 mismatches at the left (IRL) and right (IRR) ends. These are flanked by a pair of 4-bp direct repeats of the target site sequence (AAGG), indicating that transposition of IS*D1* causes a 4-bp target site duplication. IS*D1* carries two overlapping open reading frames, *orfA* and *orfB*, on its coding strand. The presence of only one putative σ^{70}-like promoter within the IRL region (Fig. 2, −35 [TAGAAT at nt 18 to 23] and −10 [GAGATT at nt 41 to 46]) indicates that *orfA*

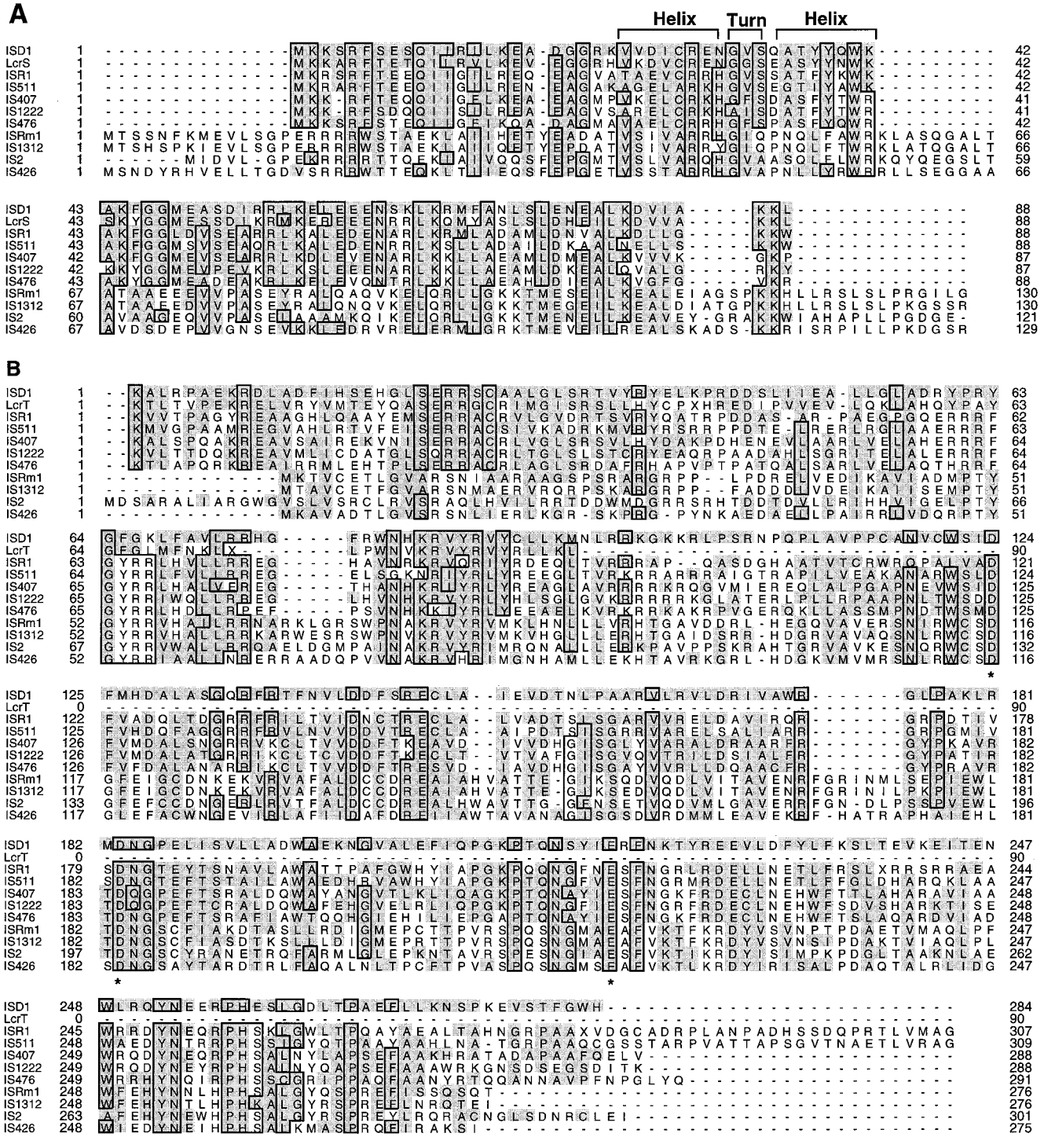


FIG. 3. Amino acid sequence comparisons of IS transposases. The putative transposase encoded by *orfA* and *orfB* of IS*D1* is compared with its counterparts in the IS*3* family and with proteins encoded by the *lcr* operon. (A) Sequence alignment of OrfA sequences and LcrS. (B) Sequence alignment of OrfB sequences and LcrT. Identical residues in 6 or more of the 11 sequences are boxed, and residues homologous to those in IS*D1* are shaded. The DNA binding helix-turn-helix motif in OrfA and the D,G-(33)-E motif in OrfB are also indicated by overlining and asterisks, respectively.

and *orfB* are cotranscribed. A BLAST search of the nonredundant protein database revealed that *orfA* and *orfB* encode a transposase of the IS*3* family (24). A multiple sequence alignment with transposases of this family, generated by using the PILEUP program of the GCG package, is shown in Fig. 3.

Sequence of an additional target site. IS*D1* is present in two copies in the *D. vulgaris* chromosome, to which we will refer as IS*D1*-A and IS*D1*-B. One of the F1SR strains, F1SR11, was unusual in having three additional IS*D1* copies on the chromosome; one was inserted in the *sacB* gene to give the 3.8-kb

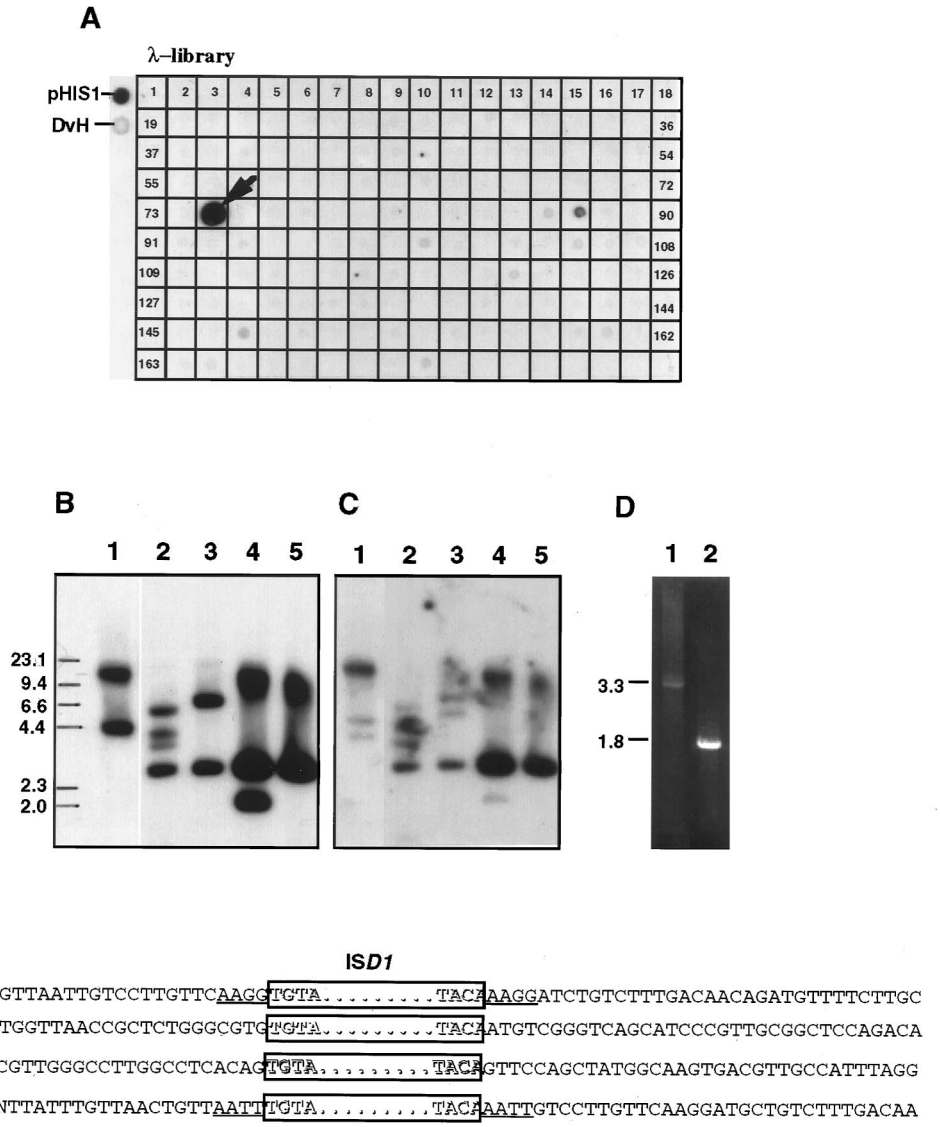
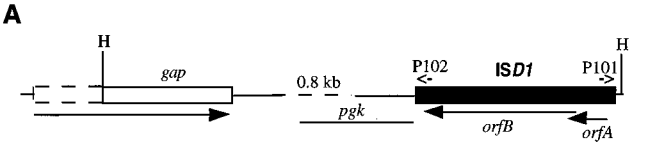


FIG. 4. Cloning and sequencing of the flanking regions of native IS*D1* from the *D. vulgaris* Hildenborough genome. (A) Dot blot hybridization of an ordered genomic library with the 0.7-kb IS*D1* probe indicated in Fig. 1C. pHIS1 DNA (pHIS1) and *D. vulgaris* Hildenborough chromosomal DNA (DvH) served as positive controls. Positive clone 75 (λ-G724) is indicated with an arrow. (B) A Southern blot of wild-type *D. vulgaris* chromosomal DNA restricted with *Cla*I (lane 1), *Acc*I (lane 2), and *Hin*cII (lane 3) and λ-G724 restricted with *Acc*I (lane 4) and *Hin*cII (lane 5) was hybridized with the 0.7-kb IS*D1* probe. The size markers are indicated in kilobases on the left. (C) Southern blot as in panel B but probed with the 1.8-kb PCR product containing the flanking regions of IS*D1*-A (Fig. 1D, lane 2). (D) PCR products amplified from *Cla*I (lane 1)- or *Hin*cII (lane 2)-restricted and circularized chromosomal DNA of wild-type *D. vulgaris* Hildenborough, using the P101-P102 primer pair. These contain the flanking regions of IS*D1*-A and IS*D1*-B, respectively. (E) Comparison of flanking region sequences for IS*D1* inserted into the *sacB* gene in *D. vulgaris* F1SR12 (i), of the two native copies IS*D1*-A (ii) and IS*D1*-B (iii), and of the one transposed in *D. vulgaris* F1SR11 (iv). The IS*D1* element is boxed. The duplicated target sites flanking transposed IS*D1* are underlined.

*Pst*I fragment, and the other two were present in 3- and 7-kb *Pst*I fragments, respectively (Fig. 1B, lane 5). *Pst*I fragments of 3 kb were isolated by agarose gel electrophoresis and circularized by ligation. PCR of the product of the ligation reaction with the pair of outward primers (Fig. 2, P101 and P102) gave a 1.9-kb PCR product. Sequence analysis of this PCR product indicated that the target site sequence of the IS*D1* transposition was AATT in this case (Fig. 4E, line iv).

Characterization of chromosomal flanking regions. Digestion of the genome with *Hin*cII, Southern blotting, and hybridization of the blot with the 0.7-kb IS*D1* probe indicated hybridizing fragments of 3 and 7 kb, corresponding to IS*D1*-A and IS*D1*-B, respectively (Fig. 4B, lane 3). Screening an ordered λ library, estimated to contain ca. 70% of the *D. vulgaris* genome (5), with the 0.7-kb IS*D1* probe gave one positive λ clone (Fig. 4A, λ-G724). Detection of the 3-kb *Hin*cII fragment in this clone (Fig. 4B, lane 5) indicated it to contain IS*D1*-A. To retrieve IS*D1*-B, *D. vulgaris* DNA was digested with *Hin*cII and circularized by ligation, and the product of the ligation reaction was used for PCR with the pair of outward primers. Agarose gel electrophoresis indicated that only a 1.8-kb PCR product, derived from the 3-kb *Hin*cII fragment, was obtained (Fig. 4D, lane 2). This PCR product hybridized to the 3-kb *Hin*cII fragments obtained from both genomic DNA and λ-G724 samples (Fig. 4C, lanes 3 and 5). The IS*D1*-B containing 7-kb *Hin*cII fragment was expected to give rise to a 5.9-kb PCR



B

```
            101                                                     150
G3P1_ECOLI  ltdetarkhi tagakkvvmt gpskdntpmf vkgaNfdkYa ..gqdIVSNA
G3P1_SALTY  ltdetarkhi tagakkvvlt gpskdntpmf vkgaNfdkYe ..gqdIVSNA
 G3P_BACST  tkredaakhl eagakkviis apakvenitv vmgvNqdkYd pkahhVISNA
 G3P-DVULG  .......... .......... .......... ...vNhhaYd pqkhhVVSNA
 Consensus  ---------- ---------- ---------- ----N---Y- ------VSNA

            151                                                     200
G3P1_ECOLI  SCTTNCLAPl AkVindnFGI ieGlMTTVHa tTatQktvDg PshkdwRggR
G3P1_SALTY  SCTTNCLAPl AkVindnFGI ieGlMTTVHa tTatQktvDg PshkdwRggR
 G3P_BACST  SCTTNCLAPf AkVlhqeFGI vrGmMTTVHs yTnnQrilDl PthkdlRgaR
 G3P-DVULG  SCTTNCLAPv AlIiqekYGI atGaMTTVHa yTndQrilDl PhkdlrRxtR
 Consensus  SCTTNCLAP- A-V----FGI --G-MTTVH- -T--Q---D- P-----R--R

            201                                                     250
G3P1_ECOLI  gAsqnIIPss TGAAkAVgkV lPELnGKLtG m.AfRVPtPn VSVVDLtvrL
G3P1_SALTY  gAsqnIIPss TGAAkAVgkV lPELnGKLtG m.AfRVPtPn VSVVDLtvrL
 G3P_BACST  aAaesIIPtt TGAAkAValV lPELkGKLnG m.AmRVPtPn VSVVDLvaeL
 G3P-DVULG  aAglnIIPts TGAAqAVakV iPELkGKFtG ysAcRVPiPt VSVVDFtaiL
 Consensus  -A---IIP-- TGAA-AV--V -PEL-GKL-G --A-RVP-P- VSVVDL---L

            251                                                     300
G3P1_ECOLI  EkaaTyEqik AavKaAAeGe MkGVLgYtEd dvVStdFnge vctSvfDAka
G3P1_SALTY  EkaaTyEqik AavKaAAeGe MkGVLgYtEd dvVStdFnge vctSvfDAka
 G3P_BACST  EkevTvEevn AalKaAAeGe LkGILaYsEe plVSrnYngs tvsStiDAls
 G3P-DVULG  ErdaTtEglr AdlKeAAqGr LsGILaFsEk plVSsdFkxx phsSxvEAey
 Consensus  E---T-E--- A--K-AA-G- --G-L-Y-E- --VS--F--- ---S--DA--

            301                                     337
G3P1_ECOLI  gialndnFvK lVswydnetg ysnkvldlia hisk...
G3P1_SALTY  gialndnFvK lV..... ..........  .....
 G3P_BACST  tmvidgkMvK vVswydnetg yshrvvdlaa yinakgl
 G3P-DVULG  ttxqdgrLaK vVawydnewg yscrladlim lmqekgl
 Consensus  -------F-K -V-------- ---------- -------
```

FIG. 5. Position of ISD1-A relative to the *gap* gene of *D. vulgaris* Hildenborough. (A) Physical map showing the 3′ end of *gap* (open box) and IS*D1*-A (filled box). The gene orientation (arrows), *Hinc*II sites (H), and locations of primers P101 and P102, used for PCR and sequencing, are indicated. The sequence of the region represented by the dashed line was not determined. The location of a putative *pgk* gene is also indicated. (B) Alignment of the amino acid sequence of *D. vulgaris* GAPDH (G3P_DVULG) with sequences from *B. stearothermophilus* (G3P_BACST), *E. coli* (G3P_ECOLI), and *Salmonella typhimurium* (G3P_SALTY). The numbering refers to the *E. coli* sequence. Residues conserved in all four species are indicated as the consensus sequence.

product. However, this was not observed, indicating bias toward amplification of the smaller product. Southern blot analysis of a *Cla*I digest of *D. vulgaris* DNA indicated that 4.4- and 15-kb fragments hybridized with the IS*D1* probe (Fig. 4B, lane 1). Hybridization with the 1.8-kb PCR product indicated that the 15-kb *Cla*I fragment contained IS*D1*-A (Fig. 4C, lane 1). PCR of a *Cla*I-restricted, circularized chromosomal DNA sample gave a 3.3-kb PCR product (Fig. 4D, lane 1), derived from the 4.4-kb fragment containing IS*D1*-B. The flanking sequences of IS*D1*-A and IS*D1*-B were determined by cycled sequencing of the two PCR products using primers P101 and P102, and the results are shown in Fig. 4E. No direct repeat sequences are flanking IS*D1*-A and IS*D1*-B in the *D. vulgaris* chromosome, implying that these flanking regions may have undergone secondary mutations following IS*D1* insertion.

IS*D1*-A is located downstream from *gap*. Sequence analysis of the 1.8-kb PCR product flanking IS*D1*-A and cloned in pNEB193 revealed the presence of a *Hinc*II site 16 nt upstream from IRL (Fig. 5A). The sequence data beyond this point therefore represented the sequence downstream from IS*D1*-A. Translation gave a truncated open reading frame coding for a polypeptide with homology to the C-terminal parts of glyceraldehyde-3-phosphate dehydrogenases (GAPDHs) from eubacteria, archaebacteria, and eukaryotes (Fig. 5B). The identity to the enzyme of *Bacillus stearothermophilus* (2, 35) is 58.5% in a 204-aa overlap. This result indicates that IS*D1*-A is located ca. 0.8 kb downstream from a *gap* gene, encoding GAPDH of *D. vulgaris* in a reverse orientation, as indicated in Fig. 5A.

Distribution of IS*D1* in other bacteria. Probing of a Southern blot of *Pvu*II-restricted genomic DNAs of 18 different bacteria with the 0.7-kb IS*D1* fragment gave the results shown in Fig. 6. Wild-type *D. vulgaris* and *D. vulgaris* F1 showed two hybridizing bands, while *D. vulgaris* F1SR12 showed an expected extra band. Among 10 other *Desulfovibrio* spp., only *D. desulfuricans* Norway showed a strongly hybridizing band, indicating the presence of an IS*D1*-like element in this organism. Weakly hybridizing bands were observed for *D. vulgaris* subsp. *oxamicus*, *D. africanus*, and *D. salexigens*, indicating the presence of distantly related IS elements in these organisms. Under these same conditions, no hybridization was found with DNA from sulfate-reducing bacteria other than *Desulfovibrio* spp. and with DNA from *S. putrefacieus*, *Clostridium* sp., and *E. coli* TG2.

DISCUSSION

Transposition of IS*D1*. IS*D1* was discovered when cells of *D. vulgaris* F1, which contains the *B. subtilis sacB* gene integrated into the chromosome at a site immediately downstream from the *dcrA* gene, were subjected to selection pressure by adding sucrose (10). Insertion of IS*D1* was in an AAGG target site of the *sacB* gene (Fig. 2). The second insertion site in the *sacB* gene (Fig. 1A, lane 3) has not been characterized. Sucrose selection did also cause IS*D1* to jump to other sites in the chromosome, and characterization of one of these indicated an AATT target site. Thus, IS*D1* transposase may have AAXX as recognition sequence, although more experiments are needed to completely define the insertion specificity. The occurrence of several simultaneous transposition events in the same cell subjected to *sacB* counterselection suggests that the transposition activity may have been induced by the *sacB* lethality in the presence of sucrose. This induction increased the frequency of random mutagenesis under stress conditions and thereby the probability of surviving the harsh selection pressures. The fact that IS*D1* can insert at more than one target site may make it a powerful natural mutagenesis tool for *D. vulgaris*. IS*D1* is not widely distributed in sulfate-reducing bacteria and may have been acquired by *D. vulgaris* Hildenborough sufficiently long ago to allow mutation of the direct repeats of the initial target sites (Fig. 5E).

Analysis and comparison of the IS*D1* sequence. The *orfA* gene of IS*D1* (nt 77 to 340) encodes a polypeptide of 88 aa, starting from an ATG codon preceded by a ribosome-binding site (Fig. 2, nt 67 to 70). The *orfB* gene overlaps with the 3′ end of *orfA*. Within the overlapping region, a putative translational frameshifting motif A_6G (nt 331 to 337) is located close to the 3′ end of *orfA*, followed by two potential stem-loop structures (Fig. 2, nt 345 to 370). This A_6G motif and the downstream secondary structures have been shown to promote the programmed translational frameshifting between *orfA* and *orfB* to produce an OrfAB fusion. This OrfAB fusion protein functions as the transposase in IS*911*, IS*150*, and IS*3* members of the IS*3* family (21, 22, 34, 39). The sequence homology suggests that this same mechanism is used by IS*D1* for regulation of transposase production and transposition of the element. An additional IR (Fig. 2, nt 50 to 93) overlapping the ribosome-binding site and start codon of *orfA* could provide additional translational regulation by forming a stable stem-loop structure at the 5′ end of the mRNA. This would sequester the *orfA* transla-

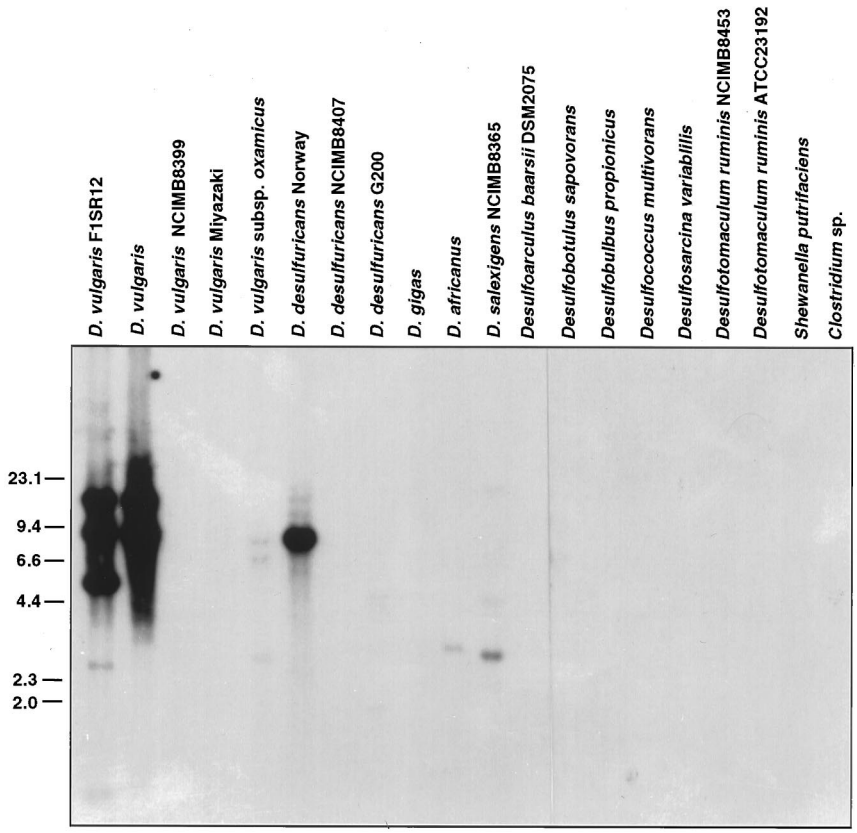


FIG. 6. Distribution of IS*D1* among different bacteria. Genomic DNA samples of *Desulfovibrio* spp. and of other bacteria as indicated above each lane were restricted with *Pvu*II, Southern blotted onto Hybond-N nylon membranes, and probed with the 0.7-kb IS*D1* probe indicated in Fig. 1C. Size markers are indicated in kilobases on the left.

tional start codon, especially when transcription is initiated from outside the IS element after its transposition into a transcriptionally active region, as in the case of IS*10* and IS*50* (4, 16) as reviewed by Chandler and Fayet (3). Otherwise, the transcript is expected to start at nt 57 and the stem-loop structure is not expected to be stable.

Multiple sequence alignment of transposases and retroviral polyproteins of the IS*3* family with OrfA and OrfB of IS*D1* revealed that IS*1222*, IS*511*, IS*R1*, IS*407*, IS*476*, IS*1313*, IS*Rm1*, IS*2*, and IS*426* are most closely related to IS*D1* (Fig. 3). OrfA of IS*D1* also showed high homology with LcrS encoded by the *lcrS* gene of plasmid pIB1 from *Yersinia pseudotuberculosis* (27), while the N terminus of OrfB of IS*D1* showed homology with LcrT, which is only 55 aa long and is encoded by the *lcrT* gene located downstream from *lcrS*. Comparison of IS*D1* and the *lcr* operon sequence (accession no. M83986) revealed the presence of an IS*D1*-like sequence at the 3′ end of the available sequence data (nt 3925 to 4521). The presence of an IRL-like sequence upstream of the *lcrS* gene and an A_7 shifty codon motif-like structure (3) between the *lcrS* and *lcrT* genes raises the possibility that these two genes originated in *Y. pseudotuberculosis* from an IS element that resembled IS*D1*. Both OrfA and OrfB of IS*D1* show highest homology to LcrS and LcrT (27). Within the IS*3* family, IS*1222* (36), IS*511* (19), IS*R1* (25), IS*407* (46), and IS*476* (15) have distinctly higher homology to IS*D1* than do IS*1312* (6), IS*Rm1* (45), IS*2* (28), IS*426* (45), and other IS*3* family members (Fig. 3). The helix-turn-helix motif suggested for OrfA of IS*2*, IS*476*, and IS*R1*, which could be involved in recognition of and binding to the terminal IRs (24), was also found in IS*D1* OrfA (Fig. 3A, aa 23 to 43). A D,D-(35)-E motif which is conserved among all IS*3* family elements identified so far as well as in retroviral integrases (18, 20) and which has been proposed to be involved directly in catalysis (20) is present in IS*D1* OrfB (Fig. 3B, aa 124, 183, and 219).

The aligned sequences for OrfA and OrfB were used to construct phylogenetic trees (not shown). The two trees implied similar phylogenetic relationships in which IS*D1* is first grouped with the *lcrS* and *lcrT* genes and then grouped with IS*R1*, IS*511*, IS*407*, IS*1222*, and IS*476*. The phylogenetic relationship of these IS elements is different from that of the hosts as derived by aligning 16S rRNA sequences. This finding implies horizontal transfer of these IS elements, which is also confirmed by the significantly different GC content of IS*D1* (52%) compared to the rest of the *D. vulgaris* genome (64%).

Insertion site of IS*D1*-A. GAPDH is one of the crucial enzymes in the Embden-Meyerhof-Parmas pathway of glycolysis. *Desulfovibrio* spp., with the exception of *D. fructosovorans* (13, 29), cannot derive energy for growth from the fermentation of hexoses, indicating the absence of a complete glycolytic pathway. However, *D. vulgaris* Hildenborough and many other *Desulfovibrio* spp. can oxidize glycerol to acetate and CO_2 through the action of a glycerol kinase, a membrane-bound glycerol-3-phosphate dehydrogenase, and glycolytic pathway enzymes to convert dihydroxyacetone phosphate to pyruvate (13). These include GAPDH, phosphoglycerate kinase (PGK), and triosephosphate isomerase (TPI). This represents the only route in *Desulfovibrio* spp. known to result in NADH production (37).

The *pgk* and *tpi* genes of several bacteria, including *Bacillus megaterium* (31) and *Corynebacterium glutamicum* (8), are located downstream from the *gap* gene in the same operon. Assuming that this same arrangement of genes is present in the *gap* operon in *D. vulgaris*, the presence of an IS*D1* element downstream from the *gap* gene suggest that the *pgk* and *tpi* genes have been inactivated by IS*D1* insertion. This idea is supported by the observation that the limited stretch of nucleotide sequence determined on the 5′ side of IS*D1*-A (Fig. 5A, *pgk*), encodes a peptide that aligns with the central part of the sequence of known PGKs (not shown). The genes surrounding IS*D1*-A may thus not encode enzymes that are currently involved in glycerol metabolism in *D. vulgaris* Hildenborough. An intriguing hypothesis is that these genes formerly provided the organism with the capacity to derive energy for growth from the glycolysis of hexoses, but that this capacity has been lost through gene inactivation (as documented here) and deletion allowing only products of glycolysis (pyruvate, lactate) supplied externally to be used as electron donors for sulfate reduction.

Potential of IS*D1* as a tool for mutagenesis of *D. vulgaris*. Wall et al. (44) have recently shown that a modified Tn*7* can be used for random mutagenesis of *D. desulfuricans* G20 at a frequency of 10^{-6} per donor. Tn*5* derivatives were also found to transpose at similar frequency following conjugation from *E. coli*. These transposons may also be useful for random mutagenesis of *D. vulgaris* Hildenborough, although initial experiments in this regard have been unsuccessful. The fact that IS*D1* is actively transposing to a frequently occurring site (AAXX) offers another route toward random mutagenesis by providing a starting point for construction of artificial transposons (38) for random mutagenesis of genes from *D. vulgaris* Hildenborough and other *Desulfovibrio* spp.

ACKNOWLEDGMENTS

This work was supported by a research grant from the Natural Science and Engineering Research Council of Canada to G.V. and by a postgraduate studentship from the Alberta Heritage Foundation for Medical Research to R.F.

We thank M. Jiang for technical assistance.

REFERENCES

1. **Bennett, P. M.** 1991. Transposable elements and transposition in bacteria, p. 323–364. *In* U. N. Streips and R. E. Yasbin (ed.), Modern microbial genetics. Wiley-Liss, New York, N.Y.
2. **Biesecker, G., J. I. Harris, J. C. Thierry, J. E. Walker, and A. J. Wonacott.** 1977. Sequence and structure of D-glyceraldehyde 3-phosphate dehydrogenase from *Bacillus stearothermophilus*. Nature (London) **266:**328–333.
3. **Chandler, M., and O. Fayet.** 1993. Translational frameshifting in the control of transposition in bacteria. Mol. Microbiol. **7:**497–503.
4. **Davis, M. A., R. W. Simons, and N. Kleckner.** 1985. Tn*10* protects itself at two levels from fortuitous activation by external promoters. Cell **43:**379–387.
5. **Deckers, H. M., and G. Voordouw.** 1994. Identification of a large family of genes for putative chemoreceptor proteins in an ordered library of the *Desulfovibrio vulgaris* Hildenborough genome. J. Bacteriol. **176:**351–358.
6. **Deng, W., M. P. Gordon, and E. W. Nester.** 1995. Sequence and distribution of IS*1312*: evidence for horizontal DNA transfer from *Rhizobium meliloti* to *Agrobacterium tumefaciens*. J. Bacteriol. **177:**2554–2559.
7. **Doak, T. G., F. P. Doerder, C. L. Jahn, and G. Herrick.** 1994. A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc. Natl. Acad. Sci. USA **91:**942–946.
8. **Eikmanns, B. J.** 1992. Identification, sequence analysis, and expression of a *Corynebacterium glutamicum* gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase. J. Bacteriol. **174:**6076–6086.
9. **Fayet, O., P. Ramond, P. Polard, M. F. Prere, and M. Chandler.** 1990. Functional similarities between retroviruses and the IS*3* family of bacterial insertion sequences? Mol. Microbiol. **4:**1771–1777.
10. **Fu, R., and G. Voordouw.** 1997. Targeted gene-replacement mutagenesis of *dcrA*, encoding an oxygen sensor of the sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. Microbiology **143:**1815–1826.
11. **Galas, D. J., and M. Chandler.** 1989. Bacterial insertion sequences, p. 109–162. *In* D. M. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
12. **Genetics Computer Group.** 1994. Program manual for the Wisconsin package, version 8. Genetics Computer Group, Madison, Wis.
13. **Hansen, T. A.** 1993. Carbon metabolism of sulfate-reducing bacteria, p. 21–40. *In* J. M. Odom and R. Singleton, Jr. (ed.), The sulfate-reducing bacteria: contemporary perspectives. Springer-Verlag, New York, N.Y.
14. **Iida, S., J. Meyer, and W. Arber.** 1983. Prokaryotic IS elements, p. 159–221. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, New York, N.Y.
15. **Kearney, B., and B. J. Staskawicz.** 1990. Characterization of IS*476* and its role in bacterial spot disease of tomato and pepper. J. Bacteriol. **172:**143–148. (Erratum, **172:**2199.)
16. **Krebs, M. P., and W. S. Reznikoff.** 1986. Transcriptional and translational initiation sites of IS*50*. Control of transposase and inhibitor expression. J. Mol. Biol. **192:**781–791.
17. **Kremer, D. A., and T. A. Hansen.** 1987. Glycerol and dihydroxyacetone dissimilation in *Desulfovibrio* strains. Arch. Microbiol. **147:**249–256.
18. **Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka.** 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell. Biol. **12:**2331–2338.
19. **Ohta, N., D. A. Mullin, J. Tarleton, B. Ely, and A. Newton.** 1990. Identification, distribution, and sequence analysis of new insertion elements in *Caulobacter crescentus*. J. Bacteriol. **172:**236–242.
20. **Polard, P., and M. Chandler.** 1995. Bacterial transposases and retroviral integrases. Mol. Microbiol. **15:**13–23.
21. **Polard, P., M. F. Prere, M. Chandler, and O. Fayet.** 1991. Programmed translational frameshifting and initiation at an AUU codon in gene expression of bacterial insertion sequence IS*911*. J. Mol. Biol. **222:**465–477.
22. **Polard, P., M. F. Prere, O. Fayet, and M. Chandler.** 1992. Transposase-induced excision and circularization of the bacterial insertion sequence IS*911*. EMBO J. **11:**5079–5090.
23. **Postgate, J. R.** 1984. The sulfate-reducing bacteria, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.
24. **Prère, M.-F., M. Chandler, and O. Fayet.** 1990. Transposition in *Shigella dysenterae*: isolation and analysis of IS*911*, a new member of the IS*3* group of insertion sequences. J. Bacteriol. **172:**4090–4099.
25. **Priefer, U. B., J. Kalinowski, B. Rüger, W. Heumann, and A. Pühler.** 1989. IS*R1*, a transposable DNA sequence resident in *Rhizobium* class IV strains, shows structural characteristics of classical insertion elements. Plasmid **21:** 120–128.
26. **Rezsohazy, 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.
27. **Rimpiläinen, M., Å. Forsberg, and H. Wolf-Watz.** 1992. A novel protein, LcrQ, involved in the low-calcium response of *Yersinia pseudotuberculosis* shows extensive homology to YopH. J. Bacteriol. **174:**3355–3363.
28. **Ronecker, H. J., and B. Rak.** 1987. Genetic organization of insertion element IS*2* based on a revised nucleotide sequence. Gene **59:**291–296.
29. **Rousset, M., Z. Dermoun, M. Chippaux, and J. P. Belaich.** 1991. Marker exchange mutagenesis of the *hydN* genes in *Desulfovibrio fructosovorans*. Mol. Microbiol. **5:**1735–1740.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Schlapfer, B. S., and H. Zuber.** 1992. Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase and triosephosphate isomerase (*gap* operon) from mesophilic *Bacillus megaturium*: comparison with corresponding genes from thermophilic *Bacillus stearothermophilus*. Gene **122:**53–62.
32. **Schwartz, E., M. Kroger, and B. Rak.** 1988. IS*150*: distribution, nucleotide sequence and phylogenetic relationships of a new *E. coli* insertion element. Nucleic Acids Res. **16:**6789–6802.
33. **Schweizer, H. P.** 1992. Allelic exchange in *Pseudomonas aeruginosa* using novel ColE1-type vectors and a family of cassettes containing a portable *oriT* and the counter-selectable *Bacillus subtilis sacB* marker. Mol. Microbiol. **6:** 1195–1204.
34. **Sekine, Y., N. Eisaki, and E. Ohtsubo.** 1994. Translational control in production of transposase and in transposition of insertion sequence IS*3*. J. Mol. Biol. **235:**1406–1420.
35. **Skarzynsky, T., P. C. E. Moody, and A. J. Wonacott.** 1987. Structure of holo-glyceraldehyde-3-phosphate dehydrogenase from Bacillus stearothermophilus at 1.8 Å resolution. J. Mol. Biol. **193:**171–187.
36. **Steibl, H. D., and F. M. Lewecke.** 1995. IS*1222*: analysis and distribution of a new insertion sequence in *Enterobacter agglomerans* 339. Gene **156:**37–42.
37. **Stokkermans, J. P.** 1993. Molecular studies on iron-sulfur proteins in *Desulfovibrio*. Ph.D. thesis. Agricultural University, Wageningen, The Netherlands.
38. **Vertes, A. A., Y. Asai, M. Kobayashi, Y. Kurusu, and H. Yukawa.** 1994. Transposon mutagenesis of coryneform bacteria. Mol. Gen. Genet. **245:**397–405.

39. **Vögele, K., E. Schwartz, C. Welz, E. Schiltz, and B. Rak.** 1991. High-level ribosomal frameshifting directs the synthesis of IS*150* gene products. Nucleic Acids Res. **19:**4377–4385.
40. **Voordouw, G.** 1988. Cloning of genes encoding redox proteins of known amino acid sequence from a library of the *Desulfovibrio vulgaris* (Hildenborough) genome. Gene **69:**75–83.
41. **Voordouw, G.** 1993. Molecular biology of the sulfate-reducing bacteria, p. 88–130. *In* J. M. Odom and R. Singleton, Jr. (ed.), The sulfate-reducing bacteria: contemporary perspectives. Springer-Verlag, New York, N.Y.
42. **Voordouw, G.** 1995. The genus *Desulfovibrio*: the centennial. Appl. Environ. Microbiol. **61:**2813–2819.
43. **Wall, J. D.** 1993. Genetics of the sulfate-reducing bacteria, p. 77–87. *In* J. M. Odom and R. Singleton, Jr. (ed.), The sulfate-reducing bacteria: contemporary perspectives. Springer-Verlag, New York, N.Y.
44. **Wall, J. D., T. Murnan, J. Argyle, R. S. English, and B. J. Rapp-Giles.** 1996. Transposon mutagenesis in *Desulfovibrio desulfuricans*: development of a random mutagenesis tool from Tn*7*. Appl. Environ. Microbiol. **62:**3762–3767.
45. **Watson, R. J., and R. Wheatcroft.** 1991. Nucleotide sequence of *Rhizobium meliloti* insertion sequence IS*Rm1*: homology to IS2 from *Escherichia coli* and IS*426* from *Agrobacterium tumefaciens*. DNA Seq. **2:**163–172.
46. **Wood, M. S., A. Byrne, and T. G. Lessie.** 1991. IS*406* and IS*407*, two gene-activating insertion sequences from *Pseudomonas cepacia*. Gene **105:** 101–105.