Association of newly discovered IS elements with the dichloromethane utilization genes of methylotrophic bacteria

M. Schmid-Appert, K. Zoller, H. Traber, S. Vuilleumier and T. Leisinger

Author for correspondence: T. Leisinger. Tel: +41 1 632 3324. Fax: +41 1 632 1148.
E-mail: leisinger@micro.biol.ethz.ch

Mikrobiologisches Institut, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zürich, Switzerland

Dichloromethane (DCM) dehalogenases enable facultative methylotrophic bacteria to utilize DCM as sole carbon and energy source. DCM-degrading aerobic methylotrophic bacteria expressing a type A DCM dehalogenase were previously shown to share a conserved 4.2 kb BamHI DNA fragment containing the dehalogenase structural gene, dcmA, and dcmR, the gene encoding a putative regulatory protein. Sequence analysis of a 10 kb DNA fragment including this region led to the identification of three types of insertion sequences identified as IS1354, IS1355 and IS1357, and also two ORFs, orf353 and orf192, of unknown function. Two identical copies of element IS1354 flank the conserved 4.2 kb fragment as a direct repeat. The occurrence of these newly identified IS elements was shown to be limited to DCM-utilizing methylotrophs containing a type A DCM dehalogenase. The organization of the corresponding dcm regions in 12 DCM-utilizing strains was examined by hybridization analysis using IS-specific probes. Six different groups could be defined on the basis of the occurrence, position and copy number of IS sequences. All groups shared a conserved 5.6 kb core region with dcmA, dcmR, orf353 and orf192 as well as IS1357. One group of strains including Pseudomonas sp. DM1 contained two copies of this conserved core region. The high degree of sequence conservation observed within the genomic region responsible for DCM utilization and the occurrence of clusters of insertion sequences in the vicinity of the dcm genes suggest that a transposon is involved in the horizontal transfer of the DCM-utilization character among methylotrophic bacteria.

Keywords: dichloromethane, dehalogenase, IS element, methylotrophic bacteria

INTRODUCTION

Some Gram-negative facultative methylotrophic bacteria utilize dichloromethane (DCM) as the sole carbon and energy source. The initial step of DCM utilization is catalysed by DCM dehalogenase, an enzyme that transforms DCM to inorganic chloride and formaldehyde, a central metabolite of methylotrophic growth. Functionally and by their amino acid sequence, bacterial DCM dehalogenases belong to the Theta-class of the large and ubiquitous superfamily of glutathione S-transferases (Leisinger et al., 1994).

Bacterial DCM dehalogenases have been shown to fall into two classes, type A and type B, which significantly differ from each other with respect to kinetic properties and amino acid sequence (Scholtz et al., 1988; Bader & Leisinger, 1994). In Methylobacterium sp. strain DM4 and other methylotrophs with a type A enzyme, the dehalogenase structural gene, dcmA, and dcmR, the gene encoding a regulatory protein governing dehalogenase expression (La Roche & Leisinger, 1990, 1991), are located on a conserved 4.2 kb BamHI fragment. In some strains, this fragment is carried on a 120 kb plasmid, whereas in others it is located on the chromosome (Leisinger et al., 1994). These findings and the observation that the ability of bacteria to utilize DCM is readily lost under non-selective conditions (Galli & Leisinger, 1988) suggest that the dcm genes have been distributed among facultative methylotrophs by hori-
zontal gene transfer. To explore this possibility in more
detail, we have analysed in *Methylobacterium* sp. strain
DM4 the DNA regions flanking the *dcm* genes. Here we
report that sequence similarity between a number of
methylotrophs with a type A DCM dehalogenase
extends over 10 kb. This conserved region comprises,
in addition to *dcmA* and *dcmR*, two ORFs of unknown
function as well as an assortment of newly discovered
insertion sequences that appear to be specifically
associated with the DCM utilization genes.

**METHODS**

**Bacterial strains, plasmids and growth conditions.** Bacterial
strains and plasmids used in this work are listed in Table 1.

Methylotrophic bacteria were grown at 30 °C in minimal
medium with methanol or DCM as carbon sources as
described elsewhere (Kohler-Staub et al., 1986). *Escherichia
coli* strains were grown on LB medium (Miller, 1972) at 37 °C.
Antibiotics for selection were added at the following concen-
trations: 200 µg ampicillin ml⁻¹, 25 µg tetracycline ml⁻¹.

**DNA manipulations.** General genetic techniques were per-
formed according to standard protocols (Ausubel et al., 1997;
Sambrook et al., 1989). Plasmids were introduced into *E. coli
strains following the procedure described by Inoue et al.

**Southern blot analysis.** Digested genomic DNA was loaded
on 1% agarose gels. The 1 kb DNA ladder used as molecular
mass marker was labelled with digoxigenin (DIG)-dUTP. The
3' ends of the marker DNA were degraded by exonuclease

**Table 1. Bacterial strains and plasmids used in this study**

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Relevant phenotype*/genotype</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Methylobacterium</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Methylobacterium extorquens</em> AM1</td>
<td>Dcm⁻</td>
<td>Green (1992)</td>
</tr>
<tr>
<td><em>Methylobacterium organophilum</em> XX</td>
<td>Dcm⁻</td>
<td>Green (1992)</td>
</tr>
<tr>
<td><em>Methylobacterium</em> sp. DM4</td>
<td>Dcm⁺</td>
<td>Galli &amp; Leisinger (1985)</td>
</tr>
<tr>
<td><em>Methylobacterium</em> sp. DM4-2cr</td>
<td>Dcm⁻</td>
<td>Galli &amp; Leisinger (1988)</td>
</tr>
<tr>
<td><em>Methylobacterium</em> sp. CM4</td>
<td>Dcm⁻</td>
<td>Doronina et al. (1996)</td>
</tr>
<tr>
<td><em>Hyphomicrobium</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Hyphomicrobium</em> sp. GJ21</td>
<td>Dcm⁺</td>
<td>D. Janssen, University of Groningen, Netherlands</td>
</tr>
<tr>
<td><em>Methylophilus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Methylophilus</em> sp. DM2</td>
<td>Dcm⁺</td>
<td>Scholtz et al. (1988)</td>
</tr>
<tr>
<td><em>Pseudomonas</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pseudomonas</em> sp. DM1</td>
<td>Dcm⁺</td>
<td>Brunner et al. (1980)</td>
</tr>
<tr>
<td><em>Methylorhabdus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Methylorhabdus multivorans</em> DM13</td>
<td>Dcm⁺</td>
<td>Doronina et al. (1995)</td>
</tr>
<tr>
<td>Unidentified methylotrophs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strains DM3, DM5-10</td>
<td>Dcm⁺</td>
<td>Galli &amp; Leisinger (1985), Galli (1986)</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5a</td>
<td>supE44 ΔlacU169 (Φ80lacZΔM15) bsdR17 recA1 endA1 relA1 bsdR RP4-2 kan::Tn5 tet::MU chromosomally integrated</td>
<td>Gibco-BRL</td>
</tr>
<tr>
<td>S17-1</td>
<td>Ap²</td>
<td>Simon et al. (1983)</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBluescript KS(+)</td>
<td>Ap², oriT from RP4</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pSUP202</td>
<td>Ap², Dcm⁻, (pVK100), 21 kb HindIII fragment encompassing the <em>Methylobacterium</em> sp. DM4 <em>dcm</em> region</td>
<td>Simon et al. (1983)</td>
</tr>
<tr>
<td>pME1540</td>
<td>Ap², (pBLS-⁺), 11 kb EcoRI–SacI fragment of IS1354</td>
<td>S. La Roche, unpublished</td>
</tr>
<tr>
<td>pME1839</td>
<td>Ap², (pBLS-⁺), 27 kb SacI–BamHI fragment with IS1355 and IS1357</td>
<td>Schmid-Appert (1996)</td>
</tr>
<tr>
<td>pME1874</td>
<td>Ap², (pBLS-⁺), 0.4 kb EcoRV fragment of IS1357</td>
<td>Schmid-Appert (1996)</td>
</tr>
</tbody>
</table>

*Dcm⁺*, DCM-utilizing; *Dcm⁻*, unable to utilize DCM.
BRL) were treated with 40 units T4 DNA polymerase for 2 hours before labelling. For this purpose 10 μg 1 kb DNA (Gibco BRL) were treated with 40 units T4 DNA polymerase for 2 min at 37 °C in 20 μl 53 mM Tris acetate (pH 7.9), 66 mM sodium acetate, 10 mM MgCl₂, 0.5 mM DTT containing 0.1 mg BSA ml⁻¹. The re-synthesis reaction was carried out for 25 min at 37 °C after addition of DIG-DUTP DNA labelling mix (Boehringer). After electrophoresis, DNA was denatured and blotted on Hybond-N membranes (Amersham) as described by the manufacturer. DNA hybridization probes were labelled by PCR amplification with alkali-labile DIG-dUTP (Roche) and blotted on Hybond-N membranes (Amersham) as described by the manufacturer. DNA hybridization probes were labelled by PCR amplification with alkali-labile DIG-dUTP and blotted on Hybond-N membranes (Amersham) as described by the manufacturer. DNA hybridization probes were labelled by PCR amplification with alkali-labile DIG-dUTP nucleotide mix. Specific DNA hybridization probes were obtained by standard protocols using 100 pmol each primer and 3 fmol template DNA. The pBluescript KS(+) derivatives pME1839, pME1841, pME1887 and pME1540, carrying appropriate sequences from IS1354, IS1355, IS1357 and IS1354 (Table 1), were used as DNA templates in the PCR reaction. The primers used to obtain specific DNA probes had the following nucleotide sequences (nucleotide positions in brackets refer to the sequence shown in Fig. 1): 5'-CGACGCGCAACCGGTT-3' (212–226), and 5'-GGGGCA-TGACGGCGAAATGGAAAG-3' (1042–1019) for IS1354; 5'-ACGGCGCATTGGAGCGAA-3' (1772–1789) and 5'-ATAAC-AGGGCACTGTATTG-3' (2390–2373) for IS1355; and universal primers 5'-ATTAACCTCCTACACTAAGG-3' (T3) and 5'-ATTACAGCTACTATAGG-3' (T7) for IS1357 and dcmA, respectively. Hybridization at 68 °C and subsequent immunochromic detection was performed following the procedure described by Engler-Blum et al. (1993). Probes were stripped from the membrane by alkali treatment for repeated hybridization of the same blot with probes of different specificity as described by the manufacturer.

**DNA sequencing and sequence analysis.** Fragments of the cosmid pME1510 were subcloned into pBluescript KS(+) for sequence analysis. DNA sequences were determined from double-stranded plasmid subclones according to a modified chain-termination method (Schaumber & Keulen, 1991). Nucleic acid and amino acid sequences were analysed by using the Genetics Computer Group program package, version 8.1 (University of Wisconsin, Madison, WI, USA). DNA and amino acid sequence similarity searches of non-redundant DNA and protein sequence databases were performed at the NCBI using the BLAST network service (Altschul et al., 1990) and the FASTA program, version 2.0.2 (Pearson, 1996). Sequence alignments were done with the GCG program PILEUP and identity scores were obtained from identity matrices of multiple sequence alignments calculated with the GCG program OLDDISTANCES. IS1354, IS1355 and IS1357 are Plasmid Reference Centre listings (E. Lederberg, Stanford, CA, USA).

**RESULTS**

**Sequence analysis of the Methylobacterium sp. DM4 dcm region**

The *dcm* genes of *Methylobacterium* sp. DM4 are located on a 4-2 kb chromosomal BamHI restriction fragment (La Roche & Leisinger, 1990; Leisinger et al., 1994). A more detailed characterization of this region in *Methylobacterium* sp. DM4 and in *Hyphomicrobium* sp. DM2 by Southern analysis indicated that similarity between these strains extended to at least 10 kb, and the corresponding region was cloned and sequenced from strain DM4. Fig. 1 gives an overview of the conserved region and indicates the location of newly sequenced DNA fragments. The region downstream of *dcmA* features two ORFs, orf192 and orf192, and IS1354. A second identical copy of this element is located at the left-hand border of the sequenced *dcm* region. IS1355 is found immediately adjacent to this element and is in turn followed by IS1357. The nucleotide sequence of the DNA fragment comprising the three newly identified insertion sequences is shown in Fig. 2, and the structural properties of these elements are summarized in Table 2.

The gene orf192 located downstream (3') of *dcmA* (Fig. 1) encodes a putative protein with a molecular mass of 40 kDa. Its sequence was preceded by a putative RBS but lacked sequences with similarity to *E. coli* consensus promoters or *Methylobacterium* promoter motifs. A database search for similar proteins gave 45% identity of Orf353 with OrfA, the deduced product of an
Fig. 2. Single strand nucleotide sequence of IS1354, IS1355 and IS1357 (accession no. X96995). IS1354 (pos. 4197—5627 in M32346), the second copy of IS1354 located downstream of orf192 in strain DM4, is identical to IS1354. Inverted repeats (IR) and direct repeats (DR) at both ends of an insertion sequence are indicated by arrows below and above the sequence, respectively. Translational products of the putative transposase genes are given below the nucleotide sequence.
Table 2. Insertion sequences IS1354, IS1355 and IS1357 of Methylobacterium sp. strain DM4

<table>
<thead>
<tr>
<th>IS element</th>
<th>Length (bp)</th>
<th>Transposase</th>
<th>Inverted repeat (bp)</th>
<th>Direct repeat (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS1354</td>
<td>1431</td>
<td>ORF421</td>
<td>48</td>
<td>8</td>
</tr>
<tr>
<td>IS1355</td>
<td>975</td>
<td>ORF153</td>
<td>9–10</td>
<td>ND*</td>
</tr>
<tr>
<td>IS1357</td>
<td>ND*</td>
<td>ORF411</td>
<td>ND*</td>
<td>ND*</td>
</tr>
</tbody>
</table>

*ND*, Not determined. The length of these sequences could not be determined, presumably due to sequence overlap between IS1354 and IS1355 and between IS1355 and IS1357.

IS1354 extended over 1431 bp and carried 48 bp imperfect terminal inverted repeats which were flanked by 8 bp direct repeats. These flanking sequence repetitions, 5'-TTGAATCA-3' for IS1354, and 5'-GCCGTCAA-3' for IS1354B, were not identical for the two copies of the element. The major ORF of IS1354 encoded a putative protein of 421 amino acids, termed Orf421. It was preceded by a putative -35/–10 promoter region and a potential RBS (Fig. 2). Orf421 exhibited similarity to putative transposases of insertions sequences from both Gram-positive and Gram-negative bacteria (Table 3). Sequence alignments of Orf421 to related transposases (not shown) led to the identification of a conserved C-terminal motif corresponding to the consensus sequence described by Eisen et al. (1994) for a family of insertion sequences including elements of plants and bacteria (Fig. 2). The size of IS1354-related elements described in the literature ranges from 1.3 to 1.7 kb. Upon insertion, they give rise to 7–9 bp direct repeats of the target site, and carry 15–41 bp inverted repeats at their extremities. Like IS1354, these elements contain a single ORF encoding a potential transposase composed of between 296 and 474 amino acids. Most of the insertion sequences homologous to IS1354 are present as multiple copies in their host strains. No sequence similarities between the target sites of these copies have been observed so that insertion of these elements is thought to be non-specific (Guilhot et al., 1992; Yates et al., 1988).

Characteristics of IS1355

IS1355 was 975 bp in length and had 9 and 10 bp terminal inverted repeats. No sequence duplications were observed in the flanking DNA. This was possibly caused by the subsequent insertion of IS1354 next to IS1355 since the 9 bp inverted repeat at the left end of IS1355 was identified within the 8 bp direct repeat flanking IS1354 (Fig. 2).

The nucleotide sequence of IS1355 revealed a short ORF, orf153, between nucleotides 61 and 520 in IS1355. It was partly overlapped by orf179, which was shifted in phase of translation by -1 nucleotide and extended from nucleotide positions 380–897 in IS1355. While orf153 was preceded by a potential –10/–35 promoter and by a putative RBS (Fig. 2), none of these structures was observed in the region 5' of orf179. Sequences similar to both Orf153 and Orf179 were found within a subgroup of the large IS5/IS4 family of insertion sequences (Rezsohazy et al., 1993). Similar transposase sequences in this group are also often encoded by two overlapping ORFs (Rezsohazy et al., 1993). The sequence 5'-AAAAAGGGGGG-3' found in the region where orf153 and orf179 overlap (Fig. 2) was very similar to potential ribosomal frameshifting signal sequences described by Iversen et al. (1994) for IS1355-related IS elements. These signals are known to direct the formation of fusion proteins as the result of -1 translational frameshifting with simultaneous slippage (Chandler & Fayet, 1993; Farabaugh, 1996). Accordingly, a frameshifting event at position 1894 of the IS1355 nucleotide sequence (Fig. 2) would result in a hypothetical fusion protein Orf153–Orf179 with a higher degree of identity to IS1355-related transposases encoded by one single reading frame than the polypeptides derived separately from orf153 and orf179 (Table 3). Interestingly, the genome of the cyanobacterium Synechocystis sp. strain PCC 6803, the sequence of which was recently reported (Kaneko et al., 1994),
Table 3. Similarity of IS1354, IS1355 and IS1357 of *Methylobacterium* sp. DM4 to other insertion sequences

<table>
<thead>
<tr>
<th>Organism</th>
<th>Insertion sequence</th>
<th>Element</th>
<th>Reference</th>
<th>Protein acc. no.*</th>
<th>Putative transposases of <em>Methylobacterium</em> sp. DM4 IS elements (percentage amino acid identity)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IS1354</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Orf421</td>
</tr>
<tr>
<td><strong>Clostridium perfringens</strong></td>
<td>IS1151</td>
<td>Daube et al. (1993)</td>
<td>S31644</td>
<td>27:1</td>
<td></td>
</tr>
<tr>
<td><strong>Mycobacterium smegmatis</strong></td>
<td>IS6120</td>
<td>Guilhot et al. (1992)</td>
<td>P35883</td>
<td>26:3</td>
<td></td>
</tr>
<tr>
<td><strong>Sinorhizobium meliloti</strong></td>
<td>ISRs5</td>
<td>Laberge et al. (1995)</td>
<td>Q52873</td>
<td>25:6</td>
<td></td>
</tr>
<tr>
<td><strong>Streptococcus thermophilus</strong></td>
<td>IS1191</td>
<td>Guédon et al. (1995)</td>
<td>Q56104</td>
<td>24:3</td>
<td></td>
</tr>
<tr>
<td><strong>Mycobacterium avium</strong></td>
<td>IS1245</td>
<td>Guerrero et al. (1995)</td>
<td>Q48368</td>
<td>24:2</td>
<td></td>
</tr>
<tr>
<td><strong>Acetobacter xylinum</strong></td>
<td>IS1032</td>
<td>Iversen et al. (1994)</td>
<td>Q44566</td>
<td>40:3</td>
<td></td>
</tr>
<tr>
<td><strong>Acetobacter xylinum</strong></td>
<td>IS1031c</td>
<td>Coucheron (1993)</td>
<td>Q44579</td>
<td>44:2</td>
<td></td>
</tr>
<tr>
<td><strong>Sinorhizobium meliloti</strong></td>
<td>ISRs4</td>
<td>Soto et al. (1992)</td>
<td>Q52958† (X65471) (D90905)</td>
<td>46:8</td>
<td>40:6</td>
</tr>
<tr>
<td><strong>Synechocystis sp. PCC 6803</strong></td>
<td>HSII1791/1792</td>
<td>Kaneko et al. (1996)</td>
<td>Q57945</td>
<td>45:0</td>
<td></td>
</tr>
<tr>
<td><strong>Synechocystis sp. PCC 6803</strong></td>
<td>Slr03500†</td>
<td>Kaneko et al. (1996)</td>
<td>Q55944/Q55945</td>
<td>45:4</td>
<td>35:4</td>
</tr>
<tr>
<td><strong>Synechocystis sp. PCC 6803</strong></td>
<td>Slr0799/8000</td>
<td>Kaneko et al. (1996)</td>
<td>Q54404</td>
<td>33:8</td>
<td>32:8</td>
</tr>
<tr>
<td><strong>Streptomyces lividans</strong></td>
<td>Tn4811</td>
<td>Chen et al. (1992)</td>
<td>P24537/P24536</td>
<td>21:1</td>
<td>18:9</td>
</tr>
<tr>
<td><strong>Burkholderia cepacia</strong></td>
<td>IS4025</td>
<td>Ferrante &amp; Lessie (1991)</td>
<td>Q44425/PQ0496</td>
<td>22:5</td>
<td>18:0</td>
</tr>
<tr>
<td><strong>Agrobacterium tumefaciens</strong></td>
<td>IS4275</td>
<td>De Meirman et al. (1990)</td>
<td>Q44425/PQ0496</td>
<td>22:5</td>
<td>18:0</td>
</tr>
<tr>
<td><strong>Synechocystis sp. PCC 6893</strong></td>
<td>HSII0667‖</td>
<td>Kaneko et al. (1996)</td>
<td>Q55974</td>
<td>24:7</td>
<td></td>
</tr>
<tr>
<td><strong>Synechocystis sp. PCC 6803</strong></td>
<td>HSII0092‖</td>
<td>Kaneko et al. (1996)</td>
<td>Q53893</td>
<td>21:0</td>
<td></td>
</tr>
<tr>
<td><strong>Synechocystis sp. PCC 6803</strong></td>
<td>Slr0099‖</td>
<td>Kaneko et al. (1996)</td>
<td>Q35871</td>
<td>20:6</td>
<td></td>
</tr>
<tr>
<td><strong>Calothrix sp. PCC 7601</strong></td>
<td>IS701</td>
<td>Mazel et al. (1991)</td>
<td>Q00461</td>
<td>19:0</td>
<td></td>
</tr>
</tbody>
</table>

* Protein database accession numbers; if not available, the gene entry accession number is given in parentheses.
† The hypothetical fusion protein Orf153–Orf179 results from a −1 translation frameshifting event with simultaneous slippage at position 1894 of the IS1355 nucleotide sequence (see text and Fig. 3).
‡ The ISRs4 sequence used here was proposed by Iversen et al. (1994). The introduction of a 1 nucleotide gap yields one continuous ORF with overall higher similarity to this group of transposase sequences.
§ Putative transposases related to IS1355 which appear to be encoding by two overlapping ORFs. Observed putative frameshift sequences: AAAAAAGGGGG (HSII1791/HSII1792, Slr0799/slr0800), AAAAAACTGG (IS402), AAAAAAGGGGG (IS427).
|| Putative transposases of *Synechocystis* as annotated by Kaneko et al. (1996). The presence of insertion sequence structures enclosing the corresponding ORFs remains to be demonstrated.
|| Four other identical putative transposase sequences and at least six further sequences differing only in a few amino acids are found in the *Synechocystis* sp. PCC 6803 genome.

1996), features many genes encoding putative transposase proteins similar to orf153 and orf179, both as single full-length ORFs and as overlapping ORFs. With regard to the latter, very similar putative frameshift signals to that found in IS1355 were found in at least two other cases (see Table 3).
IS1357, an incomplete insertion sequence

IS1357, the third insertion sequence in the dcm region of Methylobacterium sp. DM4, was located adjacent to the right hand end of IS1355 (Fig. 2). The nucleotide sequence of this element contained orf411, an ORF whose 5' region lacks a potential promoter structure and RBS. The polypeptide potentially encoded by orf411 exhibited a rather low level of identity of about 20% with putative transposases found in cyanobacteria (Table 3). Since the start of the orf411 sequence overlapped with the right-hand extremity of IS1355, the product of orf411 may have been truncated by insertion of IS1355 into IS1357. The fact that it was not possible to identify the ends of IS1357 would be consistent with such an insertion event.

Occurrence of the novel insertion sequences in methylotrophic bacteria

The distribution of the newly discovered insertion sequences among various methylotrophic bacteria was examined by Southern hybridization with the probes indicated in Fig. 3(b). The 12 DCM-utilizing methylotrophs with type A DCM dehalogenases contained copies of at least two different insertion sequences. In contrast, no hybridization signal was detected in total DNA of two Methylobacterium strains unable to grow with DCM, or in DNA of Methylobacterium sp. strain DM4-2cr (Galli & Leisinger, 1988; Table 1), a temperature-cured mutant of strain DM4 which has suffered a deletion of the entire dcm region (data not shown). The three insertion sequences were also absent
from DNA of *Methylphilus* sp. strain DM11, an organism expressing a type B DCM dehalogenase (Fig. 3).

**Organizational patterns of the dcm region in DCM-utilizing bacteria**

Southern analysis under high stringency conditions was used to examine in more detail the arrangement of the IS elements and the *dcm* genes on the genomes of the DCM-utilizing bacteria with a type A dehalogenase. An example of a hybridization experiment with *SalI*-digested total DNA of representative strains is shown in Fig. 3. Similar experiments were done with DNA that had been digested with *BamHII, BglII, EcoRI* and *PstI* (data not shown). *SalI*-digested DNA of strains DM1, DM3, DM6, DM9 and DM10, as well as DM5, DM7 and DM8, revealed identical hybridization patterns, and only representative hybridizations with DNA of strains DM1 and DM5 are presented (Fig. 3a, lanes a and d). In the case of *SalI*-digested DNA of strain DM1, the relatively strong hybridization signals obtained at about 0.8 kb with the IS1354 probe, and at about 1.3 kb with the IS1355 and IS1357 probes, respectively, suggested the presence of two hybridizing fragments of identical size in strains of the DM1 class. Additional hybridization experiments with *EcoRI* and *BglII*-digested DNA then confirmed the presence in two copies of all three IS elements in strains of this class (data not shown). It thus became clear that strains DM1, DM3, DM6, DM9 and DM10 contained two copies of the *dcm* region. Fig. 4 presents a hybridization experiment of *BglII* - and *PstI*-digested DNA carried out to unambiguously identify the two copies of the DCM dehalogenase gene in strain DM1. Since the *dcm* region had no restriction site for these enzymes, a single hybridizing band was expected, and found, for each copy of the *dcm* gene region.

Further, the larger number of hybridizing bands observed in the DNA of strains DM2 and DM13 with IS1354 and IS1357 probes, respectively, indicated that extra copies of these elements were present in these strains (Fig. 3a, lanes b and f). Subsequent Southern analysis suggested that such copies were located outside the *dcm* region (data not shown). Finally, the weak additional hybridization signals observed with the IS1355-specific probe in total DNA of strains DM2, DM4, DM13 and GJ21 (Fig. 3a, lanes b, c, f and g) suggested that more divergent IS1355-like sequences were also present in these strains.

The hybridization data were used to construct physical maps of the *dcm* region in the different strains, and these are shown in Fig. 5. In summary, six different patterns of organization were observed. The minimal structure present in all strains consisted of a conserved core region of 5.6 kb, which contains the genes *dcmR, dcmA, orf353* and *orf192* as well as IS1357. The organizational classes differ from each other with respect to the copy number and the location of additional insertion sequences flanking the conserved core region.

**Investigation of the transpositional activity of the dcm region**

The 21 kb *HindIII* fragment of strain DM4 including the 10 kb *dcm* region defined in the present work was cloned into the suicide vector pSUP202 (Simon *et al.*, 1983; see Table 1), and mobilization of the resulting recombinant plasmid pME1874 into several *Methyllobacterium dcm* strains was attempted. In no instance, however, could transconjugants be obtained that displayed the DCM-utilization phenotype, indicating the inability of the *dcm* region to transpose under the conditions of the experiment (M. Schmid-Appert & D. Gisi, unpublished results).

**DISCUSSION**

The three insertion sequences IS1354, IS1355 and IS1357 of *Methyllobacterium* sp. strain DM4 can be ascribed to class I of transposable elements (Kleckner, 1981), but clearly belong to different groups on the basis of their sequence. A close association of three different insertion sequences on a short segment of DNA, as in the present case, has rarely been observed. However, many genes encoding putative transposases similar to IS1357 and especially to IS1355 of strain DM4 (Table 3) were found in the cyanobacterium *Synechocystis* sp. PCC 6803. In one instance at least, ORFs in the *Synechocystis* genome (ORFs sll0665–sll0666) encoding polypeptides similar to those encoded by IS1355 were located immediately adjacent to an ORF (sll0667, see Table 3) similar to the C-terminal end only of Orf411 encoded by IS1357. This suggests that in *Synechocystis* sp. PCC 6803, an element similar to IS1355 inserted into a sequence similar to IS1357.

The observed accumulation of insertion sequences in the gene region involved in DCM degradation is suggestive
of a functional association between both types of genes. In addition to their ability to self-transpose during translocation to a foreign genetic context, IS elements can also play a role in gene activation (Galas & Chandler, 1989). In the present case, activation of *dcm* gene expression mediated by a promoter or regulation element internal to the insertion sequence is unlikely, since the upstream sequences of *dcm* genes are separated from the insertion sequences by the *dcm* genes themselves (Fig. 1).

Evidence for the location of specific catabolic genes on transposons is available for a number of systems (reviewed by Van der Meer et al., 1992; Wyndham et al., 1994), including several examples where the catabolic genes propagated by these elements encode dehalogenating enzymes. The chlorobenzene transposon Tn5280 of *Pseudomonas* sp. strain P51 is a typical representative of a class I composite transposon. It contains a bacterial dehalogenase gene flanked by two identical copies of an insertion sequence. The chlorobenzene transposon Tn5280 of *Pseudomonas* sp. strain P51 is a typical representative of a class I composite transposon. It contains a bacterial dehalogenase gene flanked by two identical copies of an insertion sequence. The chlorobenzene transposon Tn5280 of *Pseudomonas* sp. strain P51 is a typical representative of a class I composite transposon. It contains a bacterial dehalogenase gene flanked by two identical copies of an insertion sequence (Van der Meer et al., 1991). Nevertheless, flanking insertion sequences do not always lead to mobility of the associated catabolic genes. For example, the DNA fragment containing the genes involved in 2,4,5-trichlorophenoxyacetic acid in *Burkholderia* (Pseudomonas) cepacia AC1100 were flanked by two copies of IS931, suggesting a possible class I composite transposon structure. The corresponding DNA fragment, however, could not be demonstrated to be mobile (Haugland et al., 1990). More unusual types of mobile genetic elements involved in dehalogenation processes have also been described. These include the DEH element of *Pseudomonas putida* PP3 and a mobile DNA segment of *Xanthobacter autotrophicus* GJ10 that contains IS1247. Both elements are associated with a gene encoding a haloacid dehalogenase (Thomas et al., 1992; van der Ploeg et al., 1995), and they transpose as genetic units of variable length. Whereas the structure of the DEH element remains ill-defined, that of strain GJ10 was shown to involve only one flanking copy of IS1247.

The occurrence of two identical copies of IS1354 directly flanking genes involved in DCM utilization in strain *Methyllobacterium* sp. strain DM4 suggested at first that the genes spanned by the two copies, including the two insertion sequences IS1355 and IS1357, represent a class I composite transposon. Alternatively, the gene arrangement observed in the *dcm* region of DM4 may have been produced by cointegrate formation or by homologous recombination (Galas & Chandler, 1989). The presence of different target site duplications flanking the two copies of IS1354, however, argues against such events having taken place in the genome of DM4. Indeed, not all the DCM-utilizing strains which have been characterized contain two copies of IS1354 (Fig. 5). Of course, the analysis of the arrangement of insertion sequences in the *dcm* region of *Methyllobacterium* sp.
strain DM4 and of other related methylotrophic DCM utilizers cannot by itself explain the mechanism by which the dcm region was acquired and distributed among methylotrophs. Experiments to observe or induce the transposition of the dcm region have not been successful so far. Nevertheless, the occurrence in many DCM-utilizing methylotrophs of a highly conserved DNA fragment featuring three different insertion elements flanking a core region including the dcmA gene (Fig. 5) strongly supports the idea that the dcm region was transposed in the past.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Swiss Federal Institute of Technology, Zürich, Switzerland.

REFERENCES


Received 31 December 1996; revised 27 March 1997; accepted 8 April 1997.