

University of Groningen

Characterization of IS2112, a new insertion sequence from Rhodococcus, and its relationship with mobile elements belonging to the IS110 family

Kulakov, LA; Poelarends, Gerrit; Janssen, Dick; Larkin, MJ; Kulakov, Leonid A.; Larkin, Michael J.

Published in:
Microbiology-Uk

DOI:
[10.1099/13500872-145-3-561](https://doi.org/10.1099/13500872-145-3-561)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1999

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kulakov, L. A., Poelarends, G. J., Janssen, D. B., Larkin, M. J., Kulakov, L. A., & Larkin, M. J. (1999). Characterization of IS2112, a new insertion sequence from Rhodococcus, and its relationship with mobile elements belonging to the IS110 family. *Microbiology-Uk*, 145(3), 561-568. DOI: 10.1099/13500872-145-3-561

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Characterization of IS2112, a new insertion sequence from *Rhodococcus*, and its relationship with mobile elements belonging to the IS110 family

Leonid A. Kulakov,¹ Gerrit J. Poelarends,² Dick B. Janssen² and Michael J. Larkin¹

Author for correspondence: Leonid A. Kulakov. Tel: +44 1232 335577. Fax: +44 1232 661462.
e-mail: l.kulakov@queens-belfast.ac.uk

¹ The Questor Centre, David Keir Building, The Queen's University of Belfast, Belfast BT9 5AG, UK; and School of Biology and Biochemistry, Medical Biology Centre, The Queen's University of Belfast, Belfast BT9 7BL, UK

² Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

A new insertion sequence (IS2112) was identified in the genome of the 1-haloalkane-utilizing bacterium *Rhodococcus rhodochrous* NCIMB 13064. The insertion element is 1415 bp long, does not contain terminal inverted repeats, and is not flanked by directly repeated sequences. IS2112 belongs to the IS110 family of transposable elements, and forms a separate subfamily, along with IS116. Two copies of IS2112 were found in *R. rhodochrous* NCIMB 13064 and one, two or three copies of a similar sequence were detected in five other 1-haloalkane-degrading *Rhodococcus* strains. There were no sequences homologous to IS2112 found in the 1-haloalkane-degrading '*Pseudomonas pavonaceae*' 170 and *Rhodococcus* sp. HA1 or in several *Rhodococcus* strains which do not utilize haloalkanes. IS2112 was originally found in plasmid pRTL1 of *R. rhodochrous* NCIMB 13064, which harbours genes encoding utilization of 1-haloalkanes, and was located 5 kbp upstream of the haloalkane dehalogenase gene (*dhaA*). Although the second copy of IS2112 in strain NCIMB 13064 was also present on the pRTL1 plasmid, these sequences do not apparently comprise a single composite transposon encoding haloalkane utilization. An analysis of derivatives of NCIMB 13064 revealed that IS2112 was involved in genome rearrangements. IS2112 appeared to change its location as a result of transposition and as a result of other rearrangements of the NCIMB 13064 genome.

Keywords: *Rhodococcus*, genome rearrangements, insertion element, haloalkane dehalogenase

INTRODUCTION

The genus *Rhodococcus* includes a diverse group of Gram-positive bacteria which are noted for their ubiquity and remarkable ability to degrade a wide variety of xenobiotics, including aliphatic and aromatic compounds. Although there have been some studies related to the genetics of rhodococci, the means by which they acquire biodegradation pathways have not been investigated. However, it is clear that in other bacteria mobile DNA elements can contribute to the acquisition of biodegradation pathways by promoting mutations and

genomic rearrangements in the chromosome and in transmissible plasmids. Few transposable elements have been identified in *Rhodococcus* species: the nucleotide sequences of IS1166 and IS1295 were determined (Denome & Young, 1995; Jager *et al.*, 1995) and were shown to be very similar to *Mycobacterium smegmatis* IS6120 (Guilhot *et al.*, 1992) and *Staphylococcus aureus* IS256 (Rouch *et al.*, 1987) insertion sequences. Recently IS1415, belonging to the IS21 family, was identified in *Rhodococcus erythropolis* and its transposition in the bacterial genome was demonstrated (Nagy *et al.*, 1997).

Rhodococcus rhodochrous NCIMB 13064 degrades 1-chloroalkanes by means of a haloalkane dehalogenase (DhaA) that was shown to be encoded on a 100 kbp plasmid designated pRTL1 (Kulakova *et al.*, 1995). A

The GenBank accession number for the sequence reported in this paper is AF017179.

number of dehalogenase-negative derivatives of this strain were isolated. These mutants showed rearrangements in both the chromosome and plasmid pRTL1 (Kulakova *et al.*, 1995, 1997). Some of the derivatives were able to revert to the original phenotype (Kulakova *et al.*, 1995), suggesting the occurrence of reversible genetic rearrangements and a possible role for a transposon or an insertion element.

In this paper, we describe the presence of a DNA segment proximal of the haloalkane dehalogenase gene that shows sequence similarity to some transposable elements found in actinomycetes. The distribution of the corresponding transposable element and its role in genetic rearrangements are discussed.

METHODS

Bacterial strains and plasmids. *Escherichia coli* strain DH5 α (Sambrook *et al.*, 1989) and plasmid vectors pUC129 (Keen *et al.*, 1988) and pBluescript KS(+) (Stratagene) were used for the cloning experiments. *R. rhodochrous* NCIMB 13064 (Curragh *et al.*, 1994), *R. erythropolis* NCIMB 13065 and *Rhodococcus* strains Acr33, I1 and A1 (Kulakov *et al.*, 1998) were described previously. Of these strains, only NCIMB 13064 and NCIMB 13065 were haloalkane degraders and carried the *dhaA* gene, which encodes haloalkane dehalogenase. The haloalkane-degrading *Rhodococcus* strains GJ70 (Janssen *et al.*, 1987), HA1 (Scholtz *et al.*, 1987), m15-3 (Yokota *et al.*, 1986), Y2 (Sallis *et al.*, 1990) and TB2 (collection of Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute) and '*Pseudomonas pavonaceae*' 170 (Verhagen *et al.*, 1995) were used for the analysis of the distribution of IS2112. All *Rhodococcus* strains as well as '*P. pavonaceae*' 170 were classified on the basis of 16S rRNA gene sequence analysis (unpublished results). Derivatives of *R. rhodochrous* NCIMB 13064 were described previously (Kulakova *et al.*, 1995) as well as obtained in this work.

Media and growth conditions. *Rhodococcus*, *Pseudomonas* and *E. coli* strains were propagated in a rich (2YT or NB) or minimal (M9) medium (Miller, 1972). When required, Difco Bacto agar (1.8%, w/v) was added to the medium. Ampicillin (100 $\mu\text{g ml}^{-1}$), tetracycline (15 $\mu\text{g ml}^{-1}$), IPTG (50 $\mu\text{g ml}^{-1}$) and X-Gal (50 $\mu\text{g ml}^{-1}$) were used for the detection of recombinant plasmids.

DNA techniques. Total DNA from *Rhodococcus* strains was isolated as described by Kulakova *et al.* (1995). Recombinant DNA work was done by using standard protocols as described by Sambrook *et al.* (1989). Restriction of both total DNA and recombinant plasmids was performed using enzymes obtained from Pharmacia, according to the manufacturer's instructions. The Sephaglas BandPrep kit (Pharmacia) was used for the recovery of restriction fragments and PCR products from agarose gels.

DNA fragments were transferred to Hybond-N+ membrane (Amersham) or to a positively charged nylon membrane (Boehringer) by Southern blotting, as described by Sambrook *et al.* (1989). The membranes were then treated according to the instructions of the manufacturer. Hybridization probes were labelled using the Gene Images Random Prime Labelling Module (Amersham) or the digoxigenin-dUTP DNA Labelling and Detection kit (Boehringer). Hybridization was carried out at 64 °C and detection was performed according to the manufacturer's instructions.

DNA sequencing and analysis. Plasmids pUC129 and pBluescript KS(+) with inserts containing IS2112 and surrounding regions were used as templates in sequence reactions using the Taq Dye-Deoxy Terminator Cycle Sequencing kit (Applied Biosystems). DNA sequences were obtained using an automatic sequencer (Applied Biosystems, model 373A). The nucleotide sequences of both strands were determined. Sequencing of PCR products was performed after template purification from agarose gels. Primer synthesis was carried out by Dr C. Stevenson (School of Biology and Biochemistry, The Queen's University of Belfast). Editing and initial analysis of the sequences was performed using the DNASIS (Hitachi) software package. Searches for nucleotide and amino acid sequence similarities were done by using the FASTA and BLAST programs (Pearson & Lipman, 1988) and the EMBL and GenBank databases.

Alignments of the protein sequences were performed using the CLUSTAL W program (Thompson *et al.*, 1994). Phylogenetic analysis of the alignment of amino acid sequences was done using the PHYLIP (version 3.57c) package (Felsenstein, 1997) and the TREECON program (Van de Peer & De Wachter, 1994). For the PHYLIP analysis, bootstraps were obtained with the SEQBOOT program (100 data sets were generated) and distance matrices were generated with the PROTDIST program (PAM001 matrix) and the DNADIST program. Neighbour-joining was carried out with the NEIGHBOR program using a random input order of sequences. Parsimony analyses were done with the PROTPARS and DNAPARS programs using ordinary parsimony and randomized input order of sequences. For the analyses with the TREECON program, Tajima and Nei correction (Tajima & Nei, 1984) was used and trees were generated by neighbour-joining.

PCR. Amplification of *Rhodococcus* sequences was done with Taq+ DNA polymerase (Stratagene) in a buffer supplied by the manufacturer. Reactions were carried out in volumes of 25 μl or 50 μl with dNTPs at 200 μM concentrations, primers at 0.15 μM each, DNA – 50–100 ng and Taq+ 0.5 unit per reaction. The following temperature profile was used: denaturation at 95 °C for 3 min, followed by 30 cycles of 95 °C for 40 s, 60 °C for 30 s, and 72 °C for 1 min.

To determine nucleotide sequences adjacent to IS2112 in different derivatives of NCIMB 13064, total DNA was isolated from the corresponding strains and inverse PCR amplifications were carried out as follows. Total DNA (0.5–1.0 μg) was digested with either *Bam*HI, *Xho*I or *Eco*RI, ligated with T4 DNA ligase and used in the PCR reactions with different internal, divergent IS2112 primers. The primers were: 5' GCGCAGATCGGTGAGAC 3' (969–985 bp of the IS2112 sequence, forward primer) and 5' GCCTCCCACGCAATCTC 3' (843–859 bp, reverse primer) – used with *Bam*HI-, *Xho*I- or *Eco*RI-digested DNA; 5' GCGCAGATCGGTGAGAC 3' (969–985 bp, forward primer) and 5' CCCAATGGAGTTTCGTCG 3' (359–375 bp, reverse primer) for *Bam*HI- or *Eco*RI-digested DNA. PCR products obtained were sequenced as described above.

RESULTS AND DISCUSSION

Identification and sequence analysis of IS2112

The *R. rhodochrous* haloalkane dehalogenase gene (*dhaA*) was characterized previously (Kulakova *et al.*, 1997). A number of direct and inverted repeats were found immediately upstream of the *dhaA* coding region, and the involvement of these repeats in genomic

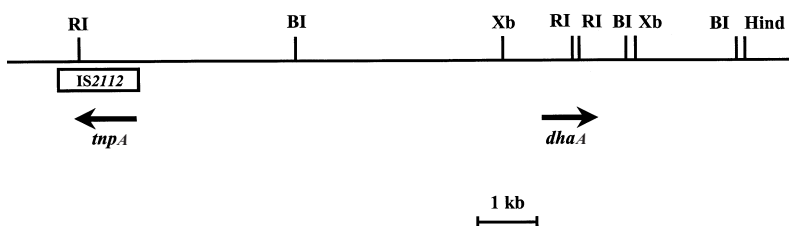


Fig. 1. Schematic representation of the pRTL1 region in the vicinity of IS2112. The position of IS2112 is indicated by a boxed area. *tnpA*, putative transposase gene of IS2112; *dhaA*, haloalkane dehalogenase gene (Kulakova *et al.*, 1997). RI, *EcoRI*; BI, *BamHI*; Xb, *XbaI*; Hind, *HindIII*.

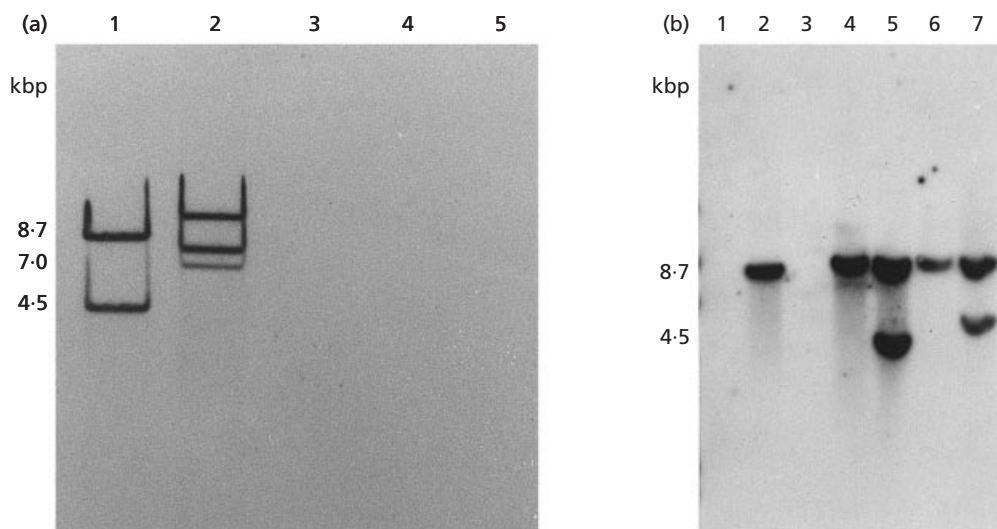


Fig. 2. Southern blot analysis of *Bam*HI-digested total DNA preparations of *Rhodococcus* strains and '*P. pavonaceae*' 170. Restriction fragments were separated in a 0.9% agarose gel and transferred onto a membrane. Hybridization was carried out with a 1050 bp internal *Sal*I fragment of IS2112. (a) Hybridization using the Gene Images kit (Amersham). Lanes: 1, NCIMB 13064; 2, NCIMB 13065; 3, *Rhodococcus* sp. I1; 4, *Rhodococcus* sp. Acr33; 5, *Rhodococcus* sp. A1. (b) Hybridization with the DNA Labelling and Detection kit (Boehringer). Lanes: 1, strain 170; 2, GJ70; 3, HAI; 4, m15-3; 5, NCIMB 13064; 6, Y2; 7, TB2.

rearrangements associated with mutations that resulted in loss of dehalogenase activity was suggested (Kulakova *et al.*, 1997). In order to further analyse genetic elements which may be involved in such rearrangements of the *R. rhodochrous* genome, overlapping clones covering approximately 40 kbp of plasmid pRTL1, including the *dhaA* gene, were obtained from the gene bank of NCIMB 13064 (Kulakova *et al.*, 1997) using the cosmid vector pLAFR5.

DNA sequencing of these cosmid clones revealed the presence of an ORF 5 kbp upstream of the *dhaA* gene transcribed in the opposite direction to the *dhaA* gene (Fig. 1). This ORF was preceded by a putative ribosome-binding site (Shine & Dalgarno, 1974) and coded for a putative protein of 400 amino acids with a calculated molecular mass of 43769 Da. A search with the FASTA program (Pearson & Lipman, 1988) in different protein databases revealed that the deduced amino acid sequence was similar to that of transposase ORFs present in mobile elements found in actinomycetes.

No inverted or direct repeats were identified in the vicinity of the analysed ORF, but it is known that some of the transposable elements found in Gram-positive

bacteria lack the terminal inverted repeats that are typical of the majority of transposons and also do not duplicate the target site in the genome upon insertion (Leskiw *et al.*, 1990; Hernandez *et al.*, 1994). Accordingly, we concluded that the identified *R. rhodochrous* NCIMB 13064 sequence is an insertion element and we named it IS2112. The exact size and ends of IS2112 were identified on the basis of the analysis of several copies of this element and its integration site (see below).

To identify the number of IS2112 copies present in NCIMB 13064, a 1050 bp *Sal*I DNA fragment comprising almost the entire putative transposase gene of IS2112 was used as a probe in Southern hybridization experiments. Two copies of homologous DNA were detected by this analysis (Fig. 2a). Two *Bam*HI fragments, each containing a copy of IS2112, were cloned from NCIMB 13064 using the pUC129 vector and analysed. The resulting sequences of the putative transposase genes and adjacent regions (197 bp upstream and 15 bp downstream) were very similar, but not identical. The nucleotide substitutions (total of 23) were evenly distributed and in all cases but two affected the third position of the triplets in the coding region.

IS2112	-----MRETHHAWVGDAGKGGHWAUVDDEAGLIVWSHKIIINDETAI	60
IS116	-----MSTRHDRIWVGIDAGKGGHWAUVDADGETLFPSTKVINDEAQQV	60
IS900	-----MAQP--VWAGVDAGKADHYCMVINDDAQRLLSQRVANDEAAL	60
IS901	-----MAEPDRVWVGIDVGGKSTHHAICAIDDTGKVVVSKKIPNEQAAI	60
IS1110	MLAGLSVGNMVDKRIIGGSMEADQLWAGADVGEHHWVVCVDDDKGTAVLSRKLTDNEQPI	60
*...*...*	
IS2112	LHALSEILALAD--ELHWAVDISGTSSTLLALLAAHQSAVYVPGRTVNRMSGAYRG-E	120
IS116	L-TLIETARERE--EVRWAVDISGRASTLLALLVAHQNVVYVPGRTVNRMSGAYKG-E	120
IS900	LELIAAVVTLADGGVETWADLNAGGAALLIALLIAAGQRLLYIPGRTVHHAAGSYRG-E	120
IS901	EDLIAQGGRIAN--HVVWAIDLTPPPAALLIALLVLLSAKREVYVPGRTVNTMASHAFRG-E	120
IS1110	RELIDEDALGCG--DVSWTVDLTVYASLLTFLADAGKSVRYLTGRAVWQASVTYRGGE	120
*...*...*	
*...*...*	
IS2112	AKTDARDAYVIAETARHR-HDLSQLIPPAQLAELALLTSRRTDLVADRVLINRLRDTL	180
IS116	GKTDKADARVIADQARMR-RDFAPLDRPPELVTLRLLTNRADLIADRVLINRLRDL	180
IS900	GKTDKADAAIADQARMR-HDLQLRAGDDIAVELRILTSSRDLVADRTRAIENARPA	180
IS901	GKTDKADARVIETARHR-RDLSPVVPGEDLVAELRSLTRYRSDLMADWVGVNRLRSM	180
IS1110	AKTDKADARVIADQSRMGADLPVLPDSDLITELRMLTAHRADLVADTRITINRLRHQL	180
	***...*...*...*	
IS2112	TGVPFALERAFDFGRHKAGLVLLTFQFQTPRAIRRRGRTRLSLWLKGRGVR--GADGVAEI	240
IS116	TGICPALERAFDYSAAGPVMVLTQYQTPAALRRITGVKRLTTLWLGRRKVR--DADTVAEK	240
IS900	AGILSALERAFDYNSRAALLTLLTGTYQTPDALRSAGGARVAFLRKRKR--NADTVAAT	240
IS901	TAIPFALEAAFDYS-TRAPLIIIVSAMCTPEIRSAGKAGVIKHLRKNRAWENNDDTIADK	240
IS1110	IAVCPALERVAQLTQDRGWVLLTRYQRPKAIHRSVGSRLTKMLGDAGVR--NAATTIAAK	240
*...*...*	
*...*...*	
IS2112	AWEAASQNVVLPGEAVAAQLVADIVRVLAVDVLARIDAQIGETFRHRPQAIVIESMP	300
IS116	AIEAARTQQVVLPGKRAKTLVCDLAHQLLALDERIKDNDREIRETFRTDDRAEIIIESMP	300
IS900	ALQAANAQHSIVPGQQLAATVVARLAKEVMALDTEIGD'TDAMIERRFRRHAEIILSMP	300
IS901	ALAAAAGQITITLPGE-EAAALIQLAARLLDLDLQIKDIDKQITNKFGHEHPAAIIIESMP	300
IS1110	AVTAAKSQVTRVLPGEDVAAGLVADLAGEVVALDDRIKST'DADIEDRFRHRPAAEVITSLP	300
*...*	
*...*	
*...*...*	
IS2112	GIGPILGAEFVVAAG-DMSAYS DAGHLASAAGLVVPRDSSGRRTGNMHRPKYSRRLRRV	360
IS116	GMGPVILGAEFVAIVG-DLSGKYDAGRLASHAGLAEVPRDSSGRRTGNVHRPQRYNRRLRWL	360
IS900	GGVILGAEFLAATGDMAAAFASADRLAGVAGLAEVPRDSSGRISGNLKRPRYDRILLRA	360
IS901	GMGPHLGAFLVITGGNMMAFTNPGRLASFAGLVVPRDSSGRITGNLHRPKYRNRRLRV	360
IS1110	GMGPRILGAEFLAAVG-DPTLIGSADQLAAWAGLAEVSRDSSGRITGRHLTPPKYSRRLRRV	360
*...*	
*...*	
*...*...*	
IS2112	FYLSAQTSSIKDGNREFYVYQKRSRGRMRHVQAVIALARRQNVNIVALRRDRLDFIVDAPA	420
IS116	FYMSAQTAMRPGPSRDYVYKRGEGLLHTQALLSLARRVDVWAMLRDKRLFTPAPPV	420
IS900	CYLSALVSIKRTDSSRTYVDRKTEGRKHHTQAVLALARRLNLVWAMLRDHAIVHPATTT	420
IS901	FYLSAALSSLIKEGSRAFYDRKRSENHIHTQALLLARRHVWVWALLRDNRTWQOQQPT	420
IS1110	MYMSALTARICDPHSAKYYQRKRDGKRIIPATLCLARRVPTSPSSVTTAPGNPHHR	420
*...*...*	
*...*...*	
*...*...*	
IS2112	LRTA-----	469
IS116	TQTA-----	469
IS900	AAA-----	469
IS901	VAAA-----	469
IS1110	SPSRRLDIFIESPFAIYALVSGVGGPVGAIVLAVVDDDDDKKAFGGR	469

Fig. 3. Amino acid sequence alignment of the transposases of IS2112 and related IS elements. Alignment was performed with the CLUSTAL W program. Residues identical in all proteins are indicated by asterisks and those identical in IS2112 and IS116 sequences are indicated by dots. The highly conserved regions 1 and 2 (Hernandez *et al.*, 1994) are indicated by underlining. IS902 transposase, which is very similar to that of IS901, is not presented in this alignment.

Although very closely related to IS2112, this copy may be designated iso-IS2112.

The base composition of IS2112, with a 60.9 mol% G+C content, was similar to that of a previously analysed region of the pRTL1 plasmid (60 mol%) (Kulakova *et al.*, 1997), but somewhat lower than the 67–72 mol% typical for *Rhodococcus* species (Finnerty, 1992) and another *R. rhodochrous* NCIMB 13064 plasmid (Kulakov *et al.*, 1997). The comparatively high G+C content of IS2112 was reflected in the preferential use of C and G residues at the 5' ends of the codons of the putative transposase gene. A similar pattern of codon usage was found for the previously characterized *dhaA* gene (Kulakova *et al.*, 1997). These findings may reflect an association between IS2112 and the conjugative plasmid pRTL1 and its relatively recent occurrence in *R. rhodochrous* NCIMB 13064.

Relatedness of IS2112 to other transposable elements

Database searches (GenBank and EMBL) revealed that the putative transposase of IS2112 was most closely related to that of IS116 (Leskiw *et al.*, 1990) with 61%

amino acid identity. Significant similarity was also detected with IS1110 (Hernandez *et al.*, 1994) and IS901 (Kunze *et al.*, 1992), both isolated from *Mycobacterium avium*, and IS900 from *Mycobacterium paratuberculosis* (Green *et al.*, 1989). These IS elements comprise a single family based upon similarity of their deduced gene products (Leskiw *et al.*, 1990). All members of this family have a single transposase gene which occupies almost the entire element (Leskiw *et al.*, 1990; Hernandez *et al.*, 1994). A lower level of similarity of IS2112 was found with the *Streptomyces* mini-circle (Henderson *et al.*, 1989; Lydiate *et al.*, 1986, 1989) and IS110 (Bruton & Chater, 1987).

An alignment of the amino acid sequences of the transposase genes of IS elements belonging to the IS110 family (Mahillon & Chandler, 1998), which was previously known as the IS900 family (Hernandez *et al.*, 1994), clearly demonstrates the relatedness between IS2112 and other family members (Fig. 3). Two highly conserved regions noticed earlier (Hernandez *et al.*, 1994) are also present in the IS2112 transposase. These regions, designated 1 and 2 in Fig. 3, are also closely related to sequences within the pilin gene inversion proteins of *Moraxella* species (Fulks *et al.*, 1990). The

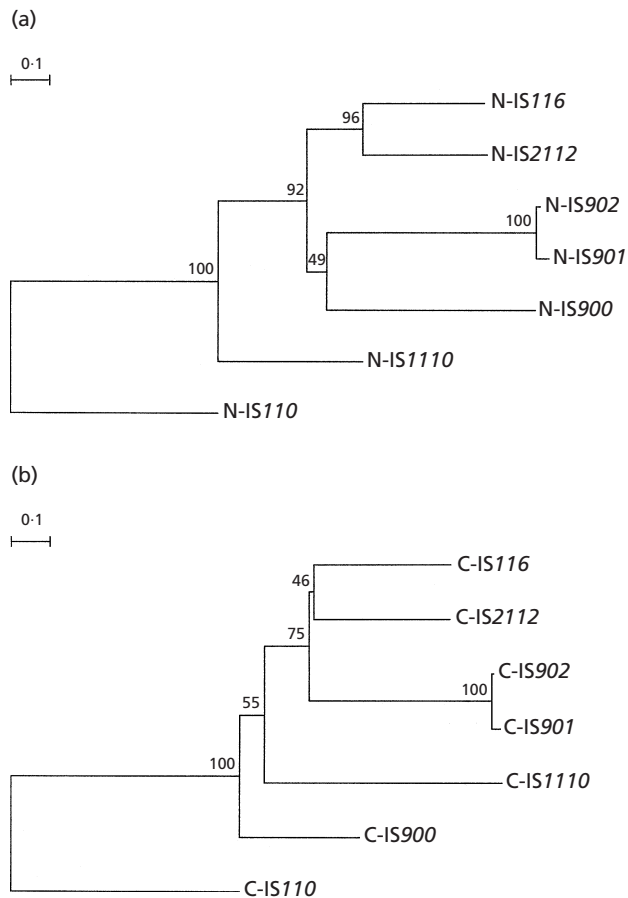


Fig. 4. Phylograms of IS110 family transposases obtained by the neighbour-joining approach. Trees were constructed with the TREECON program. (a) N-terminal parts of transposase genes; (b) C-terminal parts. Bootstrap values (%) are indicated at the nodes. The scale bars represent 0.1 substitution per site. The IS110 transposase was used as the outgroup (similar phylogenies were obtained when parsimony analysis of the same data was done).

DXXDA motif in region 1 is one of the most conserved features of the reverse transcriptases (Xiong & Eickbush, 1988).

It was noticed that when different IS110 family transposases were aligned, the extent of homology differed in the N- and C-terminal halves of some proteins (Fig. 3). A phylogenetic analysis of the alignment of amino acid sequences was performed. Two sets of data were analysed: an alignment of the N-terminal parts of the transposases corresponding to residues 1–200; and of the C-terminal parts corresponding to residues 201–400 of IS2112 (Fig. 3). This division of the analysed transposases was not based on any functional properties of the corresponding parts, but has been done solely as an approach for investigation of the possibility of recombinational events (Eltis & Bolin, 1996). The trees obtained by the neighbour-joining approach are presented in Fig. 4(a, b). This analysis confirms that IS2112 is most closely related to IS116 and that its position on

the phylogenetic tree is independent of the data set analysed. A distinctly different family result was obtained for other members of the IS110 family. Thus, the N-terminal parts of IS900 and IS901 are more closely related than their C-terminal parts. The position of IS1110 on the phylogenetic tree was also dependent on the part of the protein sequence analysed. The most likely explanation for these results is that a recombinational exchange took place in the course of evolution of members of the IS110 family, although different rates of evolution and/or parallel evolution of N- and C-terminal parts of transposases cannot be excluded. Whichever hypothesis is assumed, it may be concluded that both IS2112 and IS116 comprise a distinct subfamily in the IS110 family of transposable elements.

Analysis of NCIMB 13064 and its derivatives

R. rhodochrous NCIMB 13064 can undergo spontaneous genomic rearrangements which result in the loss of the haloalkane utilization (Dha^+) phenotype (Kulakova *et al.*, 1995, 1997). Some of the Dha^- derivatives (P200, B4) were stable, while others (S9 and P400) reverted to the original Dha^+ phenotype. In the case of strain S9, integration of the pRTL1 plasmid into the chromosome took place, whereas strain P200 had a deletion of approximately 20 kbp in the pRTL1 plasmid (Kulakova *et al.*, 1995). These mutants were analysed by Southern blotting to determine the possible role of IS2112 in rearrangements of the NCIMB 13064 genome.

Analysis of a 25 kbp region downstream of the haloalkane dehalogenase gene (*dhaA*) in NCIMB 13064 did not reveal sequences homologous to IS2112. Analysis of the deletion mutant P200 demonstrated the absence of IS2112 sequences. Assuming that the mutation was caused by a single deletion event, both copies of IS2112 may therefore be localized upstream of the *dhaA* gene. This means that the *dhaA* gene is probably not within a typical composite Tn-like structure bordered by IS2112 sequences. It is notable that either loss or relocation of the IS2112 sequences had occurred in all of the Dha^- derivatives analysed (hybridization results are not presented). Loss of both copies of IS2112 was also detected in a derivative (S14-F) which arose spontaneously in the course of laboratory cultivations of NCIMB 13064 and which had retained the Dha^+ phenotype.

Hybridization experiments showed that one copy of IS2112 which was found 5 kbp upstream of the *dhaA* gene was completely lost in strain S9. However, in the revertant strain S92 one copy of IS2112 was found on an 8.7 kbp *Bam*HI fragment as in NCIMB 13064, while a second copy of IS2112 had appeared on a smaller 2 kbp *Bam*HI fragment (results are not presented). Partial sequencing of these copies (approx. 300 bp from both ends) did not reveal any differences to the original IS2112.

To analyse the nature of these rearrangements, the DNA sequences surrounding IS2112 in NCIMB 13064 and in the S9 and S92 mutants were obtained using an inverse PCR method. The reactions were conducted using

IS2112	<u>CATGGTGC</u> ↓TCCGT..... <u>CATTGAAAC</u> ↓TCGCTT	NCIMB13064
IS2112'-site	CATGGTGCCTCGCTT	S9
IS2112	<u>CATTGAAAC</u> ↓TCCGT..... <u>CATTGAAAC</u> ↓TCGTAA	NCIMB13064
IS2112	<u>CATGCTGC</u> ↓TCCGT..... <u>CATTGAAAC</u> ↓TCGCTT	NCIMB13065
IS116	<u>CATGGTCGG</u> ↓TCCTT..... <u>CATTGAGAT</u> ↓TCTCCTGG	
IS900	<u>CATGTGTT</u> ↓TCCTT..... <u>CATTGAGAAT</u> ↓CCCCTTA	
IS900	<u>CATGTGGTGT</u> ↓TCCTT..... <u>CATTGAGAAT</u> ↓CTCCTTT	
IS901	<u>CATGCGCTGAT</u> ↓TCCTT..... <u>CATTGAGAT</u> ↓TCCTTT	
IS902	<u>CATTACAGT</u> ↓TCCTT..... <u>CATTGAGAT</u> ↓TCCTTT	
IS1110	<u>CATTGCTCCG</u> ↓TCCTT..... <u>CATTGAGAG</u> ↓TCCTTT	
CONSENSUS:	CAT (N) ₇₋₁₀ C (G/T) T.....CAT (N) ₆₋₇ C/T	

Fig. 5. Insertion sites and IS ends for the IS900 family of elements. Sequences for IS target site junctions are taken from the following sources: IS116 (Leskiw *et al.*, 1990), IS900 (Green *et al.*, 1989), IS901 (Kunze *et al.*, 1992), IS902 (Moss *et al.*, 1992), IS1110 (Hernandez *et al.*, 1994) and IS2112 (this study). The alignment of the sequences was performed manually. Not all sequences obtained for IS900, IS901, IS902 and IS2112 are presented in this figure. Vertical arrows indicate junctions between the insertion sequence and the target site sequences. The nucleotide sequence which represents an empty insertion site for IS2112 (S9 strain) is shown. Bases conserved in all known sequences and the CCT sequence conserved in all but IS2112 are underlined. The consensus sequence is indicated.

internal primers for IS2112 and total DNA preparations restricted with *Bam*HI or *Eco*RI which were subsequently ligated. Analysis of the sequences (approx. 100 bp) immediately adjacent to IS2112 in strain S9 showed that they were identical to those surrounding the corresponding copy in the parental NCIMB 13064 strain. Sequences (approx. 100 bp) surrounding both copies of IS2112 in the S92 strain were also obtained and shown to be identical to those present around the two copies of IS2112 in the original NCIMB 13064 strain. These results suggested that one copy of IS2112 was precisely deleted in strain S9, but was transposed to the same site in the revertant strain S92 as a result of a transposition of the remaining copy of IS2112. However, the results do not definitely prove replicative transposition of IS2112 and further experiments are needed to elucidate the transposition mechanism. The new fragment on which a copy of IS2112 was found in the revertant S92 probably originated from an unrelated genomic rearrangement in the adjacent region.

Analysis of IS2112 junctions with target sites

The junctions of IS2112 with the target sites were analysed on the basis of the two sequenced copies of this element from NCIMB 13064, one copy from NCIMB 13065 and also the empty integration site detected in strain S9. To obtain the sequence of the latter, PCR primers were used which anneal in the regions adjacent to the ends of the corresponding copy of IS2112 and the resulting PCR fragment (150 bp) obtained using S9 DNA was analysed. The sequences for the IS2112 junctions with target sites are presented in Fig. 5.

The relationship between IS900 and IS elements related to it is clearly pronounced when their ends and target-junction sites are compared (Fig. 5). None of these elements has terminal inverted repeats and duplication of the target site also does not occur, with the exception of IS1110, which seems to duplicate a target sequence of 5 bp (TCCTT) (Hernandez *et al.*, 1994). It is also worth noting that a precise distinction between the IS sequence and target sequences was not possible in all cases.

Although IS2112 generally conformed with all of the features typical of these elements (Hernandez *et al.*, 1994), some important differences in its ends and target site were observed (Fig. 5). The sequences CAT and CCT were considered the most typical features of the target sites and ends of the IS900 and related elements. A distance of six to nine nucleotides between these sequences is also conserved in all of the members of the IS110 family studied (Moss *et al.*, 1992). The CAT sequence located six to seven nucleotides from the insertion site was found near all three IS2112 sites (two in NCIMB 13064 and one in the NCIMB 13065 strain). However, no CCT sequences were present within 10 nucleotides of the target site. The CCT sequence at the 5' end of IS2112 was substituted by CGT, whereas CAT at the 3' end was conserved (Fig. 5). Determination of IS2112 ends allowed the precise size of this element to be calculated as 1415 bp.

Occurrence of IS2112 in different strains

To study the distribution of IS2112, *R. rhodochrous* NCIMB 13064, several 1-haloalkane-degrading *Rhodococcus* strains, *Rhodococcus* strains which are not capable of growth on 1-haloalkanes, and one 1-haloalkane-degrading *Pseudomonas* strain were analysed by hybridization experiments. A 1050 bp *Sal*I DNA fragment was used as a hybridization probe as described above. Sequences homologous to IS2112 were detected in *Rhodococcus* strains NCIMB 13064, NCIMB 13065, TB2, Y2, GJ70 and m15-3 (Fig. 2a, b). All these strains are haloalkane degraders. However, two of the haloalkane-degrading strains, *Rhodococcus* sp. HA1 and '*P. pavonaceae*' 170, did not carry sequences homologous to IS2112 (Fig. 2b). None of the other analysed *Rhodococcus* strains showed hybridization with the IS2112 probe. The simultaneous presence of IS2112 and the ability to utilize 1-haloalkanes in a number of *Rhodococcus* strains may be explained by the location of IS2112 and the *dhaA* gene close to each other on the pRTL1 plasmid. This plasmid was shown to be capable of conjugational transfer (Kulakova *et al.*, 1995) and so

may be considered responsible for the linked transfer of IS2112 and *dhaA* to various *Rhodococcus* strains. Indeed, a plasmid of a similar size as that of pRTL1 (100 kbp) was detected in strain NCIMB 13065.

One copy of IS2112 was found in GJ70, m15-3 and Y2, whereas two copies were present in NCIMB 13064 and TB2 (Fig. 2a, b). Three copies were found in NCIMB 13065, one of which showed a hybridization band of noticeably lower intensity, suggesting that a sequence with a lower degree of homology to IS2112 was present in this strain (Fig. 2a).

Concluding remarks

The newly identified insertion element IS2112 belongs to the IS110 family and is one of the first characterized transposable elements of *Rhodococcus*. IS2112 is most closely related to IS116 and these elements probably comprise a separate subfamily. IS2112 is a member of a group of transposable elements found in actinomycetes which do not possess terminal repeats and in most cases do not duplicate their target sequences upon integration. When considering these differences from the majority of transposable elements, the mechanism of transposition used by this group of transposons is of considerable interest. Although little information is available on their transposition mechanism, it is worth noting, however, that the *S. coelicolor* mini-circle element was initially found in a circular form and as such integrated with up to 100 times higher frequency than when cloned as a linear copy (Henderson *et al.*, 1989). Mahillon & Chandler (1998) communicate that DNA fragments carrying abutted IS900 ends have been detected *in vivo*. Their appearance was dependent on an intact transposase gene, and their nucleotide sequence was consistent with the formation of a circular form of the element (although it was also consistent with the presence of a head-to-tail dimer).

IS2112 has so far been found only in 1-haloalkane-degrading *Rhodococcus* strains. Using derivatives of *R. rhodochrous* NCIMB 13064, we were able to show that IS2112 is involved in the genome rearrangements characteristic for this strain and that some of these rearrangements were connected with the expression of the haloalkane dehalogenase gene. However, there is not yet information available which shows a direct link between these rearrangements and the IS2112 transposition.

ACKNOWLEDGEMENTS

This project was supported by The Queen's University Environmental Science and Technology Research Centre, European Regional Development Fund – Technology Development Programme and by EC Framework IV Environment and Climate grant ENV4-CT95-0086.

REFERENCES

Bruton, C. L. & Chater, K. F. (1987). Nucleotide sequence of IS110, an insertion sequence of *Streptomyces coelicolor* A3(2). *Nucleic Acids Res* **15**, 7053–7065.

Curragh, H., Flynn, O., Larkin, M. J., Stafford, T. M., Hamilton, J. T. G. & Harper, D. B. (1994). Haloalkane degradation and assimilation by *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology* **140**, 1433–1442.

Denome, S. A. & Young, K. D. (1995). Identification and activity of two insertion sequence elements in *Rhodococcus* sp. strain IGTS8. *Gene* **161**, 33–38.

Eltis, L. D. & Bolin, J. T. (1996). Evolutionary relationship among extradiol dioxygenases. *J Bacteriol* **178**, 5930–5937.

Felsenstein, J. (1997). PHYLIP (Phylogeny Inference Package), version 3.57c. Department of Genetics, University of Washington, Seattle.

Finnerty, W. R. (1992). The biology and genetics of the genus *Rhodococcus*. *Annu Rev Microbiol* **46**, 193–218.

Fulks, K. A., Marrs, C. F., Stevens, S. P. & Green, M. R. (1990). Sequence analysis of the inversion region containing the pilin genes of *Moraxella bovis*. *J Bacteriol* **172**, 310–316.

Green, E. P., Tizard, M. L., Moss, M. T., Thompson, J., Winterbourne, D. J., McFadden, J. J. & Hermon-Taylor, J. (1989). Sequence and characteristics of IS900, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. *Nucleic Acids Res* **17**, 9063–9073.

Guilhot, C., Gicquel, B., Davies, J. & Martin, C. (1992). Isolation and analysis of IS6120, a new insertion sequence from *Mycobacterium smegmatis*. *Mol Microbiol* **6**, 107–113.

Henderson, D. J., Lydiate, D. J. & Hopwood, D. A. (1989). Structural and functional analysis of the mini-circle, a transposable element of *Streptomyces coelicolor* A3(2). *Mol Microbiol* **3**, 1307–1318.

Hernandez, P. M., Fomukong, N. G., Hellyer, T., Brown, I. N. & Dale, J. W. (1994). Characterization of IS1110, a highly mobile genetic element from *Mycobacterium avium*. *Mol Microbiol* **12**, 717–724.

Jager, W., Schafer, A., Kalinowski, J. & Puhler, A. (1995). Isolation of insertion elements from Gram-positive *Brevibacterium*, *Corynebacterium* and *Rhodococcus* strains using the *Bacillus subtilis* *sacB* gene as a positive selection marker. *FEMS Microbiol Lett* **126**, 1–6.

Janssen, D. B., Jager, D. & Witholt, B. (1987). Degradation of n-haloalkanes and dihaloalkanes by wild-type and mutants of *Acinetobacter* sp. strain GJ70. *Appl Environ Microbiol* **49**, 673–677.

Keen, N. T., Tamaki, S., Kobayashi, D. & Trollinger, D. (1988). Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. *Gene* **70**, 191–197.

Kulakov, L. A., Larkin, M. J. & Kulakova, A. N. (1997). Cryptic plasmid pKA22 isolated from naphthalene degrading derivative of *Rhodococcus rhodochrous* NCIMB13064. *Plasmid* **38**, 61–69.

Kulakov, L. A., Delcroix, V. A., Larkin, M. J., Ksenzenko, V. N. & Kulakova, A. N. (1998). Cloning of new *Rhodococcus* extradiol dioxygenase genes and study of their distribution in different *Rhodococcus* strains. *Microbiology* **144**, 955–963.

Kulakova, A. N., Stafford, T. M., Larkin, M. J. & Kulakov, L. A. (1995). Plasmid pRTL1 controlling 1-chloroalkane degradation by *Rhodococcus rhodochrous* NCIMB13064. *Plasmid* **33**, 208–217.

Kulakova, A. N., Larkin, M. J. & Kulakov, L. A. (1997). The plasmid-located haloalkane dehalogenase gene from *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology* **143**, 109–115.

Kunze, Z. M., Portaels, F. & McFadden, J. J. (1992). Biologically distinct subtypes of *Mycobacterium avium* differ in possession of insertion sequence IS901. *J Clin Microbiol* **30**, 2366–2372.

- Leskiw, B. K., Mevarech, M., Barritt, L. S., Jensen, S. E., Henderson, D. J., Hopwood, D. A., Bruton, C. J. & Chater, K. F. (1990). Discovery of an insertion sequence, IS116, from *Streptomyces clavigerus* and its relatedness to other transposable elements from actinomycetes. *J Gen Microbiol* **136**, 1251–1258.
- Lydiate, D. J., Ikeda, H. & Hopwood, D. A. (1986). A 2.6 kb DNA sequence of *Streptomyces coelicolor* A3(2) which functions as a transposable element. *Mol Gen Genet* **203**, 79–88.
- Lydiate, D. J., Ashby, A. M., Henderson, D. J., Kieser, H. M. & Hopwood, D. A. (1989). Physical and genetic characterization of linear copies of the *Streptomyces coelicolor* mini-circle. *J Gen Microbiol* **135**, 941–955.
- Mahillon, J. & Chandler, M. (1998). Insertion sequences. *Microbiol Mol Biol Rev* **62**, 725–774.
- Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Moss, M. T., Malik, Z. P., Tizard, M. L., Green, E. P., Sanderson, J. D. & Hermon-Taylor, J. (1992). IS902, an insertion element of the chronic-enteritis-causing *Mycobacterium avium* subsp. *silvaticum*. *J Gen Microbiol* **138**, 139–145.
- Nagy, I., Schoofs, G., Vanderleyden, J. & De Mot, R. (1997). Transposition of the IS21-related element IS1415 in *Rhodococcus erythropolis*. *J Bacteriol* **179**, 4635–4638.
- Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* **85**, 2444–2448.
- Rouch, D. A., Byrne, M. E., Kong, Y. C. & Skurray, R. A. (1987). The *aacA-aphD* gentamicin and kanamycin resistance determinant of Tn4001 from *Staphylococcus aureus*: expression and nucleotide sequence analysis. *J Gen Microbiol* **133**, 3039–3052.
- Sallis, P. J., Armfield, S. J., Bull, A. T. & Hardman, D. J. (1990). Isolation and characterization of a haloalkane halohydrase from *Rhodococcus erythropolis* Y2. *J Gen Microbiol* **136**, 115–120.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Scholtz, R., Schmuckle, A., Cook, A. M. & Leisinger, T. (1987). Degradation of eighteen 1-monoalkanes by *Arthrobacter* sp. strain HA1. *J Gen Microbiol* **133**, 267–274.
- Shine, J. & Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to non-sense triplets and ribosomal binding sites. *Proc Natl Acad Sci USA* **71**, 1342–1346.
- Tajima, F. & Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* **1**, 269–285.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.
- Van de Peer, Y. & De Wachter, R. (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Applic Biosci* **10**, 569–570.
- Verhagen, C., Smit, E., Janssen, D. B. & van Elsas, J. D. (1995). Bacterial dichloropropene degradation in soil; screening of soils and involvement of plasmids carrying the *dhla* gene. *Soil Biol Biochem* **27**, 1547–1557.
- Xiong, Y. & Eickbush, T. H. (1988). Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns. *Mol Biol Evol* **5**, 675–690.
- Yokota, T., Fuse, H., Omori, T. & Minoda, Y. (1986). Microbial dehalogenation of haloalkanes mediated by oxygenase or halohydrase. *Agric Biol Chem* **50**, 453–460.
-
- Received 12 August 1998; revised 10 November 1998; accepted 25 November 1998.