A new subgroup of the IS3 family and properties of its representative member ISPpy1

Mayya Petrova, Natalya Shcherbatova, Zhosephine Gorlenko and Sofia Mindlin

Institute of Molecular Genetics, Russian Academy of Sciences, Kurchatov sq. 2, Moscow 123182, Russia

Recently, we described a novel insertion element, ISPpy1, isolated from a permafrost strain of Psychrobacter maritimus. In this work, we demonstrated that ISPpy1 is a member of a novel subgroup of the IS3 family of insertion sequences (ISs) that was not identified and characterized previously. IS elements of this subgroup termed the ISPpy1 subgroup are broadly distributed among different taxa of Eubacteria, including Geobacteraceae, Chlorobiaceae, Desulfobacteraceae, Methylobacteriaceae, Nitrosomonadaceae and Cyanobacteria. While displaying characteristic features of the IS3-family elements, ISPpy1 subgroup elements exhibit some unusual features. In particular, most of them have longer terminal repeats with unconventional ends and frameshifting box with an atypical organization, and, unlike many other IS3-family elements, do not exhibit any distinct IS specificity. We studied the transposition and mutagenic properties of a representative member of this subgroup, ISPpy1 and showed that in contrast to the original P. maritimus host, in a heterologous host, Escherichia coli K-12, it is able to transpose with extremely high efficiency into the chromosome, either by itself or as a part of a composite transposon containing two ISPpy1 copies. The majority of transposants carry multiple chromosomal copies (up to 12) of ISPpy1. It was discovered that ISPpy1 is characterized by a marked mutagenic activity in E. coli: its chromosomal insertions generate various types of mutations, including auxotrophic, pleiotropic and rifampicin-resistance mutations. The distribution of IS elements of the novel subgroup among different bacteria, their role in the formation of composite transposons and the horizontal transfer of genes are examined and discussed.

INTRODUCTION

The distribution and abundance of insertion sequences (ISs) among bacterial genomes are highly variable (Nagy & Chandler, 2004; Sawyer et al., 1987; Sigier et al., 2006). Many strains of the same species can either contain multiple IS copies or be completely devoid of ISs. For instance, the study by Sawyer et al. (1987) of the distribution of six different ISs in 71 natural isolates of E. coli revealed that the majority of isolates have no or few copies of a given IS element and that only a few isolates contain a large number of copies of this element. Similarly, IS distribution was found to vary greatly among different E. coli subspecies. For example, it was shown that E. coli K-12 and E. coli W contain IS150, whereas this IS element was absent in E. coli B and E. coli C (Schwartz et al., 1988); IS2 was present in E. coli K-12 and E. coli B, but absent in E. coli C (Hu & Deonier, 1981).

Based on such experiments, in conjunction with results of works demonstrating a high homogeneity of IS copies in the same genome (Biseric & Ochman, 1993; Lawrence et al., 1992; Nyman et al., 1981), it was suggested that IS elements are often recently acquired (Lawrence et al., 1992; Sawyer et al., 1987; Wagner, 2006). In accordance with this view, ISs are periodically eliminated from bacterial populations and are reintroduced into the same bacterial species via horizontal transfer (Lawrence et al., 1992; Wagner, 2006). It was also proposed that the changes that reduce the IS copy number and decrease the transposition activity of already established elements have a selective advantage because they decrease the deleterious effects of IS elements on host fitness and viability (Nagy & Chandler, 2004; Wagner, 2006). It should be noted that, despite the many efforts that have been made to elucidate the mechanisms underlying the enormous variation in genomic abundance and distribution of IS elements (e.g. Wagner, 2006), many aspects of this phenomenon remain poorly understood (Touchon & Rocha, 2007).

Abbreviation: IS, insertion sequence.

The GenBank/EMBL/DDBJ accession numbers for the new insertion sequences ISPpy1, ISCysp23, ISAe17, ISGie6, and ISGpy1 are AM992204, CP000807, CP002985, JN616388, and AP012160, respectively.
Recently, we described a novel IS3-family insertion element, ISppy1 (Fig. 1a). This element was found on the plasmid pKLH80 in the ancient bacterial strain Psychrobacter maritimus MR29-12 (previously designated P. psychrophilus) isolated from a 15,000–40,000 year old permafrost sample (Mindlin et al., 2008; Petrova et al., 2009). Intriguingly, ISppy1 exhibited pronounced transposition activity in cells of heterologous hosts, such as Acinetobacter calcoaceticus and E. coli (Petrova et al., 2009, 2012). It was demonstrated that, in both species, ISppy1 can generate novel composite transposons (designated Tn5080 and Tn5080a, Fig. 1b) flanked by two copies of ISppy1 and containing streptomycin- and tetracycline-resistance genes derived from pKLH80. It was also shown that the composite transposon Tn5080a translocates efficiently into different target plasmids and into the E. coli chromosome (Petrova et al., 2012).

In this work, we used the ISppy1 element as a model to study the mechanisms of IS transposition in heterologous bacteria and investigated the properties and behaviour of this element in E. coli in greater detail. We studied the frequency and consequences of ISppy1 insertions into the E. coli chromosome and demonstrated a strong mutagenic effect produced by this element. We also studied the target preference of ISppy1 by analysing independent insertions of ISppy1 and of the composite transposon Tn5080a into vector plasmids. Finally, we analysed the distribution of ISppy1 and related elements among different bacteria and demonstrated that ISppy1 belongs to a novel subgroup of IS3 elements that are widely distributed among various groups of bacteria.

**METHODS**

**Media and growth conditions.** The cultures were grown in Luria-Bertani (LB) broth and LB agar containing 20 g Bacto agar l−1 at 30 °C. The plates were supplemented with selective agents at the following concentrations (µg ml−1): chloramphenicol (Cm), 20; streptomycin (Sm), 50; tetracycline (Tc), 15; gentamicin (Gm), 4; ampicillin (Ap) 50; nalidixic acid (Nal), 20; and rifampicin (Rif), 20–25. The antibiotics were obtained from the following sources: ampicillin, from Biokhimik; gentamicin, tetracycline, nalidixic acid, from Sigma; streptomycin, from Kraspharma; chloramphenicol, from Reachim; and rifampicin, from Pharma Synthesis. Adams minimal agar medium (0.1% NH₄Cl, 0.15% K₂HPO₄, 0.35% NaH₂PO₄, 0.01% MgSO₄·7H₂O, 0.2% glucose, pH 7.2, 20.0% agar (Difco)) was used to isolate auxotrophic mutants.

**Measurement of growth rate of bacteria.** To measure the growth rate of different strains, the overnight cultures grown in LB broth were diluted 20-fold into fresh medium and additionally incubated for 6 h. To determine cell density, optical density (OD) was measured using a spectrophotometer (Thermo Spectronic Genesys 10 UV) at a wavelength of 600 nm using 100 mm cuvettes. The OD of sterile medium was subtracted from the measured OD.

**Bacterial strains and plasmids.** The E. coli K-12 strains used were JF238 (prototroph, gyrA91) and the RifR derivative of UB5201 (pro met recA56 gyrA rif-r) (Kholodii et al., 2000). The recombinant plasmids pBR325-80.2 (SmR TcR CmR) and pBR325-Ppy1 (TcR CmR), which are derivatives of pBR325 (Petrova et al., 2012) containing the transposons Tn5080a and ISppy1 inserted into the ampicillin-resistance gene (bla), respectively, were used in transposition experiments. Plasmid pBR325 (ApR TcR CmR) was used as a control. Plasmid pGEM-5SZII (−) (derivative of pUC19, ApR) was used as a control.

**Transposition of Tn5080a and ISppy1 from pBR325 into the E. coli chromosome.** The strategy used for the selection of Tn5080a

---

**Fig. 1.** General organization of ISppy1 (a) and Tn5080a (b). The solid vertical boxes indicate the left (IRl) and right (IRr) terminal IRs. orfA and orfB are the two open reading frames of ISppy1. The position of the frameshifting box (FB) is shown in the 3′-area of orfA, strA-strB, streptomycin resistance genes; tetR-tetH, genes conferring resistance to tetracycline. (c) The strategy used to analyse the orientation of ISppy1 and Tn5080a insertions and to obtain the PCR products for sequencing. Black arrows show primers 1–4 used in this work; bla, ampicillin-resistance gene in pBR325 and pGEM5. On direct insertions, PCR primer pairs 3 + 1 and 2 + 4 were used; on reverse insertions, 3 + 2 and 1 + 4 were used.
translocations into the chromosome of *E. coli* was as follows: cultures of JF238 and UB520 rif strains containing the recombinant plasmid pBR325-80.2 were grown in LB at 30 °C with shaking for 48–72 h, and with several successive transfers into fresh LB, followed by proper dilution and plating on LB agar supplemented with Nal or Rif. The colonies were then checked for resistance to Cm, Sm, and Tc. Clones with the CmR SmR TcR phenotype were selected, and transposition frequency was calculated. A similar procedure was used for the analysis of transpositions of *IS*<sub>py</sub>1. pBR325-Ppy1 was chosen as a donor plasmid. The strain JF238 (pBR325-Ppy1) was grown in LB at 30 °C for 18 h, diluted twice a day with fresh LB for a period of 2 days, and plated on LB agar supplemented with Nal. Clones with the Cm<sup>R</sup> Tc<sup>R</sup> phenotype (presumably containing no plasmids) were selected, and the presence of *IS*<sub>py</sub>1 in their genome was tested using a colony-hybridization method. Chromosome localization of some of them was supported by blot hybridization. Confirmation of plasmid loss in these experiments was obtained by PCR with a specific primer on pBR325 (L08855) backbone sequences (positions from 1519 to 2487). We observed 100% correlation (20/20) between loss of Cm<sup>R</sup> and plasmid backbone sequences.

**Isolation of Tn5080a insertions into the bla gene.** We used two JF238 derivatives carrying the independent Tn5080a chromosome insertions JF147 and JF148 (Petrova et al., 2012). The plasmid pGEM-SZ(−) was introduced into these strains via transformation. After two serial transfers of Ap<sup>R</sup> transformants, Ap<sup>R</sup> transformants were selected and checked for resistance to Ap; Ap<sup>S</sup> colonies presumably carrying the insertion of Tn5080a into the bla gene were selected and their orientation and lengths of the direct repeats were determined (see below). To isolate independent Tn5080a insertions, only one or two recombinant plasmids were chosen from each experiment. In addition to derivatives of pGEM-SZ(−), insertions of IS<sub>py</sub>1 and Tn5080a into plasmid pBR325, which were obtained in our previous work, were also analysed (Petrova et al., 2012). In both plasmid sets, only insertions into the *bla* gene were investigated.

**Isolation of mutant *E. coli* strains with IS<sub>py</sub>1 insertions.** To isolate auxotrophic strains, Cm<sup>R</sup> clones arising in the cultures of JF238 (pBR325-Ppy1) and containing IS<sub>py</sub>1 in their chromosome were checked for the ability to grow on Adams minimal medium. Clones that did not grow on this medium were retained for further investigations.

To isolate Rif<sup>R</sup>, strains, cultures of JF238 (pBR325-Ppy1) obtained after prolonged incubation in LB broth were plated using appropriate dilutions on LB agar in the absence or in the presence of 20 µg ml<sup>−1</sup> rifampicin. The colonies that grew in the presence of rifampicin were checked for resistance to ampicillin and gentamicin and for their ability to grow on minimal medium. The Rif<sup>R</sup> mutants that did not grow on minimal medium were selected for further studies.

To isolate prototrophic revertants, the overnight cultures of auxotrophs grown in LB broth were plated using appropriate dilutions on Adams minimal agar medium and on LB agar for colony counting. The colonies that grew on the minimal medium were checked repeatedly for growth on the same medium. The revertant frequency was determined in three independent experiments. The prototrophic clones were selected after 2–3 days of incubation of drops of auxotroph culture plated on Adams minimal medium. DNA from the Rif<sup>R</sup> auxotrophic clones and its prototrophic revertants was isolated and studied by Southern blot hybridization.

**DNA methodology.** Routine methods were used (Sambrook et al., 1989). Genomic DNAs were extracted using a Genomic DNA purification kit (Fermentas). DNA sequencing was performed at the Inter-Institute Centre for Collective Use ‘GENOME’ at the Institute of Molecular Biology, and Moscow Screening of bacterial strains for the presence of IS<sub>py</sub>1-like elements was performed using the colony-hybridization method (Petrova et al., 2002). The copy number of IS<sub>py</sub>1 was estimated using Southern blot hybridization. For this purpose, all DNAs analysed were digested with *Avall* (in this restriction reaction, the number of hybridized bands corresponds to the IS<sub>py</sub>1 copy number). The probe used for Southern blot hybridization corresponding to IS<sub>py</sub>1 was obtained by PCR as described in our previous work (Petrova et al., 2009, 2012) and contained 845 bp of the central part of IS<sub>py</sub>1 (position from 286 to 1130 according to AM992204). The orientations of the Tn5080a and IS<sub>py</sub>1 insertions into the *bla* gene of pGEM-SZ(−) and pBR325 were determined by PCR using primers 1–4, which are specific to the sequences of IS<sub>py</sub>1 and the plasmid ampicillin-resistance gene *bla* (Fig. 1c). The insertion sites and target DNA duplications were determined via the sequencing of the resulting PCR products.

**Bioinformatics analysis.** Nucleotide sequences were analysed at the website of the National Center for Biotechnology Information using the BLAST network service (Altschul et al., 1997) and the Vector NTI Advance v. 11 (Life Technologies). The terminal inverted repeats were verified manually. Phylogenetic analysis was conducted using the Phylogeny.fr platform (Dereeper et al., 2008) available at http://www.phylogeny.fr. The ‘Advanced’ mode was used to construct phylogenetic trees: MUSCLE for multiple alignment, GBLOCKS for alignment curation, PhyML for phylogeny and TreeDyn for tree drawing. Trees were constructed by the maximum-likelihood method. Bootstrap values were calculated based on 100 computer-generated trees. All ISs protein sequences were obtained from IS-finder (https://www.is-biotoul.fr). To analyse the subgroup affiliation of disputable ISs and demonstrate strong differences between IS150 and IS<sub>py</sub>1 subgroups, IS elements with identity from 90 to 30% belonging to these subgroups were chosen. In addition we analysed the elements that have been previously included in the IS150 subgroup, but contained consensus DD(35)E motif, typical for subgroup IS<sub>py</sub>1.

**RESULTS**

**General features and distribution of IS<sub>py</sub>1**

We have shown previously that IS<sub>py</sub>1 displays features that are typical of the IS sequences of the IS3 family (Fig. 1a). However, some features of IS<sub>py</sub>1 distinguish it from other IS3-family elements (Table 1) (Petrova et al., 2012).

To obtain further insight into the distribution of IS<sub>py</sub>1 in natural bacterial populations, we attempted to identify additional IS<sub>py</sub>1 copies in bacterial strains available in our collection. For this purpose we examined the distribution of IS<sub>py</sub>1 among different species of Proteobacteria, particularly those related to *P. maritimus*; we screened seven strains belonging to the *Psychrobacter* genus, 20 strains of the closely related genus *Acinetobacter*, and 26 bacterial strains belonging to *Pseudomonas* and *Stenotrophomonas* genera for the presence of sequences with homology to this IS element. However, none of the more than 50 strains tested contained sequences with homology to IS<sub>py</sub>1.

**ISP<sub>py</sub>1 is a member of a novel subgroup of the IS3 family of insertion elements**

To investigate the origin of IS<sub>py</sub>1 and to identify related elements, we performed a search for homologous...
Table 1. Comparative structures and distribution of typical IS3 family elements and ISs belonging to IS\textsuperscript{Ppy1}-subgroup

<table>
<thead>
<tr>
<th>Name</th>
<th>Original host (no. of copies)</th>
<th>Localization</th>
<th>Length (bp)</th>
<th>Terminal residues</th>
<th>Frameshifting signals</th>
<th>Other hosts (no. of copies)</th>
<th>Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS3 family</td>
<td>–</td>
<td>–</td>
<td>1200–1550</td>
<td>5′-TG–CA-3′</td>
<td>(A)(_9)G</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>IS\textsuperscript{Ppy1}</td>
<td>\textit{Psychrobacter maritimus} MR29-12 (1)</td>
<td>pKLH80</td>
<td>1275</td>
<td>5′-TA–TA-3</td>
<td>(A)(_9)G</td>
<td>–</td>
<td>AM992204</td>
</tr>
<tr>
<td>ISC\textsuperscript{sp23}</td>
<td>\textit{Cyanothecaceae sp.} ATCC 51142 (1)</td>
<td>Linear chromosome</td>
<td>1283</td>
<td>5′-TG–TG-3′</td>
<td>(A)(_9)G</td>
<td>–</td>
<td>CP000807</td>
</tr>
<tr>
<td>ISA\textsuperscript{fe17}</td>
<td>\textit{Acidithiobacillus ferrooxidans} SS3 (5)</td>
<td>Chromosome</td>
<td>1334</td>
<td>5′-TA–TA-3′</td>
<td>(A)(_9)G</td>
<td>–</td>
<td>CP002985</td>
</tr>
<tr>
<td>ISA\textsuperscript{ba14}</td>
<td>\textit{Acinetobacter baumannii} 1656-2 (2)</td>
<td>Chromosome</td>
<td>1282</td>
<td>5′-AG–CT-3′</td>
<td>(U)(_8)A†</td>
<td>A. baumannii, chromosome (2); A. \textit{lwoffii}, plasmids (1), (1); A. \textit{johnsonii}, plasmid (1); \textit{P. maritimus} MR29-12, plasmid pKLH80 (1); \textit{Providencia stuartii}, plasmid (1)</td>
<td>CP001921 CP001937 JQ364968 JQ001791 JN616388 JX443123 JN687470</td>
</tr>
<tr>
<td>ISA\textsuperscript{ju2}</td>
<td>\textit{Bradyrhizobium japonicum} USDA110 (9)</td>
<td>Chromosome</td>
<td>1364</td>
<td>5′-AG–CT-3′</td>
<td>(U)(_8)G†</td>
<td>–</td>
<td>NC004463</td>
</tr>
<tr>
<td>ISA\textsuperscript{li5}</td>
<td>\textit{Azospirillum lipoferum} 4B (1)</td>
<td>Main chromosome</td>
<td>1266</td>
<td>5′-TG–CA-3</td>
<td>(U)(_8)G†</td>
<td>–</td>
<td>JQ001791</td>
</tr>
<tr>
<td>ISA\textsuperscript{go6}</td>
<td>\textit{Geobacter lovleyi} SZ (1)</td>
<td>pGLOV01</td>
<td>1281</td>
<td>ND</td>
<td>(A)(_8)U†</td>
<td>–</td>
<td>JN616388</td>
</tr>
<tr>
<td>ISA\textsuperscript{ge1}</td>
<td>\textit{Gluconacetobacter xylinus} NBRC 3288 (1)</td>
<td>pGXY010</td>
<td>1313</td>
<td>5′-CA–TC-3′</td>
<td>(U)(_8)A†</td>
<td>\textit{Gluconobacter oxydans} 621H chromosome (8)</td>
<td>AP012160 CP000009</td>
</tr>
</tbody>
</table>

Accession numbers (ACs) of original host strains are in bold; ACs of host strains found later in italics.
ND, Not determined.
*Summarized data.
†Predicted frameshifting sites that were not described previously in IS elements.
sequences in public databases. The search did not identify any full-length IS\textit{Ppy1} copies in available DNA sequences. Surprisingly, however, a homology search in the DNA database revealed the existence of two contigs containing portions of an IS\textit{Ppy1}-like element in a recently submitted partial sequence of the whole genome of \textit{Psychrobacter} sp. 1501 (2011) isolated from human blood. The first contig (AFHU01000198) contains an IR sequence at its 3’ end that is completely identical to the left IR of IS\textit{Ppy1} and the entire \textit{orfB} gene, with 98% nucleotide identity to IS\textit{Ppy1} (18 mismatches over 848 nt). The other contig (AFHU01000091) contains a portion of the \textit{orfA} gene at its 5’ end with 99% identity to IS\textit{Ppy1} (three mismatches over 233 nt) and an IR that is completely identical to the right IR of IS\textit{Ppy1}. Thus, although the entire sequence of the genome has not yet been subjected to final National Center for Biotechnology Information (NCBI) review, we suppose that the complete variant of IS\textit{Ppy1} is actually present in \textit{Psychrobacter} sp. 1501 (2011).

In addition, we identified several IS elements with moderate levels of identity to IS\textit{Ppy1}. One of these elements was found in the linear chromosome of \textit{Cyanothece} sp. ATCC 51142 (Welsh et al., 2008). We termed this novel IS IS\textit{Cysp23} and compared its structure with that of IS\textit{Ppy1}. We found that the identity of these two elements at the amino acid sequence level was as great as 45% and 56% for \textit{OrfA} and \textit{OrfB}, respectively, with identity at the most conserved region reaching 62%. Interestingly, the translation frameshifting signal typical of the IS3 family, A6(G), is extended in both IS\textit{Ppy1} and IS\textit{Cysp23}, which have two and three additional adenines, respectively, at their upstream ends. It is worth noting that only a single copy of IS\textit{Cysp23} was found in \textit{Cyanothece} 51142, which was present on the 429701 bp linear chromosome; no additional copies were detected on the 4934 kb circular chromosome or on its four plasmids.

To identify the evolutionary relationships of IS\textit{Ppy1}, IS\textit{Cysp23}, and related elements, we constructed a phylogenetic tree using the best-studied IS3 family members via multiple alignments of \textit{OrfA} and \textit{OrfB} transposase sequences. This analysis revealed that IS\textit{Ppy1} and IS\textit{Cysp23} exemplify a novel subgroup of the IS3 family, unidentified previously. The IS elements from this subgroup are found in genomic sequences in numerous bacterial taxa. In particular, they are present in chromosomes and plasmids of \textit{alpha} (Gluconacetobacter and \textit{Azospirillum}), \textit{beta} (\textit{Nitrosomonas, Azorarcus, Poloromonas}), \textit{gamma} (\textit{Acidithiobacillus, Rhizobium, Marinobacter, Halorhodospira}) and \textit{delta} (\textit{Geobacter, Pelobacter, Desulfatibacillus, Desulfovibrio}) subdivisions of Proteobacteria. They were also detected in other groups of bacteria such as \textit{Deferribacteres}, \textit{Firmicutes}, \textit{Bacteroidetes}, \textit{Spirochaetes}, \textit{Chlorobi, Actinobacteria}, \textit{Nitrospirae} and \textit{Cyanobacteria} (\textit{Cyanothece, Acaryochloris}). Several of these IS elements are shown in Table 1. The IS3-family dendrograms based on \textit{OrfA} (Fig. 2a) and \textit{OrfB} (Fig. 2b) sequences revealed that IS\textit{Ppy1} and related elements form a distinct subgroup of the IS3 family, in addition to the five IS3-family subgroups that have been described (Mahillon & Chandler, 1998). Therefore, we have described this novel subgroup as the IS\textit{Ppy1} subgroup.

As was shown previously by Mahillon & Chandler (1998), all IS3 elements contain a common fixed set of conserved amino acids in their catalytic domain, the DD(35)E motif, whereas each IS3 subgroup is characterized by a unique set of additional conserved residues (Fig. 2c). We found that the DD(35)E motif of the IS\textit{Ppy1} subgroup contains its own set of characteristic amino acid residues, distinguishing it from other IS3 subgroups. In particular, significant differences from other subgroups were detected in a region which contained seven amino acids downstream from the E residue (‘EK region’), and in the D(35)E region (Fig. 2c).

It should be pointed out that many IS elements belonging to IS\textit{Ppy1} subgroup were identified by other authors as members of IS150-subgroup (see legend to Fig. 2). However, careful analysis of molecular structure of these elements, particularly the amino acid composition of their catalytic domain, unambiguously testifies their belonging to the IS\textit{Ppy1} subgroup.

Interestingly, most of the elements from the IS\textit{Ppy1} subgroup analysed here contained a frameshifting box, located near the 3’ end of \textit{orfA} with an organization that is not typical of IS elements (Baranov et al., 2002). In addition to boxes containing additional adenines, (A)\textsubscript{6}G or (A)\textsubscript{3}G instead of (A)\textsubscript{6}G in IS\textit{Ppy1} and IS\textit{Cysp23}, we detected motifs that had not been described previously for IS elements, including (U)\textsubscript{6}G, (U)\textsubscript{6}A, and (A)\textsubscript{6}U (Table 1).

**Distribution of IS\textit{Ppy1} subgroup elements in bacterial populations**

We studied the distribution and abundance of seven IS sequences belonging to IS\textit{Ppy1}-subgroup among bacterial genomes. The results of alignments including these elements are summarized in Table 1. One can see that these elements differ in behaviour from each other. Some of them were found only once and in one copy. In particular, the original host \textit{P. maritimus} MR29-12 contained only a single copy of IS\textit{Ppy1}, localized on the plasmid (Fig. 3a). Other IS\textit{Ppy1}-subgroup elements were found in different bacteria both in plasmids and in chromosomes, in the latter case in many copies. For instance IS\textit{Gxy1} was found to be widely distributed among acetic acid bacteria. In particular, near identical copies have been found at eight loci in the genome of \textit{Gluconobacter oxydans} 621H. IS\textit{Ab14} was detected not only in the original host, \textit{Acinetobacter baumannii} 1656-2 but also in other \textit{Acinetobacter} spp. strains, in the pKHL80 plasmid of \textit{Psychrobacter maritimus} MR29-12, and in the pMR0211 plasmid of \textit{Providencia stuartii}. 

---

**Downloaded from www.microbiologyresearch.org by**

**IP: 54.70.40.11**

Transposition of ISPpy1 and Tn5080a into the E. coli K-12 chromosome

Previously we demonstrated high transposition activity of ISPpy1 in the heterologous hosts Acinetobacter calcoaceticus BD413 and E. coli K-12 (Petrova et al., 2012). In this paper, we focused on studying the translocations of ISPpy1 into the chromosome of E. coli K-12 and on the analysis of the consequences of these events.

In the first series of experiments, translocations of the composite transposon Tn5080a containing two copies of ISPpy1 into the E. coli chromosome were examined. In this case, single plasmid-free colonies were isolated and checked for the presence of Tn5080a. The prolonged cultivation of cells in LB broth initiated two processes: (i) the loss of the plasmid, as evidenced by the appearance of CmS cells, and (ii) the insertion of the composite transposon into the chromosome, as evidenced by the appearance of CmS SmR TcR cells. Both events occurred in recA+ (JF238) and in recA- (UB5201rif) strains with high efficiency. Although the loss of the plasmid occurred more frequently in the recA strain than it did in the recA+ strain, the efficiency of

Fig. 2. IS3 family. Dendrogram based on the alignment of the amino acid sequences of predicted OrfA (a) and Orf B (b) from elements of different subgroups of the IS3 family (Mahillon & Chandler, 1998). ISDsp4, ISMch1, ISAba14, ISRj1, ISKpn11, lSMex7, ISAli5 and ISAzo18 previously were considered as IS150 subgroup elements. The subgroups are indicated by brackets. (c) DD(35)E consensus of the ISPpy1 subgroup in comparison with other IS3 family subgroups. The representation is based on a scheme suggested by Mahillon & Chandler (1998) with modifications and additions. Amino acids forming part of the conserved motif are shown as large, upper-case, bold letters, which indicate conservation within the entire IS3 family; capital non-bolded letters indicate conservation within a subgroup, and lower-case letters indicate that the particular amino acid is predominant in this subgroup. The numbers in parentheses show the distance in amino acids between the sequences of the conserved motif.
transposon insertion in both strains differed only by a small extent (Table 2).

In the second series of experiments, we studied the ability of IS\textsubscript{Ppy1} to translocate on its own into the \textit{E. coli} chromosome. The frequency of plasmid-free clones carrying copies of IS elements in the bacterial chromosome was estimated using the colony-hybridization method. Surprisingly, almost all of the plasmid-free clones tested carried IS\textsubscript{Ppy1} in their chromosome (Table 2).

**ISPpy1-mediated mutagenesis in E. coli**

We suggested that IS\textsubscript{Ppy1} like other IS elements can induce mutations upon insertion into the \textit{E. coli} chromosome. We tested the colonies containing IS\textsubscript{Ppy1} insertions for auxotrophy and antibiotic resistance. We chose the mutations resistant to rifampicin because of their phenotypic variability and the great diversity in resistance mechanisms (Tupin \textit{et al.}, 2010). Our assumption was confirmed and we succeeded in creating two libraries of mutants generated by insertions of IS\textsubscript{Ppy1} into the JF238 (\textit{recA}+) strain. The first one consisted of randomly isolated plasmid-free auxotrophic derivatives of JF238 (pBR325-Ppy1), as revealed by their inability to grow on minimal media. We found that such mutants emerged with a high frequency (which reached 1%).

The second library consisted of plasmid-free clones that were resistant to low concentrations of rifampicin. As

---

**Table 2. Translocation of Tn5060\textsubscript{a} and IS\textsubscript{Ppy1} into chromosome of \textit{E. coli} strains JF238 (\textit{recA}+) and UB5201 (\textit{recA}) strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Exp. 1</th>
<th>Exp. 2</th>
<th>Exp. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>JF238 (pKLH80.2)</td>
<td>20</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>UB5201 (pKL180.2)</td>
<td>20</td>
<td>55</td>
<td>10</td>
</tr>
<tr>
<td>JF238 (pBR325-Ppy1)</td>
<td>100</td>
<td>53</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>Total number Cm\textsuperscript{s} (%)</th>
<th>Total number Cm\textsuperscript{s} (%)</th>
<th>With insertions</th>
</tr>
</thead>
<tbody>
<tr>
<td>JF238 (pKLH80.2)</td>
<td>20</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>UB5201 (pKL180.2)</td>
<td>20</td>
<td>55</td>
<td>10</td>
</tr>
<tr>
<td>JF238 (pBR325-Ppy1)</td>
<td>100</td>
<td>53</td>
<td>48</td>
</tr>
</tbody>
</table>

*Plasmid-free colonies with insertions of Tn5060\textsubscript{a} or IS\textsubscript{Ppy1} into bacterial chromosome.
indicated in Table 3, the number of Rif\(^R\) colonies was 20 to 40 times greater than that obtained from experiments using an *E. coli* strain that carried the vector plasmid pBR325 without IS\(^{Ppy1}\). The library of Rif\(^R\) variants was also tested for auxotrophy. Up to 10% of Rif\(^R\) variants did not grow on minimal medium (see below).

**Characteristics of mutant *E. coli* strains generated by IS\(^{Ppy1}\) insertions into the chromosome**

We then performed further functional characterization of several auxotrophic variants generated by IS\(^{Ppy1}\) transpositions. The majority of the auxotrophic variants tested (seven out of eight) formed slow-growing colonies on different media. Further studies suggested that some of these strains can be regarded as pleiotropic mutants. In particular, we tested the resistance of auxotrophic strains to the antibiotics ampicillin and gentamicin, and found that some of them were resistant to both antibiotics (e.g. strain D1 from the first library of mutants, Table 4). Furthermore, many of the Rif\(^R\) strains were also resistant to ampicillin, exhibited an auxotrophic phenotype (see above), and were characterized by poor growth on complete medium (Table 4).

We determined the frequency of occurrence of revertant prototrophic colonies in five of the pleiotropic variants (A3, D1, Rf27, RF36, and Rf47) and studied the phenotype of prototrophs in comparison with the original auxotrophic variants (see Methods). Four variants among the five tested produced phenotypic prototrophs at high frequencies (from \(1.6 \times 10^{-5}\) to \(1.4 \times 10^{-5}\)). We also showed that certain prototrophs had a restored sensitivity to one or both antibiotics and were characterized by a normal growth rate (Table 4).

The number of copies of IS\(^{Ppy1}\) present on the chromosome of auxotrophic and prototrophic *E. coli* variants was determined via Southern blot hybridization. The results of this analysis showed that insertions of IS\(^{Ppy1}\) were present in the DNA of all variants tested and that most of them contained several copies of IS\(^{Ppy1}\) (up to 10) on their chromosome. The prototrophs differed from their respective auxotrophs in number and position of ISs (Fig. 3).

**IS\(^{Ppy1}\) target-DNA-sequence specificity**

Our next set of experiments was aimed at the analysis of the sequence specificity of IS\(^{Ppy1}\) insertions. Previously, we developed a method to isolate plasmids carrying transposon insertions into the plasmid *bla* gene and obtained a set of three insertions of Tn5080a and IS\(^{Ppy1}\) into pBR325 (Petrova et al., 2012). In the present work, these experiments were extended (see Materials), leading to the creation of an independent set of seven Tn5080a insertions into pGEM-5Z. The resulting 10 insertions were used for the analysis of the transposition specificity of IS\(^{Ppy1}\). We sequenced the site of insertion for all plasmids and determined both the orientation of IS\(^{Ppy1}\) and the length of direct repeats (DR). We found that all insertions were located within the *bla* gene in a relatively random fashion, and both orientations were observed in nearly equal proportions (Fig. 4a). We also found that insertions resulted in duplications of 2–5 bp in the target DNA (Fig. 4b).

**DISCUSSION**

IS3-family elements are among the most numerous and widely distributed IS elements of eubacteria and archaea. They are common in the genomes of Gram-negative and Gram-positive bacteria, and their structure, transposition mechanisms, and distribution among different bacteria have been investigated thoroughly (Mahillon & Chandler, 1998; Sigui\(e\) et al., 2006; Touchon & Rocha, 2007).

IS\(^{Ppy1}\) is a novel IS3-family element that was isolated recently from the permafrost strain of *Psychrobacter maritimus* MR29-12. We have shown in this study that IS\(^{Ppy1}\) is a representative member of a large group of IS3-family elements that are moderately related to IS\(^{Ppy1}\), termed the IS\(^{Ppy1}\) subgroup. This novel subgroup is clearly separated from other known IS3-family subgroups and is defined by deep branching in the dendrogram resulting from the cluster analysis of their OrFa and OrFB amino acid sequences (Fig. 2a, b). The IS\(^{Ppy1}\)-subgroup elements differ from other members of the IS3 family regarding several structural features, including longer terminal repeats, different sequence motifs at their ends, and a specific amino acid signature in their catalytic domains. We found that the IS elements belonging to this subgroup are broadly spread among different bacterial taxa and are located on bacterial plasmids and chromosomes. Based on the great diversity of IS\(^{Ppy1}\)-subgroup elements and their broad spreading across different bacterial species, we suggest that the common ancestor of these IS elements evolved very early in the evolutionary history of eubacteria. However, the cloning of these elements and the analysis of their transposition activities had not been performed previously. Thus, IS\(^{Ppy1}\) exemplifies the only functionally characterized member of this subgroup of IS3 elements.

We failed to find any evidence in favour of the effective transposition of IS\(^{Ppy1}\) in its original host, *P. maritimus* MR29-12. A similar trend was observed for several other elements of the IS\(^{Ppy1}\) subgroup such as IS\(^{Cysp23}\), ISA\(fe\)13, IS\(Al\)i5, IS\(^{Glo6}\) (Table 1). We found only a single copy of

---

**Table 3. Isolation of rifampicin-resistance variants from *E. coli* strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Frequency of Rif(^R) colony occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp. A</td>
</tr>
<tr>
<td>JF238 (pBR325-Ppy1)</td>
<td>(6.0 \times 10^{-5})</td>
</tr>
<tr>
<td>JF238 (pBR325)</td>
<td>(3.2 \times 10^{-5})</td>
</tr>
</tbody>
</table>
each of them. A quite distinct situation is observed for other members of IS\(^{Ppy1}\) subgroup. Particularly, we revealed several elements such as IS\(^{Aba14}\) and IS\(^{Gxy1}\) that are characterized by wide distribution in nature and clinical settings (Table 1).

The most acceptable explanation of variations in modes of IS elements distribution in our opinion is offered in the work of Wagner (2006). According to his suggestion, ‘most insertion sequences in a typical genome are evolutionarily young and have been recently acquired. They may undergo periodic extinction in bacterial lineages. By implication, they are detrimental to their host in the long run’. According to this representation, the variability in distribution of different elements is a result of two different processes: first, potentially high functional activity of the IS elements, leading to wide distribution of its copies between different bacteria; and secondly, a process of coevolution of the bacterial cell with acquired IS elements leading to elimination or degradation of its copies and genome stabilization. Both options of interaction of a bacterial cell with IS elements are demonstrated by the behaviour of the IS\(^{Ppy1}\) subgroup elements, in particular the behaviour of IS\(^{Ppy1}\).

In original host, \(P.\) maritimus IS\(^{Ppy1}\) is functionally inert suggesting a long period of coevolution. In the heterologous host, \(E.\) coli, this element shows a high functional activity resulting in strong deleterious effect. It is possible to assume that the subsequent process of stabilization of a genome will inevitably lead to elimination of the majority of IS\(^{Ppy1}\) copies.

It is necessary to note that IS\(^{Ppy1}\) subgroup elements actively participate in the emergence of composite transposons carrying various genes, including antibiotic-resistance genes, which promotes their horizontal transfer. Among a small number of IS\(^{Ppy1}\) subgroup elements that we tested we revealed two elements, IS\(^{Ppy1}\) and IS\(^{Aba14}\) possessing such ability. IS\(^{Ppy1}\) was shown to participate in the formation of the composite transposons Tn\(^{5080}\) and Tn\(^{5080a}\) (Petrova et al., 2009, 2012) while IS\(^{Aba14}\) was revealed in the strain RAB of \(A.\) baumanii as a part of the composite transposon Tn\(^{2114}\) (Bonnin et al., 2012).

**Table 4. Phenotypic characters of different mutant derivatives of E. coli JF238 generated by IS\(^{Ppy1}\)**

<table>
<thead>
<tr>
<th>Strain*</th>
<th>Growth on minimal medium</th>
<th>Resistance to†</th>
<th>OD‡</th>
<th>Prototroph frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>JF238</td>
<td>+</td>
<td>S S S</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>–</td>
<td>S R R</td>
<td>1.8</td>
<td>(1.6 \times 10^{-5})</td>
</tr>
<tr>
<td>D1-P</td>
<td>+</td>
<td>S S S</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>–</td>
<td>S S S</td>
<td>1.8</td>
<td>ND</td>
</tr>
<tr>
<td>A3-P</td>
<td>+</td>
<td>S S S</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Rf27</td>
<td>–</td>
<td>R R R</td>
<td>1.4</td>
<td>(1.4 \times 10^{-3})</td>
</tr>
<tr>
<td>Rf27-P1</td>
<td>+</td>
<td>S/R S</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Rf27-P2</td>
<td>+</td>
<td>S S S</td>
<td>1.4</td>
<td>(1.3 \times 10^{-4})</td>
</tr>
<tr>
<td>Rf36</td>
<td>–</td>
<td>R R R</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Rf36-P</td>
<td>+</td>
<td>S S S</td>
<td>1.5</td>
<td>(6.9 \times 10^{-5})</td>
</tr>
<tr>
<td>Rf47</td>
<td>–</td>
<td>R R S</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Rf47-P</td>
<td>+</td>
<td>R S S</td>
<td>1.6</td>
<td></td>
</tr>
</tbody>
</table>

ND, Not determined.

*Strains from the first library; Rf27, Rf36, Rf47, strains from the second library; strains with P, P1, P2, prototrophic revertants.
†R, Resistant; S, sensitive; S/R, intermediate.
‡Growth on LB for 6 h at 30 °C (Methods D1 and A3).

**Fig. 4.** IS\(^{Ppy1}\) target-DNA-sequence specificity. (a) Positions of IS\(^{Ppy1}\) and Tn\(^{5080a}\) insertion sites in \(bla\) gene of pBR325 and pGEM-5Zf(−). Horizontal arrows indicate insertion sites and orientation of the mobile element: white, IS\(^{Ppy1}\); black, Tn\(^{5080a}\). Numbers above arrows indicate clone numbers. Numbers beneath arrows indicate clone numbers according to the pBR325 sequence (L08855). (b) Nucleotide sequences around target sites of the insertions. Numbers at the left indicate clone numbers; the sequences of direct repeats are shown in the centre.

**M. Petrova and others**

1908

Microbiology 159

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
To understand the functional properties of ISPpy1 in greater detail, we studied the transposition and mutagenic activity of ISPpy1 and its target-DNA-sequence specificity in *E. coli*. We found that ISPpy1: (i) does not show a distinct target-DNA-sequence specificity, in contrast with that of many other IS3-family elements (Clément et al., 1999; Craig, 1997; Tobes & Pareja, 2006), and (ii) produces auxotrophic and pleiotropic mutants containing multiple ISPpy1 copies and inserts into different sites of the bacterial chromosome.

During this study, we also isolated ISPpy1-induced mutants characterized by resistance to rifampicin. Resistance to rifampicin occurs most frequently by mutations in the *rpoB* gene, which encodes the β subunit of RNA polymerase (Ovchinnikov et al., 1983). However, other mechanisms of Rif resistance have been reported, including resistance by decreasing the permeability of the membrane barriers (Hui et al., 1977) and by involvement of efflux pumps in the detoxification of multiple antibiotics (Hagman et al., 1995; Siddiqi et al., 2004; Tupin et al., 2010). These mechanisms may be responsible for the emergence of Rif resistance mutations generated by the ISPpy1 insertions. Many of the Rif resistant strains obtained in this study are pleiotropic mutants also characterized by resistance to other antibiotics, low growth rates, and inability to grow on minimal medium. In addition, these auxotrophic mutants can generate prototrophic revertants that are sensitive to rifampicin and display no growth defects. Interestingly, in most auxotroph–prototroph pairs, prototrophs contained more copies of ISPpy1 than did the parent auxotrophic strain. It can be suggested that excision of ISPpy1 is likely to permanently render the original insertion site inactive due to the DR of 2–5 bp. That is why most of the characterized prototrophs contained additional insertions, and are most likely due to mutations (including suppressor mutations) generated by additional ISPpy1 insertions (Fig. 3a, b).

The recent discovery of ISPpy1 in a *Psychrobacter* clinical strain suggests that this and other ISPpy1-subgroup elements may have a role in horizontal gene transfer from environmental to clinical strains and in genetic diversity in clinically relevant bacteria. Further studies are required to identify the possible mutagenic potential of these elements in the environment and in clinical settings.

**ACKNOWLEDGEMENTS**

We are grateful to A. V. Kulbachinsky and two anonymous reviewers for providing invaluable comments and suggestions; N. A. Khachikian and E.I. Molchanova for expert technical assistance; V.A. Shcherbakova for providing the set of *Psychrobacter* strains and for help in the identification of *Psychrobacter marinus* MR29-12 strain. This work was partially supported by the Russian Foundation for Basic Research, grants 08-04-00263 and 11-04-01217 and by the Russian Academy of Sciences Presidium Programmes ‘Wildlife: Current State and Development’ (grant to M. P.) and ‘Molecular and Cellular Biology’ (grant to A. Kulbachinsky).

**REFERENCES**


Edited by: E. Hartland