A comparison of the multiple alleles of \textit{xyl}S carried by TOL plasmids pWW53 and pDK1 and its implications for their evolutionary relationship

SUSAN J. ASSINDER, PAOLO DE MARCO,† DAVID J. OSBORNE,‡ CHIT LAA POH,§ LINDA E. SHAW, MICHAEL K. WINSON¶ and PETER A. WILLIAMS*

School of Biological Sciences, University of Wales, Bangor, Gwynedd LL57 2UW, Wales, UK

(Received 26 June 1992; revised 18 September 1992; accepted 2 November 1992)

Both of the independently isolated TOL plasmids pWW53 and pDK1 contain multiple regions homologous to the \textit{xyl}S regulatory gene of the archetypal TOL plasmid pWW0. The three homologues on pWW53 vary in the extent of their homology to \textit{xyl}S\textsubscript{pww0}. \textit{xyl}S\textsubscript{pww3} is 99\% identical to \textit{xyl}S\textsubscript{pww0} and is located relative to the single copy of \textit{xyl}R\textsubscript{pww3} in exactly the same way as \textit{xyl}S and \textit{xyl}R on pWW0. The DNA sequence of \textit{xyl}S\textsubscript{pww3} is 87\% identical to the \textit{xyl}S\textsubscript{pww0} sequence within the coding region but the non-coding DNA upstream is not homologous. There is a frame-shift change at the end of the coding region which causes the C terminus of \textit{xyl}S\textsubscript{pww3} to be extended by an additional 10 amino acids relative to \textit{xyl}S\textsubscript{pww0}. \textit{xyl}S\textsubscript{pww5} is anomalous and appears to encode a truncated pseudogene lacking the first 525 bases found in the other \textit{xyl}S genes. Evidence is presented to show that both \textit{xyl}SI\textsubscript{pww3} and \textit{xyl}S\textsubscript{pww3} act as regulators of \textit{meta} pathway operons. Plasmid pDK1 carries two homologues of \textit{xyl}S. \textit{xyl}SI\textsubscript{pdk1} is functional and is a hybrid gene: its 5' end and the upstream sequences are highly homologous to both \textit{xyl}SI\textsubscript{pww3} and \textit{xyl}S\textsubscript{pww0}, whereas its 3' end is identical to \textit{xyl}S\textsubscript{pww3}. The sequence of \textit{xyl}S\textsubscript{pdk1} is identical to that of the anomalous truncated \textit{xyl}S\textsubscript{pww3}. Comparison of the organization and the restriction maps of the \textit{xyl} catabolic operons on pDK1 and pWW53, together with the nucleotide sequences presented here, indicates that the catabolic DNA on pDK1 has derived from a replicon on which the \textit{xyl} genes are organized similarly to pWW53 and that a genetic rearrangement has taken place involving a reciprocal recombination internal to two of its \textit{xyl}S homologues.

Introduction

The molecular biology of the TOL pathway for the catabolism of toluene and the xylenes has been characterized in greatest detail on the 117 kbp plasmid pWW0 from \textit{Pseudomonas putida} mt-2. However, a number of other plasmids encoding the same catabolic functions have been reported (see Assinder & Williams, 1990) and the gene organization on two, pWW53 (Keil \textit{et al.}, 1985\textit{a, b}; Osborne \textit{et al.}, 1988) and pDK1 (Shaw & Williams, 1988) has been described in some detail. This evidence suggests that the genetic organization of the \textit{xyl} catabolic genes found on pWW0 (Franklin \textit{et al.}, 1981), namely single operons for the upper pathway (Harayama \textit{et al.}, 1989), and the \textit{meta} pathway genes (Harayama & Rekik, 1990), is not typical of other TOL plasmids. Not only are the \textit{xyl} operons arranged in different orders and orientations but there is increasing evidence that many carry more than one copy either of whole operons or of individual genes (Keil \textit{et al.}, 1985\textit{b}; Chatfield & Williams, 1986; O'Donnell & Williams, 1991). In particular we have shown that the 107 kbp plasmid pWW53 carries two homologous, but distinguishable, direct repeats of the \textit{meta} pathway operon \textit{xylXYZLTFJQIH} both of which are functional (Osborne \textit{et al.}, 1988). In this paper we show the...
existence on two TOL plasmids of multiple copies of the xylS regulatory gene, responsible for the activation of the meta pathway operon by benzoate or methyl-benzoates (tolulates). Consideration of the nucleotide sequences, together with the overall xyl gene organization, suggests that plasmid pDK1, which was isolated by Kunz & Chapman (1981) in Minneapolis-St Paul, USA, appears to have derived from a major recombinational rearrangement from a plasmid similar in its organization, suggests that plasmid pDK1, which was isolated by Kunz & Chapman (1981) in Minneapolis-St Paul, USA, appears to have derived from a major recombinational rearrangement from a plasmid similar in its organization.  

Methods

**Bacterial strains and plasmids.** Bacterial strains and plasmids used are listed in Table 1. Bacteria were grown and maintained according to Worsey & Williams (1975) and Keil et al. (1985a, b).

<table>
<thead>
<tr>
<th>Strain</th>
<th>Plasmid</th>
<th>Notes and reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pseudomonas</em> strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PaW130</td>
<td>pWW53</td>
<td>Keil et al. (1985b)</td>
</tr>
<tr>
<td>MT53</td>
<td>pWW53</td>
<td>Keil et al. (1985a)</td>
</tr>
<tr>
<td>MT53-1</td>
<td>pWW53-1</td>
<td>Osborne et al. (1988)</td>
</tr>
<tr>
<td>MT53-2</td>
<td>pWW53-2</td>
<td>Osborne et al. (1988)</td>
</tr>
<tr>
<td>HSI</td>
<td>pDK1</td>
<td>Wild type isolate (Kunz &amp; Chapman, 1981)</td>
</tr>
<tr>
<td>PpCM1</td>
<td>pDKM1</td>
<td>Benzolate-selected segregant of HSI (Kunz &amp; Chapman, 1981)</td>
</tr>
</tbody>
</table>

| Recombinant plasmids from pWW53 | | |
| pWW53-3002 | Osborne et al. (1988) |
| pWW53-3004 | Osborne et al. (1988) |
| pWW53-3508 | Keil et al. (1987b) |
| pWW53-3510 | 6.7 kbp HindIII-H of pWW53 in pKK220 |
| pWW53-1013 | 14.0 kbp KpnI-C of pWW53 in pKK220 |
| pWW53-1201 | 53 kbp Smal-KpnI fragment obtained by Smal digestion of pWW53-1013 |
| pWW53-1001 | 3.2 kbp BamHI fragment of pWW53-3010 cloned into pFBI.1 (xylS2,pww53) |
| pWW53-1002 | 3.2 kbp BamHI fragment of pWW53-1001 cloned into pWW53-3002 |
| pWW53-1003 | 3.2 kbp BamHI fragment of pWW53-1001 cloned into pWW53-3004 |

| Recombinant plasmids from pDK1 | | |
| pDK2-3351 | HindIII-S in pUC18 in opposite orientation to 3351 |
| pDK2-3352 | HindIII-S in pUC18 in opposite orientation to 3351 |
| pDK2-3353 | 1.1 kbp HindIII-U in pUC18 |
| pDK2-3376 | 1.45 kbp XhoI-C in pUC18 |
| pDK2-3377 | XhoI-C in pUC18 in opposite orientation to 3376 |
| pDK2-3364 | 3.9 kbp BamHI-D in pUC18 |

**DNA manipulations.** Vector and recombinant plasmid DNA isolation, construction of recombinant DNA, Escherichia coli transformation, nick-translation to radiolabel DNA with 32P and DNA–DNA hybridizations were by standard procedures (Sambrook et al., 1989), pWW53, pDK1 and their derivatives were extracted according to Wheatcroft & Williams (1981).

**DNA subcloning strategies and sequencing.** Sequencing was performed on both DNA strands using the double-stranded dideoxy chain-termination procedure (Sanger et al., 1977) either with M13 forward and reverse primers, or with custom-synthesized oligonucleotide primers. Nested deletions of clones for sequencing were obtained using S1 nuclease and exonuclease III (Henikoff, 1984) using a Pharmacia kit.

(i) Sequencing of xylS1,pww53. The 3.9 kbp XhoI fragment of pWW53 contained in pWW53-3508 (Keil et al., 1987b) was used as a source of xylS1 for DNA sequencing. Contiguous 0.6 kbp BglII and 1.2 kbp BglII-Smal fragments were subcloned into pUC18 in both orientations.

(ii) Sequencing of xylS2,pww53. The 1.40 kbp KpnI fragment KC of pWW53 was cloned into pUC19 giving pWW53-1013. Internal Smal fragments were excised by digestion with Smal and religation to generate pWW53-1201. This carries the xylS2,pww53 gene on a 5.3 kbp Smal–KpnI fragment.

(iii) Sequencing of xylS3,pww53. The 3.2 kbp BamHI fragment carrying xylS3,pww53 in pWW53-1001 was subcloned by transferring the 2.0 kbp BamHI–BglII fragment into the BamHI site of pUC19 in both orientations.

(iv) Sequencing of xylS1,pdk1. Determination of the sequence of xylS1 started with 3.9 kbp BamHI-D. This was cloned into pUC19 in both orientations to give plasmids pDK2-3364 and pDK2-3365. A small (0.7 kbp) Smal fragment was deleted from pDK2-3364 between the Smal site in the multicloning site of the vector and the internal Smal site in the insert. Similarly a PstI fragment was excised from pDK2-3365 using the multicloning site and an internal site in the insert leaving a 2.4 kbp PstI–BamHI insert.

(v) Sequence of xylS2,pdk1. The sequence of xylS2 was initiated from three fragments cloned into pUC18, the adjacent 1.35 kbp HindIII-S (pDK2-3351 and pDK2-3352) and 1.1 kbp HindIII-U fragments (pDK2-3353), and the overlapping XhoI-C fragment (pDK2-3376 and pDK2-3377).

Results

**Identification of three xylS homologues on pWW53**

A 300 bp fragment internal to the xylS,pww53 gene (gift of F. C. H. Franklin, University of Birmingham, UK) was radiolabelled and hybridized to pWW53 DNA cut with a range of restriction endonucleases. Two regions carrying putative xylS homologues were identified in addition to that already shown to exist at the end of operon 1 (to be referred to as xylS1,pww53) (Keil et al., 1987b; Osborne et al., 1988). One of these regions, denoted xylS2,pww53, was at the downstream end of...
Multiple xylS genes on TOL plasmids

Fig. 1. Restriction maps of the areas containing the xyl genes of TOL plasmids pDK1 and pWW53. The maps are aligned to highlight the structural relationships between the two plasmids. The restriction map for XhoI and HindIII for pDK1 is taken from Shaw & Williams (1988) and represents the pDK1 insert in the RP4:pDK1 cointegrate pDK2. The restriction map for pWW53 has not been published previously in this form but some of the data is from Osborne et al. (1988). The two maps have been aligned on the two upper pathway operons (xylCMABN; upper). The extent of homology around this region is shown by the two bold arrows. The single HindIII site in this region which is present on pWW53 but not on pDK1 is shown as a dashed line. The second region of homology extends between xylS3 (S3) past xylS2 (S2), including meta pathway operon 2 (meta 2) on pWW53 and from xylS1 (S1) and past xylS2, including the meta pathway operon 1 (meta 1) on pDK1 and is denoted by the bold dashed arrows: there are no differences in restriction sites for nine enzymes within this region. The three xylS homologues which are proposed as being involved in the rearrangement of catabolic DNA from a pWW53-like plasmid to give pDK1 are shown as cross hatched boxes: xylSZpDK1. The two xylS2 homologues which have been shown to have identical sequences are shown as black boxes. The two meta pathway operons which are indistinguishable by restriction mapping and presumably have very similar, if not identical sequences, are shown as shaded boxes. The directions of transcription of the various transcription units are shown as light arrows. Relevant restriction fragments mentioned in the text are lettered. The extent of the deletions in pWW53 which have occurred in the formation of the plasmids pWW53-1 and pWW53-2 are shown above the pW53 map. R, xylR.

A 1.5 kbp BglII-SmaI fragment of pWW53 HindIII-A (HA), shown by Keil et al. (1987b) to carry part of xylR of pWW53 was used as a probe against digests of DNA from pWW53, pWW53-1 and pWW53-2. It hybridized only to the region adjacent to xylS1 of pWW53 at the downstream end of operon 1 on pWW53, implying that xylR is present in a single copy on pWW53. As confirmation of this, no hybridization was observed to the deletion derivatives of pWW53, pWW53-1 and pWW53-2, since the downstream end of operon 1 is deleted in these plasmids (Fig. 1).

Functional analysis of xylS3 of pWW53

The ability of xylS3 of pWW53 to act as a regulator of meta pathway operons was examined using the following strategy (Fig. 2). HindIII fragment H from pWW53 which carries xylS3 (Fig. 1) was cloned into pKT230 (pWW53-3010). Using the unique vector BamHI site, the 3.2 kbp BamHI fragment was subcloned into pFBI. The BamHI fragment from pWW53 was cloned into pWW53-3002 (consisting of pWW53 HB cloned into the HindIII site of pKT230; Osborne et al., 1988) to give plasmid pWW53-1002. This construct allowed us to study the interaction between xylS3 of pWW53 and the promoter of meta pathway operon 2, Pm2 of pWW53. It was not possible to make an exactly
Fig. 2. Cloning strategy for the functional analysis of xylS3pw53. The sizes of the vectors are indicated within the circles. The location of xylS3pw53 is shown as a black box. The location of pWW53 meta operon 2 is shown as a hatched box. H, HindIII; B, BamHI. The origins of the plasmids and accurate sizes of the cloned inserts are given in Table 1.

Table 2. Specific activities of meta pathway enzymes in Paw130 carrying the xylS3pw53 gene and either the hybrid meta pathway operon from pWW53-2 (HA') or meta operon 2 from pWW53 (HB)

<table>
<thead>
<tr>
<th>xylS3 operon</th>
<th>meta operon</th>
<th>m-Toluate</th>
<th>Specific activity*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m-Toluate</td>
<td>C23O</td>
<td>HMSD</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>HA'</td>
<td>36</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>HA'</td>
<td>139</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>H B</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>H B</td>
<td>17</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>H B</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>H B</td>
<td>183</td>
<td>130</td>
</tr>
</tbody>
</table>

* Specific activities are expressed as milliunits (mg protein)-1 and represent the mean of three replicates. C23O, catechol 2,3-oxygenase; HMSD, 2-hydroxymuconic semialdehyde dehydrogenase; HMSG, 2-hydroxymuconic semialdehyde hydrolase.

Identification of two xylS homologues on pDK1

In our earlier paper on pDK1 (Shaw & Williams, 1988) we demonstrated that at the upstream end of the upper pathway operon xylCMABN of pDK1 there were two adjacent BglII sites as found on pWW0 in the centre of the divergently transcribed xylRS gene pair (Spooner et al., 1986). These sites were located on a 3-9 kbp BamHI fragment D (BD) and the overlapping 2-8 kbp XhoI fragment S (XS) (around coordinates 25–27, Fig. 1). Both fragments hybridized with the xylSpwo-specific probe (see above) and, when BD was inserted into a recombinant plasmid containing the cloned pDK1 meta pathway operon, the meta pathway enzymes were inducible by m-toluate (Shaw & Williams, 1988). BD appears therefore to carry a functional xylS gene which we have termed xylS1pdk1.

Plasmid pDKM1 is a spontaneous deletion derivative of pDK1 (Kunz & Chapman, 1981) which carries an inducible meta pathway and supports growth on m- and p-toluates but has lost the upper pathway operon as part of its deletion. Restriction digests showed that the fragments around the xylS1pdk1 gene were not present on pDKM1 (results not shown) having been deleted with the adjacent upper pathway operon and its xylR gene (see Fig. 1). However, the xylSpwwo-specific probe did hybridize to Southern blots of digested pDKM1 at positions corresponding to the 3-8 kbp XhoI fragment J (XJ) and the overlapping 1-35 kbp HindIII–fragment S (HS); we have termed this region of homology xylS2pdk1.

Sequence determination of the pWW53 and pDK1 xylS genes

For clarity only the DNA sequences corresponding to the ORF of xylSspwo plus a short upstream region are...
Multiple xylS genes on TOL plasmids

Fig. 3 (for legend see page 563)
Fig. 3 continued (for legend see facing page)
Multiple xylS genes on TOL plasmids

A total of 1603 bp around xylS1pWW53 has been sequenced, starting at the BglII site within xylIRpWW53 and extending 86 bp beyond the end of the XylS encoding region. The putative ORF gives a protein of 321 amino acids with an $M_t$ of 36580. For xylS2pWW53, 1447 bp were sequenced, similarly commencing at a BglII site upstream of the gene. Optimum alignment with xylS1pWW53 shows homology only with the 3' end of the gene within which region there is only one potential XylS-like ORF (Fig. 3) of 157 amino acids ($M_t = 17894$). A sequence of 1346 bp surrounding xylS3pWW53 was determined in which there is a single ORF of 331 amino acids ($M_t = 37845$).

In the case of xylS1pWW53, 1679 bp were sequenced, with a putative ORF of 331 amino acids ($M_t = 37924$). The sequence of xylS2pWW53 was identical to that of xylS2pWW53 and therefore encodes a possible truncated XylS-like protein of 157 amino acids ($M_t = 17894$).

Discussion

Regulation of meta pathway operons on pWW53

The arrangement of catabolic genes on TOL plasmid pWW53 is clearly more complex than that on pWW0 (Franklin et al., 1981). In addition to the large duplication of the meta pathway operon identified previously (Osborne et al., 1988), there exist three homologues of xylS1 and xylS3. Two of these (xylS1 and xylS3) have definitively been shown to encode positive regulator proteins for the meta pathway genes. The expression of the meta pathway enzymes must involve interactions between these regulatory proteins and the operator–promoter regions of the two meta pathway operons. Although both meta pathway operons are functional, there are major differences in the specific activities of their encoded enzymes: the specific activities encoded by operon 1 are approximately 10-fold higher than those on operon 2, even when both operons are cloned on
Fig. 4. Comparison of the deduced amino acid sequences of the XylS proteins from pWW53 and pDK1 with the published sequences of the proteins from pWW0 (Inouye et al., 1986; Spooner et al., 1986). Colons indicate identity with the homologous pWW0 protein.

equivalent HindIII fragments downstream of an identical vector promoter (Osborne et al., 1988). It has been suggested that the lower activities of the operon 2 enzymes reflect intrinsic differences in the kinetic properties of the enzymes, rather than a difference in the expression of the two operons (Williams et al., 1990).
DNA sequencing of the two operon promoters, Pm1pww53 and Pm2pww53, as part of this study (data not shown) supports this view by demonstrating that both promoters are very similar and both highly homologous to Pm<sub>pww0</sub>.

Functional analysis of the pWW53 and pDK1 xylS genes

The regulatory function of xylS<sub>1</sub><sup>pww53</sup> has been demonstrated previously by the inducibility of the meta pathway enzymes in a P. putida strain carrying HA cloned into pKT230 upon which it is the only homologue (Keil et al., 1985a). The activity of xylS<sub>2</sub><sup>pww53</sup> was demonstrated in this study by the inducibility of meta pathway enzymes in strains carrying this gene in conjunction with either HB of pWW53 or the novel HindIII fragment HA′ from pWW53-2 (Table 2). The data are consistent with xylS<sub>2</sub><sup>pww53</sup> encoding a functional gene product which can interact with the promoters of both meta pathway operons.

The regulatory function of xylS<sub>1</sub><sup>pdk1</sup> has been demonstrated in an earlier study (Shaw & Williams, 1988).

It is impossible to state with any certainty whether the xylS2 genes from either plasmid can act as functional activators. In both cases there is very indirect evidence which points to possible functionality. This comes from deleted plasmids generated in vivo in which the xylS2 genes are the only homologues present. In the case of plasmid pWW53 the deleted derivatives are pWW53-1 and pWW53-2 (see Fig. 1; Osborne et al., 1988), and for pDK1 the deletion is in plasmid pDKM1 (Kunz & Chapman, 1981). In P. putida strains carrying these plasmids the meta pathway genes are induced by benzoate or toluates (Osborne et al., 1988; Kunz & Chapman, 1981). However, we have as yet been unable to demonstrate that either of the cloned xylS2 genes can function as benzoate- or toluate-stimulated activators in an expression system. Furthermore, a recent paper has demonstrated that a gene in the P. putida chromosomal, which is probably benR, can substitute for xylS in such an expression system although only with unsubstituted benzoate as inductor (Jeffrey et al., 1992). It is possible that the inducibility of the meta pathway genes on pWW53-1, pWW53-2 and pDKM1 could be due to this chromosomal regulator gene.

Sequence analysis of the pWW53 and pDK1 xylS genes

There is striking homology between xylS<sub>1</sub><sup>pww53</sup> and xylS<sub>pww0</sub> at both the DNA (99%) (Fig. 3) and amino acid (99%) sequence level (Fig. 4).

For xylS<sub>2</sub><sup>pww53</sup>, there is a high level of DNA (87%) and amino acid (83%) sequence homology with xylS<sub>pww0</sub> within the coding portion of the gene, but there is significant divergence in the 5′ leader sequence (Fig. 3) and downstream of the ORF (data not shown). The upstream breakpoint in homology starts only about 23 bases upstream of the presumed translational start codon. There is a frame-shift mutation towards the end of the coding region (base 1146) which causes the C terminus of the XylS<sub>pww0</sub> protein to be elongated by an additional 10 amino acid residues relative to XylS<sub>pww0</sub>. The deduced amino acid sequence of XylS<sub>1</sub><sup>pdk1</sup> is the same length (331 amino acids) as XylS<sub>3</sub><sup>pww53</sup> and has the same C-terminal extension.

The DNA upstream of xylS<sub>1</sub><sup>pww53</sup> and xylS<sub>1</sub><sup>pdk1</sup> is highly homologous to that of xylS<sub>pww0</sub> to the extent that it is possible to identify the −12, −24 promoter (Fig. 3) and a number of the other regulatory elements which have been demonstrated or proposed for xylS<sub>pww0</sub> (data not shown). Within the non-homologous upstream region of xylS<sub>3</sub><sup>pww53</sup>, there is a sequence showing homology to the −12, −24 promoter of xylS<sub>pww0</sub> (Fig. 3) but no evidence has been obtained as yet to demonstrate its functional significance.

The two xylS2 sequences are identical and both present the same problem of interpretation. There is 82% identity between the xylS2 region and the pWW0 ORF downstream of base 420 but no homology upstream (Fig. 3). The possibility that xylS2 encodes a single XylS protein with a divergent N terminus is ruled out by the presence of an in-frame TAA stop codon (bases 344–346) (Fig. 3). In the region that is homologous to xylS<sub>pww0</sub>, there is an ORF which runs from bases 719 to 1192 with a potential ribosome-binding site (RBS) centred 10 bp upstream of the ATG which shows good homology with other Pseudomonas RBS sequences (Spooner et al., 1986). This ORF would give rise to a protein of 157 amino acids homologous to the C-terminal portion of XylS<sub>pww0</sub> but lacking its N-terminal half. However, if this does encode a truncated version of XylS, it is surprising that there is still such marked homology with the xylS<sub>pww0</sub> sequence upstream of its start codon. It might be expected that any similarity in DNA sequence would have been lost outside of the shortened coding region, unless the divergence of xylS2 from xylS<sub>pww0</sub> was very recent in evolutionary terms. Analysis of both xylS2 sequences upstream of their regions of homology to the pWW0 xylS ORF has shown the presence in each case of a complete ORF with a high degree of homology to a number of recombinase and transposon resolvase genes (Assinder et al., 1992). At present the most probable explanation is that the two xylS homologues represent the residues of complete xylS genes into which a transposon or insertion sequence has been inserted recently.
The presence of xylIR genes on pDK1 and pWW53

Hybridization results have suggested that both plasmids (Keil et al., 1987; Shaw & Williams, 1988) contain only one region that is homologous to xylIRwpw0. In both plasmids this region is adjacent to their xylIS1 genes. This is confirmed by the sequences upstream of both xylIS1pw53 and xylIS1pdk1 (data not shown). In both there is the start of an ORF which is highly homologous to the start of xylIRpw53 and similarly reads in the opposite direction. No such homology is found upstream of the other xylIS homologues.

The relationship between pDK1 and pWW53

The major point of interest concerning the two xylIS homologues on pDK1 comes when we compare them with the three homologues from pWW53, taking into account the genetic organization and restriction maps of the two plasmids (Fig. 1). The plasmids share two large areas which have identical or virtually identical restriction maps. The meta pathway operon (xylXYZLATEGFJQKIH) of pDK1 has been cloned on a HindIII fragment (HA, Fig. 1) measured at 15.5 kbp (Shaw & Williams, 1988). One of the two meta pathway operons on pWW53, referred to as meta pathway operon 2, is found on HindIII fragment HB, measured at 15.6 kbp (Osborne et al., 1988). The two fragments have in common 48 restriction sites for eight different enzymes. Furthermore, the identity of the two restriction maps extends about 6 kbp upstream of the fragment carrying the operon and about 5 kbp downstream, taking in a further 25 restriction sites (Fig. 1). Within a total length of about 26 kbp we have found no difference in the restriction fragmentation of the two plasmids. This paper has shown the identity at the nucleotide level for the xylIS2 regions which are internal to these long indistinguishable regions on the two plasmids (Fig. 3).

The second region of homology shared by pWW53 and pDK1 covers the 7 kbp of their two upper pathway operons (xylCMABN) together with about 2 kbp downstream and 4 kbp upstream, which includes the xylIR genes on each plasmid. Within these regions they share 33 common sites for nine restriction enzymes. In this case, however, there is a single difference as pWW53 has one additional HindIII site towards the end of the operon (Fig. 1) which is not present on pDK1. It could be that this homology extends further downstream of the operons but we have not mapped pDK1 any further in that direction.

The two homologous regions shared by pDK1 and pWW53 differ in their relative orientations on the two plasmids and on pDK1 they are contiguous whereas on pWW53 they are separated (Fig. 1). This can be explained when the sequence of xylIS1pdk1 (Fig. 3) is examined. The gene is located at the boundary of the two regions of homology which pDK1 shares with pWW53 (Fig. 1). Comparison of its nucleotide sequence with those of xylIS1pw53 and xylIS3pw53 reveals that it appears to be a hybrid of the two (Fig. 5). From the first common nucleotide sequenced up to approximately base 1340 (as numbered on Fig. 5) the sequences of xylIS1pdk1 and xylIS1pw53 are highly homologous. Within the coding

![Fig. 5. Comparison of the nucleotide sequence of xylIS1pdk1 and the flanking DNA with the homologous DNA of (a) xylISpw0, (b) xylIS1pw53 and (c) xylIS3pw53. The sequences were optimally aligned using DNASIS (Pharmacia). The number of mismatches between the two sequences was determined within a 60 bp window of the xylIS1pdk1 sequence. Successive bar lines represent the windows displaced along the sequence by 30 bp and the bar is drawn at a position on the corresponding to the mid-point of the 60 bp window. The numbering on the x-axis corresponds to the numbering of the bases in the Genbank file for xylIS1pdk1, which extends 440 bases upstream of the beginning of the sequence as presented in Fig. 3(a) and 633 bases upstream of its ORF. All three comparisons extend to base 1650, 24 bp after the end of the ORF. The arrows represent the point at which it was possible to start the comparison because of the length of the sequence information available on the other strands: (b) starts at base 121, and (c) at base 361. The locations of xylIRpw0 (R) and xylISpw0 (S) and their intergenic region is shown at the top of (a), and the direction of transcription is indicated by the horizontal arrows. The compared sequence for (a) is taken from Spooner et al. (1986).]
region there is an even distribution of base differences of around 0–3 bases per 60 nucleotides. In the intergenic region between *xylR* and *xylS* there is a higher level of difference which is due almost entirely to two separate insertions of 8 bp which are found only in the *xylRSI*<sub>pww53</sub> sequence. In fact the sequence of *pDK1* DNA in this region and for most of the coding region of *xylSI* is even closer to that of *xylS*<sub>pww0</sub> DNA (Fig. 5a) than that of *xylSI*<sub>pww53</sub> (Fig. 5b).

Over the same region (up to about base 1340) there is a much-lower degree of homology between the *xylSI*<sub>pdk1</sub> DNA and the DNA around *xylS*<sub>pww53</sub> (Fig. 5c). Upstream of the XylS ORFs there is no significant homology with base differences of around 35 per 60 bases: this agrees with the apparent absence of a XylR gene upstream of *xylS*<sub>pww53</sub>. Within the ORFs there are approximately 5–10 differences per 60 bases, showing a substantially lower degree of homology than found when the comparison is with the coding sequences of *xylSI*<sub>pww53</sub> and *xylSI*<sub>pdk1</sub> (Fig. 5b). The 5' end of the *xylSI*<sub>pdk1</sub> ORF is therefore very similar to both *xylS*<sub>pww0</sub> and *xylSI*<sub>pww53</sub> (but marginally closer to the former) and is significantly different to *xylS*<sub>pww53</sub>. However, from about base 1340 the level of homology with both *xylSI*<sub>pww53</sub> and *xylS*<sub>pww0</sub> drops (Fig. 5a, b) to less than 90% until at the end of the ORF and into the short stretch of the downstream regions which have been sequenced it becomes insignificant. Over this same region, *xylS*<sub>pww53</sub> and *xylSI*<sub>pdk1</sub> are identical (Fig. 5c).

The hybrid nature of the *xylSI*<sub>pdk1</sub> gene (*xylSI*<sub>pww53⋅xylS*<sub>pww53</sub>) exactly parallels its location relative to the two large regions of homology which are shared by *pDK1* and *pWW53*. Whereas the 5' end of the gene is adjacent to a XylR gene and is upstream of the upper pathway operon, in exactly the same way as *xylSI*<sub>pww53</sub>, its 3' end is located identically to *xylS*<sub>pww53</sub>; both genes being upstream of regions carrying indistinguishable *meta* pathway operons.

Both the hybrid nature of *xylSI*<sub>pdk1</sub> and its location can be explained if *pDK1* is derived from a plasmid similar to *pWW53*. A single reciprocal recombination event between two copies of *xylS* on *pWW53* (in this case *xylSI* and *xylS3*) in the vicinity of base 900 (Fig. 3) (or codon 232, Fig. 4) would produce a gene with the structure of *xylSI*<sub>pdk1</sub> and would also cause a substitution of a *meta* pathway operon like (xy*ZXZTYLFKGJQKIH*), in place of and in the opposite orientation to (xy*XYZZTGFJQKIH*), giving exactly the gene arrangement found on *pDK1* (Fig. 1).

Two possible events could lead to the structure of *pDK1*. The recombination could occur either within a plasmid like *pWW53* or between two TOL plasmids co-existing within the same strain. In either case, one ancestor of *pDK1* would require more than one copy of a *xylS* gene to give the structure found. It had occurred to us that such a recombination event might have taken place in our laboratory since both strains were being studied on adjacent benches: this would make a weaker case for the occurrence of such an event being relevant to evolution in the wild. However, one argument can be brought to bear against the formation of *pDK1* being artefactual. The 5' end of *xylSI*<sub>pdk1</sub> is unique and is not identical to the 5' ends of either *xylSI*<sub>pww53</sub> or *xylS*<sub>pww0</sub> (Figs 4 and 5). Furthermore, the entire sequence of the *meta* pathway operon and the *xylRSI* region of *pDK1* has been independently determined and the two sequences are identical (R. C. Benjamin, personal communications).

Duplications appear to be a common feature of the catabolic genes of TOL plasmids and possible reasons that they might confer a selective advantage on a plasmid have been discussed elsewhere (O'Donnell & Williams, 1991). Whether or not there is any selective advantage, the results in this paper illustrate that a consequence of the presence of such duplications is that they can serve as points at which recombination can occur to give rise to both new hybrid alleles of the genes and major rearrangements of DNA.

Some of this work was carried out with funding from the Science and Engineering Research Council. M. K. W. acknowledges receipt of a Sir William Roberts Scholarship from the University College of North Wales. C. L. P. was in receipt of sabbatical leave from the National University of Singapore.

References


