The giant linear plasmid pHG207 from *Rhodococcus* sp. encoding hydrogen autotrophy: characterization of the plasmid and its termini

**JUTTA KALKUS, CHRISTIANE DÖRRIE, DORIS FISCHER, MICHAEL REH and HANS GÜNTER SCHLEGL**

Institut für Mikrobiologie der Georg-August-Universität, Grisebachstraße 8, 37077 Göttingen, Germany

(Received 23 April 1993; accepted 25 May 1993)

As described previously, in *Rhodococcus* sp. (formerly *Nocardia opaca*) strains MR11 and MR22, the ability to grow as an aerobic hydrogen bacterium (the Aut character) is located on giant conjugative linear plasmids – on pHG201 (270 kb) in strain MR11 and on pHG205 (280 kb) in strain MR22. In an autotrophic transconjugant originating from MR22 a smaller plasmid, pHG207 (225 kb), was detected and shown to be a recombination product of the wild-type plasmids pHG204 and pHG205. A donor carrying pHG207 as the sole plasmid transferred the Aut marker at a 1000-fold frequency compared to the wild-type plasmid pHG205. Analysis of the plasmid ends revealed that plasmid pHG207 carries proteins at both ends; the proteins are linked to the 5' ends of the strands. The cloned end fragments of about 2 kb were sequenced and found to contain highly homologous sequences within the terminal 583 bp (left end part) and 560 bp (right end part). Several potential reading frames were detected, but database searching gave no indication about possible functions.

**Introduction**

*Rhodococcus* sp. strains MR11 and MR22 (formerly *Nocardia opaca*) are nutritionally versatile, strictly aerobic Gram-positive bacteria which are able to grow heterotrophically on various sugars, organic acids or hydrocarbons as well as lithoautotrophically on gaseous hydrogen and carbon dioxide. These bacteria contain a single hydrogenase, a cytoplasmic, heterotetrameric, NAD*-reducing enzyme, which aroused interest for investigation of intramolecular electron transport (Schneider et al., 1984; Zaborosch et al., 1989). The ability to grow autotrophically on hydrogen as sole energy source, denoted Aut', is transferable by conjugation to non-autotrophic mutants and various other heterotrophic species of *Rhodococcus* (Reh & Schlegel, 1975). The conjugative transfer and the instability of the Aut character in wild-type strains and transconjugants indicated the presence of an extrachromosomal Aut element (Reh & Schlegel, 1981). This assumption was confirmed by determining the specific activities of relevant enzymes in the donors and autotrophic transconjugants (Ecker et al., 1986). Three ccc DNA plasmids detected in *Rhodococcus* sp. MR11 and MR22 turned out not to be involved in lithoautotrophy (Reh, 1981). Even supposedly plasmid-free transconjugants functioned as efficient donors of the Aut character (Sensfuß et al., 1986).

The application of pulsed-field electrophoresis techniques combined with a gentle method of DNA preparation led to the detection of several giant linear plasmids both in wild-type strains and in various transconjugants (Kalkus et al., 1990). The wild-type strain MR11 contains the linear plasmids pHG201 (270 kb), pHG202 (400 kb) and pHG203 (420 kb), and strain MR22 contains pHG204 (180 kb), pHG205 (280 kb) and pHG206 (510 kb). In the autotrophic transconjugant MR2253 originating from MR22 another linear plasmid, pHG207 (225 kb), was found. The plasmids pHG201, pHG205 and pHG207 were shown to be carriers of the genes of cytoplasmic hydrogenase and ribulose-bisphosphate carboxylase (Kalkus et al., 1990).

To study the structure of the linear plasmids of *Rhodococcus* sp. MR11 and MR22 we undertook a physical characterization of pHG207, the smallest among the Aut elements. In this paper, we present restriction

---

**Abbreviations**: Aut, lithoautotrophic growth; ITR, inverted terminal repeats; PFE, pulsed-field electrophoresis.

The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession numbers L14442 and L14443.
maps of pHG207, pHG204 and pHG205, indicating that pHG207 is a product of recombination of pHG204 and pHG205. Furthermore, we demonstrated that the linear plasmid pHG207 has proteins at both ends which are probably covalently linked to the 5' termini of the plasmid DNA. Cloning and nucleotide sequence analysis of both ends revealed inverted terminal repeats.

Methods

Bacterial strains, phage and plasmids. Sources and references of strains of Rhodococcus used in this work are listed in Table 1. Escherichia coli* was purchased from Stratagene. For preparation of single-stranded DNA, the plasmid was digested with XhoI and XbaI (New England Biolabs, Beverly, MA, USA) and purified by phenol/chloroform extraction with a QIAQuick Gel Extraction Kit (Qiagen, Hilden, Germany). For selective conditions, media were supplemented with streptomycin (250 μg ml⁻¹) or thallium acetate (50 μg ml⁻¹) or X-Gal (40 μg ml⁻¹). Cloning and nucleotide sequence analysis of both ends revealed inverted terminal repeats.

Mating experiments. For construction of transconjugants a qualitative (agar mating) or a quantitative method (filter mating) were used. Matings occurred on hexadecane agar (Sensfuß et al., 1989) at 37 °C. If required, the medium was supplemented with ampicillin (100 μg ml⁻¹), IPTG (50 μg ml⁻¹) or X-Gal (40 μg ml⁻¹). Cloning of the terminal fragments of pHG207 in Bluescript SKM13 revealed the two hybrid plasmids pHG207-L (containing the left terminal 2.2 kb KpnI fragment from pHG207) and pHG207-R (containing the right terminal 2.3 kb Sall fragment from pHG207).

Restriction and ligation of DNA and isolation of DNA fragments. DNA, either dissolved or in agarose plugs, was digested with various restriction endonucleases as described by Sambrook et al. (1989). Restriction fragments were isolated from agarose gels either by using the Geneclean treatment kit (Biorad, Hercules, CA) or by electroelution of highly concentrated DNA plugs. After electroelution DNA was precipitated with ethanol. The redissolved DNA was dropped on a nitrocellulose filter (Millipore) floating on water and was dialysed against water (Marusyk & Sergeant, 1980). For preparation of non-proteolytically-treated pHG207 proteinase K treatment was omitted.

Electrophoresis. Conventional agarose gel electrophoresis was carried out using slab gels containing 1% (w/v) agarose in TAE (40 mM-Tris/acetate, 2 mM-EDTA; pH 7.9) at 4-6 V cm⁻¹. Pulsed-field electrophoresis (PFE) was performed using the Pulsaphor system (Pharmacia). For all separations 1% (w/v) agarose slab gels (15 × 15 × 0.4 cm) in 0.5 × TBE (Sambrook et al., 1989) were used. Buffer was kept at 11 °C, and gels were run at 6 V cm⁻¹. SDS-PFE was carried out by adding 0.2% SDS to buffer and agarose gel. As molecular mass standards HindIII- or PstI-cut λ DNA, a mixture of SacI- Apal- and KpnI-cut λ DNA or concatemers of λ DNA (Anand, 1986; Waterbury & Lane, 1987) were used.

For preparation of cells of E. coliXL1 blue and electrotransformation of freshly prepared or frozen cells was performed as described by Ausubel et al. (1987). For electroporation transformation the Gene-Pulser apparatus from Bio-Rad with 0.2 cm cuvettes was used.

Hybridization experiments. For fragmentation of large DNA molecules agarose gels were treated with HCl (Wahl et al., 1979). DNA was denatured by incubating the gel twice in 0.5 M-NaOH, 1.5 M-NaCl for 15 min. After neutralization (30 min in 0.5 M-Tris/HCl, 1.5 M-NaCl; pH 7.2) DNA was transferred with 20 × SSC (3 M-NaCl, 0.3 M-

### Table 1. Strains of Rhodococcus sp.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant phenotype</th>
<th>Plasmids</th>
<th>Parent strains</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhodococcus sp. (formerly Nocardia opaca)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR22</td>
<td>Aut+T'F*</td>
<td>pHG204, pHG205, pHG206</td>
<td>Wild-type</td>
<td>DSM 3346, Sensfuß et al. (1986)</td>
</tr>
<tr>
<td>MR226</td>
<td>Aut+T'P'Rif'</td>
<td>pHG33</td>
<td>MR22</td>
<td>Sensfuß et al. (1986)</td>
</tr>
<tr>
<td>MR2226</td>
<td>Aut+T'Str'</td>
<td>pHG33, pHG205</td>
<td>MR22</td>
<td>Sensfuß et al. (1986)</td>
</tr>
<tr>
<td>MR2223</td>
<td>Aut+P'T'</td>
<td>pHG33</td>
<td>MR22</td>
<td>This work</td>
</tr>
<tr>
<td>MR249</td>
<td>Aut+T'Str'</td>
<td>pHG33, pHG205</td>
<td>MR22 × MR2226</td>
<td>This work</td>
</tr>
<tr>
<td>MR253</td>
<td>Aut+T'Str'</td>
<td>pHG33, pHG205</td>
<td>MR22 × MR2226</td>
<td>Kalkus et al. (1990)</td>
</tr>
<tr>
<td>MR281</td>
<td>Aut+T'Str'</td>
<td>pHG33</td>
<td>MR22 × MR2226</td>
<td>This work</td>
</tr>
</tbody>
</table>

*AF, treated with acriflavine.
Exonuclease and mung bean nuclease digestion of pHG207. Isolated plasmid DNA (40 ng µl⁻¹) was embedded in an equal volume of 2% (w/v) low melting point agarose. After preparation of agarose plugs containing approximately 100 ng plasmid DNA, each test two plugs were equilibrated twice with exonuclease buffer (0.05 M-glycylglycine, 1 mM-MgCl₂, 1 mM-mercaptoethanol; pH 8.5). The buffer was replaced by 50 µl fresh buffer, and 30 U exonuclease III or 4 U A exonuclease were added. After 5-40 min at 37°C the reaction was terminated by adding 10 mM-EDTA and replacing the buffer with mung-bean nuclease buffer (10 mM-sodium acetate, 0.1 mM-zinc acetate, 50 mM-NaCl, 5%, w/v, glycerol, 1 mM-mercaptoethanol; pH 50). Agarose plugs were equilibrated as described above. Remaining single-stranded plasmid DNA was removed by adding 0.06 U mung-bean nuclease and incubation at 37°C for 30 min. For termination the buffer was exchanged with TE buffer (10 mM-Tris/HCl, 1 mM-EDTA; pH 8.0).

*DNA sequence analysis.* DNA sequencing was carried out by the dideoxy-chain-termination method (Sanger et al., 1977) with single-stranded or alkaline denatured double-stranded plasmid DNA. Single-stranded DNA from Bluescript derivatives was prepared according to Sambrook et al. (1989). For non-radioactive sequencing an Applied Biosystems model 370A apparatus in combination with a Sequenase version 2.0 sequencing kit (BRL) as recommended by the manufacturer. Hybridization and detection of biotinylated probes were carried out as described by Oelmüller et al. (1990). The buffer was replaced with 50 µl fresh buffer, and 30 U exonuclease III or 4 U A exonuclease were added. After 5-40 min at 37°C the reaction was terminated by adding 10 mM-EDTA and replacing the buffer with mung-bean nuclease buffer (10 mM-sodium acetate, 0.1 mM-zinc acetate, 50 mM-NaCl, 5%, w/v, glycerol, 1 mM-mercaptoethanol; pH 50). Agarose plugs were equilibrated as described above. Remaining single-stranded plasmid DNA was removed by adding 0.06 U mung-bean nuclease and incubation at 37°C for 30 min. For termination the buffer was exchanged with TE buffer (10 mM-Tris/HCl, 1 mM-EDTA; pH 8.0).

**Table 2. Characterization of thallium-resistant transconjugants and autotrophic transconjugants resulting from matings of the wild type strain MR22 with its plasmid-free derivative MR2226**

<table>
<thead>
<tr>
<th>Selective markers</th>
<th>Frequency of transconjugants per donor cell</th>
<th>Type of transconjugant</th>
<th>Relevant phenotype</th>
<th>Plasmids detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIPr</td>
<td>3 x 10⁻¹</td>
<td>1</td>
<td>Aut⁻ Tl⁻</td>
<td>pHG204</td>
</tr>
<tr>
<td>Aut⁻ Str⁻</td>
<td>7 x 10⁻³</td>
<td>2a</td>
<td>Aut⁻ Tl⁻</td>
<td>pHG205, pHG204, pHG33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2b</td>
<td>Aut⁻ Tl⁻</td>
<td>pHG205, pHG204</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2c</td>
<td>Aut⁻ Tl⁻</td>
<td>pHG205, pHG33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2d</td>
<td>Aut⁻ Tl⁻</td>
<td>Linear 225 kb plasmid</td>
</tr>
</tbody>
</table>

**Results**

*Correlation of the Aut⁻ and Tl⁻ phenotypes with the presence of linear plasmids.* When, in conjugative matings, *Rhodococcus* sp. wild-type strain MR22 harbouring the three linear plasmids pHG204, pHG205, pHG206 and the circular plasmid pHG33 was used as the donor and the plasmid-free derivative strain MR2226 served as recipient, transconjugants were obtained. When transconjugants for the Aut character and for thallium resistance were selected for, the transfer frequency was rather high in both cases (Table 2). Among the autotrophic transconjugants there were two groups: one which, in addition to the Aut character, had received the thallium resistance Tl⁻ marker, and one which had not. In contrast, all of the transconjugants selected for thallium resistance were heterotrophic, i.e. had not received the Aut character.

From each of the three phenotypically distinguishable groups – Aut⁻ Tl⁻, Aut⁻ Tl⁻ and Aut⁺ Tl⁻ – about a dozen transconjugants were analysed with respect to the presence of the linear plasmids and the circular plasmid. All strains selected for thallium resistance (Table 2, group 1) contained pHG204 as the sole plasmid. The detection of these strains, which did not harbour the circular plasmid pHG33, contradicts the assumption that this plasmid carries the Tl⁻ determinants (Sensfuß et al., 1986). So far no phenotypic marker can be assigned to pHG33.

All strains selected for the Aut character contained the linear plasmid pHG205 or a linear plasmid of a smaller size. The Aut⁺ Tl⁻ transconjugants contained two linear plasmids, pHG204 and pHG205; the majority in addition harboured pHG33 (group 2a). The linear plasmid pHG205 had previously (by DNA-DNA hybridization) been identified as the autotrophy-carrying plasmid of strain MR22 (Kalkus et al., 1990).
Among the Aut+ T1' transconjugants two subgroups could be differentiated. Group 2c contained, in addition to the linear plasmid pHG205, the circular plasmid pHG33. Cells of subgroup 2d contained, instead of plasmid pHG205, another linear plasmid intermediate in size between pHG204 and pHG205. As a representative of this type of transconjugant the strain MR2253 was selected, and its intermediate-size plasmid was designated pHG207. This strain MR2253 turned out to be a good donor for the Aut character and was chosen for further exploration of the origin of the pHG207.

For estimation of the transfer frequency of the linear plasmids pHG204, pHG205 and pHG207, donors containing only one of the three plasmids were constructed by conjugation. Only pHG205 was accompanied by the plasmid pHG33 as an additional extrachromosomal element. In a quantitative mating the Aut+ and T1' markers were transferred to the plasmid-free recipient MR2226 (Table 3). The results indicate that the markers were transferred by pHG204 and pHG205 at a 10^(-2)-fold frequency compared to the wild-type plasmid pHG205. If we assume that the transfer frequency is a plasmid-specific property, plasmid pHG207 represents a combination of the properties of plasmid pHG204 (high transfer frequency, T1') and pHG205 (low transfer frequency, Aut+). These results aroused the suspicion that the plasmid pHG207 is not a deletion product of pHG205, but is a recombinant plasmid.

pHG207 is a recombination product of wild-type plasmids pHG204 and pHG205

For exploration of the origin of the pHG207, samples of the isolated plasmid DNA of pHG207, pHG205 and pHG204 were digested using the restriction enzyme SnaBI, and the restriction fragments were separated by using PFE (Fig. 1). Only two of the eight visible bands of pHG207 were found to have partners in pHG205. Five of the eight bands of pHG207 agree in respect of their size to bands of pHG204. Corresponding results were obtained when the restriction enzymes AsnI, SpeI and XbaI were applied. The comparison indicates that plasmid pHG207 is composed of fragments detectable in pHG205 and pHG204. This presumption was confirmed by hybridization experiments using labelled pHG207 DNA as a probe. Results of the physical mapping with the enzymes SnaBI, AsnI, SpeI and XbaI, which cut DNA with high GC content rarely, are summarized in Fig. 2. The restriction maps show that plasmid pHG207 is a recombination product consisting of the left part of pHG204 and the right part of pHG205.

The assumption of a privileged recombination between plasmids pHG204 and pHG205 results from another observation. Of 36 heterotrophic mutants isolated after UV irradiation, 12 contained plasmids which can be assumed to be products of reciprocal recombination. These plasmids had a uniform size of about 230 kb, and consist of DNA sequences of pHG204 and pHG205 as hybridization analyses have shown.

Estimation of the copy number of pHG207

The copy number of pHG207 in strain MR2253 was estimated by comparison of differently treated samples of total DNA. For PFE total DNA of MR2253 was treated with SpeI and put into the agarose gel slot. A
Fig. 2. Macorstriction maps of pHG204, pHG207 and pHG205. cbbL, fragment contains the gene of the large subunit of ribulose-bisphosphate carboxylase; hox, fragments contain the hydrogenase structural genes; unknown order, order of these restriction fragments was not determined.

parallel sample of total DNA was left untreated before it was put into the slot. During electrophoresis intact chromosomal DNA does not migrate while plasmid DNA does. Then the fluorescence of the plasmid band and the restriction fragment bands was determined densitometrically. The fluorescence intensity of the plasmid band was compared to the intensity of bands of corresponding size originating from the digested total DNA. This kind of measurement indicated a copy number of 1 to 2 pHG207 molecules per chromosome.

Among the autotrophically growing strains of Rhodococcus sp. which have been examined so far, no variant has been encountered which does not contain linear extrachromosomal DNA. If the restriction pattern obtained after digestion of total DNA of strain MR2253 by SpeI was hybridized with labelled DNA probes of pHG207, only the known pHG207 fragments showed signals (plasmid pHG207 is cut only once by SpeI). This observation indicates that the DNA of plasmid pHG207 is present in strain MR2253 exclusively in the extrachromosomal and linear form. This conclusion has, however, to be limited because if a small minority of cells would contain plasmids in the integrated state, the signal would have remained hidden.

Structure of the ends of pHG207

So far neither the mechanism of replication nor the structure of the ends of the linear, Ant-character-harbouiring plasmids of Rhodococcus sp. had been studied. Therefore we started with a physical study of the ends of pHG207 to see whether they are linked with proteins, whether they consist of homologous DNA regions and whether any known open reading frames are detectable.

Exonuclease treatment. To explore the structural
Fig. 4. Change in electrophoretic mobility of restriction fragments of pHG207 after proteolytic treatment and Southern blot analysis with the cloned left end fragment as probe. (a) PFE separation of proteolytically or non-proteolytically treated digested pHG207. PFE separation programme: 30 min pulse time 0.1 s; 30 min pulse time 1 s; 11 h pulse time ramp 1–10 s. (b) Southern blot analysis with a biotin-labelled 1.9 kb EcoRI fragment from pHG207-L as probe. Lane 1, pHG207 digested with EcoRI and treated with proteinase K; 2, the same as 1, but without proteolytic treatment; 3, pHG207 digested with SalI and treated with proteinase K; 4, the same as 3, but without proteolytic treatment; 5, pHG207 digested with KpnI and treated with proteinase K; 6, the same as 5, but without proteolytic treatment. The arrows mark those fragments which showed a changed mobility. The arrowheads indicate the fragments that were found to be terminal fragments by the Southern blot analysis (b).

composition of the ends of pHG207 the degradability of the plasmid by a 3' and a 5' end-specific DNA exonuclease was examined. The isolated, proteinase-K-treated plasmid DNA was embedded in agarose gel and was incubated with either exonuclease III or λ exonuclease for 5–40 min at 37 °C. Subsequently, mungbean nuclease was added in order to augment the effect of the exonucleases. The result is shown in Fig. 3. Intact DNA of pHG207 was degraded only by the action of exonuclease III but not by the λ exonuclease (lanes 5–8 and 9–12), whereas fragmented DNA was degraded by both exonucleases (lanes 15, 16). The result shows that the 5' ends of the DNA were blocked whereas the 3' ends of pHG207 were freely accessible. The result suggests that the 5' ends of the native plasmid are covalently linked with a protein as described for Phi29 (Ito, 1978) and pSLA2 (Hirochika et al., 1984).

Evidence of terminal proteins (retardation experiment). If there are terminally bound proteins, the electrophoretic mobility of the end fragments of pHG207 in samples not treated with proteinase K should be lower than in proteolytically treated samples. DNA of pHG207 was prepared without a proteolytic treatment. After digestion of the DNA with KpnI, SalI or EcoRI, an aliquot of each sample was treated with 1% SDS and proteinase K (0.1 mg ml⁻¹). The digestions were analysed by PFE. As indicated in Fig. 4(a) a 2.2 kb KpnI, a 2.3 kb SalI and a 1.9 kb EcoRI band showed different mobilities. In the lanes containing samples that had been treated with proteinase K (lanes 2, 4 and 6), these bands appeared slightly below the position of the corresponding bands of untreated samples. The fluorescence intensity of the bands of the retarded fragments was significantly lower compared to the neighbouring bands as can be easily seen in the KpnI endonuclease lane (lane 6) indicating the unequal banding of the protein-carrying fragments. This difference in fluorescence intensity was not apparent when both KpnI lysates were electrophoretically separated in a 0.2% (w/v) SDS-containing gel, although the 2.2 kb KpnI fragment of samples that
had not been treated with proteinase K was retarded even under these conditions. Retardation of the protein-carrying fragments even in the presence of SDS indicates that the protein is covalently linked to the DNA.

In the experiment described, in each pattern only a single restriction fragment retarded by a covalently linked protein moiety became visible. All the following are possible explanations and had to be examined: (i) only one end of pHG207 is connected with a terminal protein; (ii) both ends of pHG207 are connected with protein, but (a) both end fragments are of equal size and are banding together, or (b) the second end fragment, due to its small size, disappeared during electrophoresis or remained hidden in the band pattern. It was improbable that both end fragments banded in the same position because the bands of all three lysates did not show 2-fold fluorescence intensity. To differentiate between the possibilities mentioned the 24 kb SpeI fragment, known from the restriction map to be the left side end fragment, was isolated, and samples were cut with endonucleases KpnI, EcoRI or SalI. In the resulting patterns of these lysates, the 2.2 kb KpnI and the 1.9 kb EcoRI fragment became visible. However, the 2.3 kb SalI fragment, which was characterized to be an end fragment by the retardation experiment, did not become visible. This experiment shows that among the fragments retarded by linked proteins the 2.2 kb KpnI and the 1.9 kb EcoRI fragment originated from the left side of the plasmid and the 2.3 kb SalI fragment originated from the right side of the plasmid. Thus, the experiment showed that pHG207 carries terminal proteins at both ends.

Cloning of the terminal fragments. For further analysis, especially for determining the DNA nucleotide sequence, of the left and right side ends of pHG207, the 2.2 kb KpnI and the 2.3 kb SalI fragment were isolated and were cloned although the 5' blockage still existed after proteinase K treatment. Ligation of the end fragments into the bluescript vector occurred via two unequal termini. The 2.2 kb KpnI fragment was ligated with its KpnI end to the KpnI position of the vector. The opposite end of the fragment was ligated to the blunt end of the vector produced by an EcoRV endonuclease cut. The 2.3 kb SalI fragment was cloned in an analogous manner. Both chimeric plasmids were introduced into E. coli XL1 blue by electrotransformation. Among the five and six, respectively, transformants isolated there was one clone, the plasmid of which contained either the 2.2 kb KpnI or the 2.3 kb SalI end fragment. The restriction maps of both plasmids, designated pHG207-L and pHG207-R, are shown in Fig. 5.

Hybridization of fragments with end fragment DNA probes. The cloned end fragments were used to identify the bands of the end fragments missing in the retardation experiment. As indicated in the restriction map of pHG207-R, the right side EcoRI end fragment was not recognized in the retardation experiment, probably due to its small size. But it should be possible to discover the right side KpnI and the left side SalI end fragment by hybridization. SalI- and KpnI-cut DNA, respectively, from the retardation experiment was probed with a biotinylated 1.9 kb EcoRI fragment from pHG207-L. In each lane two hybridization signals became visible indicating that the right and left side end fragments share homologous nucleotide sequences (Fig. 4b). These homologies were examined in detail by nucleotide sequence analysis. The hybridization experiment lead to the identification of a 40 kb SalI and a 70 kb KpnI fragment as further end fragments.

Nucleotide sequences of the terminal regions of pHG207. Sequence analysis of the cloned end fragments of pHG207 gave totals of 2235 bp for the 2.2 kb KpnI fragment and 2349 bp for the 2.3 kb SalI fragment (Fig. 6). Comparison of both sequences revealed a high

---

Fig. 5. Restriction maps of pHG207-L and pHG207-R. The restriction maps are based on nucleic acid sequence analysis data. The numbers below the restriction sites indicate the sequence position. The indicated restriction sites were confirmed by restriction analysis. The arrows represent the inverted terminal repeats (ITR).
Fig 6. Nucleotide sequences of pHG207-R and pHG207-L and alignment of the inverted terminal repeats (ITR). The arrows indicate short inverted and direct repeats.
homology of 91% between the 583 bp region of the left side end and 560 bp region of the right side end. The nucleotides in position 234–254 of pHG207-L are completely lacking in the sequence of pHG207-R. The homologous regions are sharply separated from adjoining sequences which do not show any similarities to each other. Thus, the terminal regions of plasmid pHG207 are incomplete inverted terminal repeats (ITR). The terminal hundred basepairs of both ends contain several palindromic sequences as indicated in Fig. 6.

The ITR sequences are characterized by astonishingly great variations in GC content (Fig. 7). Three successions of regions of high GC content alternate with regions of low GC content. In the sequences adjoining the ITR, the variations are significantly less pronounced, and the GC content alternates with a small amplitude around an average GC content of 62% in pHG207-L and 67% in pHG207-R.

The nucleotide sequences of both fragments were screened for potential open reading frames which would encode polypeptides with more than 100 amino acids. Four potential ORFs were found in the left side fragment and eight ORFs in the right side fragment; some of them are overlapping. Comparison of the derived polypeptide sequences with the nucleotide sequences of the September 1992 version of the EMBL and GenBank databases gave no indication about possible functions.

Discussion

The conjugative linear plasmid pHG207 from *Rhodococcus* sp. strain MR2253 turned out to be very interesting due to its special properties: (i) it carries the *Aut* character and thus permits transfer of the ability to grow as a hydrogenotroph to recipient cells; (ii) the transfer frequency is unusually high. The present study has shown that pHG207 is not a wild-type plasmid but originated from the wild-type plasmids pHG204 (characterized by thallium resistance and high transfer frequency) and pHG205 (*Aut* and low transfer frequency) by recombination. As plasmids of the size (about 225 kb) and with the properties of pHG207 originated in various independent experiments, the recombination of pHG204 and pHG205 is apparently reproducible and the result of a programmed gene arrangement. So far, we have not encountered *Aut* transconjugant cells or mutant cells which contain the *Aut* plasmid pHG205 as the sole extrachromosomal element. The conjugative transfer of pHG205 was always linked to the cotransfer of pHG33 or pHG204 (M. Reh, unpublished data). In contrast, pHG207 is mobilizable without any helper plasmid and is transferred at a very high frequency. In contrast to pHG205 the linear *Aut* plasmid pHG201 from the wild-type strain MR11 frequently occurs as a single plasmid in transconjugant cells and can be transferred singly (J. Kalkus, unpublished results). Whether the recombinant plasmid pHG207 is already present in the donor cells or is formed during conjugation is not known. An indication was provided by detecting faint bands of linear plasmids of the size of pHG207 in plasmid preparations from MR22 cells which were not taken from single colonies. Thus plasmid recombination may have already occurred in the donor cells (J. Kalkus, unpublished data).

The ends of plasmid pHG207 carry terminal proteins which are probably linked with the respective 5’ end of the DNA. These conclusions are based on the resistance of the plasmid DNA to *A* exonuclease and the retardation of the terminal fragments during gel electrophoresis. Furthermore, the nucleotide sequence of the cloned end fragments revealed inverted incomplete terminal repeats of 583 bp (ITR left side) and 560 bp (right side). These structures are similar to the well-studied linear genomes of adenoviruses and the linear plasmids pSLA2, SCP1 and pSC11 isolated from streptomyces. On the basis of its properties pHG207 can be assigned to the class of invertrons which comprise the linear plasmids, linear viral genomes and transposable elements as proposed by Sakaguchi (1990).

The length of the ITR sequences of pHG207 is similar to those of other linear plasmids such as pSLA2 (17 kb) from *Streptomyces rochei* (about 600 bp; Hirochika et
al., 1984), pMC3-2 (6 kb) from *Morchella conica* (about 700 bp; Rohr et al., 1991) and pSCL1 (11.7 kb) from *Streptomyces clavuligerus* (about 900 bp; Wu & Roy, 1993). These plasmids are significantly smaller than pHG207, however. The only large-size linear plasmid so far studied is SCP1 (350 kb) from *Streptomyces coelicolor*, which contains much larger ITR sequences of about 81 kb (Kinashi & Shimaji-Murayama, 1991; Kinashi et al., 1991). Thus, among the ITR sequences of linear replicons there is neither a uniform length nor a recognizable correlation between plasmid size and the length of the ITR sequence.

The mean GC content of the pHG207-L and pHG207-R end fragments amounts to 64.2 mol% G + C and confirms the molar GC content of *Rhodococcus* sp. MR1 as determined by thermal denaturation of the DNA (64.8 mol% G + C; Reh & Schlegel, 1975). Within the ITR sequences one region of about 150 bp is noticeable, because of its very low GC content. The termini of invertrons are a potential origin of replication, and such an AT rich region might facilitate strand separation at the initial stage of replication.

The structural similarities of ITR sequences of linear plasmids of different organisms or linear viral genomes are apparently restricted to frequent palindromic sequences at their immediate ends. They are thought to be potential binding sites for proteins required for replication. In the adenovirus genome, such a terminal palindromic sequence, which is essential for replication, in the terminal sequence of the linear plasmid SCP1. Further studies are required to elucidate the function of these structures and their role in replication.

The ITR sequences of both sides of pHG207 are not completely homologous to each other. In contrast, the ITR sequences of pSCL1 are identical over a wide range (689 bp), and small deviations occur only in an adjoining range (about 200 bp). Regrettably, the ITR sequences of two other similar plasmids have not been analysed completely. Of plasmid pSLA2, only parts of the sequence of the right side end were determined; as these parts were identical to the sequence of the left side end, Hirochika et al. (1984) assumed that both ends are identical. When the end fragments of SCP1 were cloned the two ends were not differentiated. However, because six clones were sequenced in parallel, presumably both ends were analysed. The only difference encountered concerned the number of terminal guanines (four to six bases). Thus the results lead to the conclusion that the terminal sequences of several hundred nucleotides of both end fragments are completely identical (Kinashi et al., 1991). We would emphasize again that pHG207 is a recombinant plasmid and the two ends originated from two different native plasmids. The non-identity of the ITR sequences of pHG207 can therefore be explained.

Central questions remained unanswered by this study. Each of the two wild-type strains, MR11 and MR22 of *Rhodococcus* sp., contains three linear plasmids, the end fragments of which lend themselves to settle the question as to the basic structure of neighbouring native plasmids. Furthermore, in both strains some plasmids carry selective markers such as Aut character, thallium resistance and transfer frequency. Another outstanding metabolic property is the capacity for fast growth on propane and other hydrocarbons of medium chain length. Thus, continuation of research on *Rhodococcus* sp. seems rewarding.

This investigation was supported by grants from the Deutsche Forschungsgemeinschaft (Sch 50/30-7, Re 842/1-1) and by the Max Buchner Forschungsstiftung. We are grateful to Dr M. Zimmer, Max Planck Institut für Experimentelle Medizin, Göttingen, for providing the opportunity to use the automated sequencer.

References


Sakaguchi, K. (1990). Invertrons, a class of structurally and functionally related genetic elements that include linear DNA plasmids, transposable elements, and genomes of adenovirus-type viruses. Microbiological Reviews 54, 66–74.


