JOURNAL OF BACTERIOLOGY, Nov. 1998, p. 5505–5514 Vol. 180, No. 21
0021-9193/98/$04.00+0

Int-B13, an Unusual Site-Specific Recombinase of the Bacteriophage P4 Integrase Family, Is Responsible for Chromosomal Insertion of the 105-Kilobase *clc* Element of *Pseudomonas* sp. Strain B13

ROALD RAVATN,† SONJA STUDER, ALEXANDER J. B. ZEHNDER, AND JAN ROELOF VAN DER MEER*

Swiss Federal Institute for Environmental Science and Technology (EAWAG) and Swiss Federal Institute for Technology (ETH), CH-8600 Dübendorf, Switzerland

Received 12 June 1998/Accepted 25 August 1998

***Pseudomonas* sp. strain B13 carries the *clcRABDE* genes encoding chlorocatechol-degradative enzymes on the self-transmissible 105-kb *clc* element. The element integrates site and orientation specifically into the chromosomes of various bacterial recipients, with a glycine tRNA structural gene (*glyV*) as the integration site. We report here the localization and nucleotide sequence of the integrase gene and the activity of the integrase gene product in mediating site-specific integration. The integrase gene (*int-B13*) was located near the right end of the *clc* element. It consisted of an open reading frame (ORF) of maximally 1,971 bp with a coding capacity for 657 amino acids (aa). The full-length protein (74 kDa) was observed upon overexpression and sodium dodecyl sulfate-polyacrylamide gel electrophoresis separation. The N-terminal 430 aa of the predicted Int-B13 protein had substantial similarity to integrases from bacteriophages of the P4 family, but Int-B13 was much larger than P4-type integrases. The C-terminal 220 aa of Int-B13 were homologous to an ORF flanking a gene cluster for naphthalene degradation in *Pseudomonas aeruginosa* PaK1. Similar to the bacteriophages ϕR73 and P4, the *clc* element integrates into the 3′ end of the target tRNA gene. This target site was characterized from four different recipient strains into which the *clc* element integrated, showing sequence specificity of the integration. In *Pseudomonas* sp. strain B13, a circular form of the *clc* element, which carries an 18-bp DNA sequence identical to the 3′-end portion of *glyV* as part of its attachment site (*attP*), could be detected. Upon chromosomal integration of the *clc* element into a bacterial attachment site (*attB*), a functional *glyV* was reconstructed at the right end of the element. The integration process could be demonstrated in RecA-deficient *Escherichia coli* with two recombinant plasmids, one carrying the *int-B13* gene and the *attP* site and the other carrying the *attB* site of *Pseudomonas putida* F1.**

Pseudomonas sp. strain B13 is a sewage isolate capable of utilizing 3-chlorobenzoate (3CBA) as its sole carbon and energy source (14). The degradation of 3CBA involves an initial oxidation to chlorocatechols, which are subsequently converted to 3-oxoadipate by the action of four enzymes of the modified *ortho* cleavage pathway, encoded by the *clcABDE* genes (15). The *clc* genes have been transferred from strain B13 to different *Pseudomonas* recipient bacteria, thereby enabling the recipients to degrade chlorocatechols as well (22, 26, 27, 40, 41). We have recently demonstrated that the *clc* genes are located on a 105-kb mobile element (named the *clc* element) which is the transfer determinant and is capable of integrating in the chromosome (25, 34). The original host *Pseudomonas* sp. strain B13 also carries two nonadjacent chromosomal copies of the *clc* element, although the isolation of small amounts of a 110-kb plasmid (pB13) carrying the *clc* genes in strain B13 has been reported elsewhere (10). The *Eco*RI restriction patterns of pB13 and the integrated *clc* element were basically identical, and the apparent 5-kb size difference was due only to inaccurate sizing of the largest *Eco*RI fragments (25). This suggested that pB13 and the integrating *clc* element exist in two different forms of the same entity, i.e., an integron and a free "plasmid."

The chromosomal location of the *clc* element was demonstrated by Southern hybridization on digested chromosomal DNAs separated by pulsed-field gel electrophoresis for transconjugants of *Pseudomonas putida* F1, *P. putida* BN10, *Burkholderia cepacia* WR401, *Alcaligenes eutrophus* CH34, and *Ralstonia* spp. (34). Some transconjugants carried only one chromosomal copy of the *clc* element, others carried two, and the F1 transconjugants carried up to eight copies (25, 34). Interestingly, chromosomal integrations in the F1 transconjugants occurred in two loci, with tandem amplification mainly in one locus. Integration of the *clc* element was shown to be RecA independent and site specific and should therefore have been mediated by functions encoded on the element itself (25). The integration sites in F1 were both identified as glycine tRNA structural genes, and the integrations appeared to take place at the 3′ end of the tRNA gene. A wide variety of genetic elements are known to integrate into the host chromosome by means of site-specific recombinases which use tRNA genes as their target sites. Such elements include the bacteriophages ϕR73 (17, 37), P4 and P22 (24), and T12 (21); insertional actinomycete plasmids (11); virulence determinants of *Dichelobacter nodosus* (11, 12) and of *Vibrio cholerae* (18); and the *Bacteroides* NBU1 element (33).

* Corresponding author. Mailing address: EAWAG, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland. Phone: (41) 1-823-5438. Fax: (41) 1-823-5547. E-mail: vdmeer@eawag.ch.

† Present address: Department of Molecular Genetics and Microbiology, UMDNJ-Robert Wood Johnson Medical School, Piscataway, NJ 08854.

TABLE 1. Plasmids used in this study

Plasmid	Source of insert DNA	Relevant characteristics
pRR104	4.2-kb *Eco*RI fragment of cosmid insert 3G3 cloned into pUC18Not (25)	Left junction (*attL1*) of the integrated *clc* element at *attB1-F1*
pRR108	4.1-kb *Nhe*I-*Eco*RI fragment of cosmid insert 2B1 cloned into pUC18Not (25)	*int-B13* plus right junction (*attR2-F1*)
pRR123	871-bp PCR product obtained from strain F1 DNA by using primers RR301 and RR303.	Integration site *attB1-F1* in strain F1 prior to integration
pRR146	476-bp PCR product obtained from strain B13 DNA by using primers RR316a and RR319.	*attP*; junction between left and right ends of the *clc* element
pRR165	3.5-kb *Nco*I-*Eco*RI fragment of pRR108 cloned into pUC28; *Sfi*I site in polylinker was removed.	*int-B13* plus right junction (*attR2-F1*)
pRR169	Integrase gene cloned into pET3c	*int-B13* overexpression
pRR169ΔNot	pRR169 with frameshift mutation in *Not*I site	*int-B13* (Δ*Not*I) overexpression
pRR171	3.5-kb *Eco*RI-*Nco*I fragment from pRR108 cloned into pACYC184	*int-B13* plus *attR2-F1*
pRR172	780-bp *Aat*II-*Eco*RI fragment in pRR171 exchanged for 200-bp *Aat*II-*Eco*RI fragment from pRR146	*int-B13* plus *attP*
pRR172ΔNot	pRR172 with frameshift mutation in *Not*I site	*int-B13* (Δ*Not*I) plus *attP*

In this paper, we present the characterization of a novel, unusually long recombinase gene (*int-B13*) of the phage P4 integrase family and demonstrate its function in site-specific integration of the *clc* element. To our knowledge, this is the first time that a bacteriophage-related integrase has been shown to be associated with horizontal transfer of genes involved in degradation of aromatic substances, further demonstrating the importance of this class of genetic elements in bacterial evolution.

MATERIALS AND METHODS

Bacterial strains, media, and growth conditions. *Escherichia coli* DH5α (31) was used for routine cloning experiments with plasmids. *E. coli* BL21(DE3) (pLysS) (35) was used as a host strain for inducible protein overexpression of pET3c-derived plasmids (29). *E. coli* cultures were grown at 37°C on Luria-Bertani medium (31), which was supplemented with the following antibiotics when appropriate: ampicillin, 100 μg/ml; chloramphenicol, 10 μg/ml; and tetracycline, 20 μg/ml. For induction of *int-B13* in *E. coli* BL21, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the medium at a 1 mM concentration.

Plasmids. The plasmid constructs used in this study are listed in Table 1. The plasmids pUC18Not (16), pUC28 (4), and pACYC184 (New England Biolabs, Beverly, Mass.) were used as general cloning vehicles. Plasmid pET3c (29) is an ATG vector derived from pBR322 which contains the ϕ_{10} promoter, ribosome binding site, and terminator optimized for T7-directed protein expression. The linearized vector pGEM-T Easy (Promega Corporation, Madison, Wis.), containing single 3′-thymidine overhangs, was used for cloning of DNAs amplified by PCR.

For overexpression of the integrase gene *int-B13* in *E. coli*, a translational fusion of the gene was constructed by using the ATG triplet in the *Nde*I site located downstream of the ϕ_{10} promoter and the ribosome binding site in pET3c as the start codon. The start of the *int-B13* gene was taken as position 262 (Fig. 1). First, a 379-bp PCR product was generated with primers RR330 and RR331 (Table 2) and subsequently digested with *Nde*I and *Sfi*I to create ligation termini. In a three-point ligation, the resulting DNA fragment (*Nde*I/*Sfi*I) and a fragment from pRR165 with the remainder of the *int-B13* open reading frame (ORF) (*Sfi*I/*Bam*HI) were cloned into pET3c (*Bam*HI and *Nde*I digested) to produce pRR169. The PCR-derived part of pRR169 was sequenced and confirmed to be identical to the original nucleotide sequence. Plasmid pRR169ΔNot is identical to pRR169 except for a frameshift mutation in the unique *Not*I site within the *int-B13* coding sequence. The frameshift mutation was introduced by digestion of pRR169 with *Not*I, filling in of the 3′ recessed ends, and religation. The presence of 4 additional nucleotides (nt) was confirmed by sequencing of this region of pRR169ΔNot.

Plasmid pRR171 contained the right end of the *clc* element with the *int-B13* gene plus the *attB1* site of *P. putida* F1 cloned in pACYC184. Plasmid pRR172 was constructed by replacing the 780-bp *Aat*I-*Eco*RI fragment in pRR171 with a 200-bp *Aat*I-*Eco*RI fragment from pRR146, thereby combining *int-B13* and *attP* as on the original circular form of the *clc* element. Plasmid pRR172ΔNot contained a frameshift mutation in *int-B13* as described above for pRR169ΔNot.

Expression of *int-B13* in *E. coli*. *E. coli* BL21(DE3)(pLysS) harboring plasmid pRR169 or pRR169ΔNot was grown in Luria-Bertani medium to an optical density at 540 nm of 0.45 to 0.55. Subsequently, cells were induced by the addition of 1 mM IPTG, and the culture was grown for another 90 min. Bacterial cells (1 ml) were harvested by centrifugation, resuspended in 50 μl of protein loading buffer (19), and boiled for 5 min. After 1 min of centrifugation (at 15,000 × *g*), samples of 5 to 10 μl were used directly for analysis by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), which was performed according to the method of Laemmli (19).

DNA isolations and manipulations. Plasmid DNA isolations (from *E. coli* DH5α), transformations, and other DNA manipulations were carried out according to established procedures (31). Total DNA was isolated with the Easy-DNA kit (Invitrogen, Carlsbad, Calif.) or by the method of Marmur (20). Restriction enzymes and other DNA-modifying enzymes were purchased from Amersham Life Science (Little Chalfont, Buckinghamshire, United Kingdom) and used according to the specifications of the manufacturer.

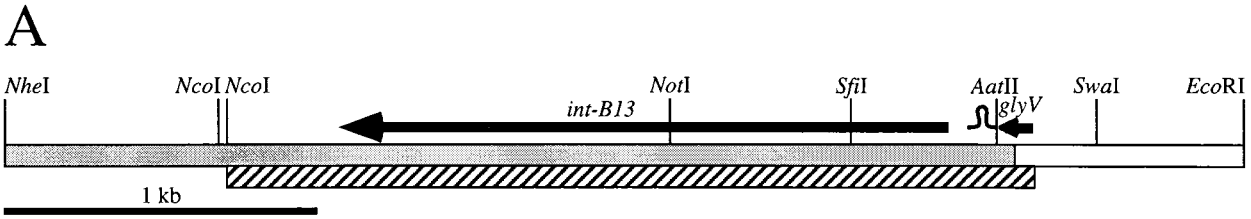
DNA amplification by PCR. PCRs were performed with *Taq* polymerase according to the descriptions of the supplier (Life Technologies, Basel, Switzerland). PCR primers used in this study (Table 2) were purchased from MWG Biotech (Ebersberg, Germany) or Microsynth (Balgach, Switzerland). The method referred to as colony PCR was performed as follows. One bacterial colony from an agar plate was transferred with a sterile toothpick into a 0.5-ml PCR tube containing 100 μl of distilled sterile water. The sample was heated to 98°C for 6 min in order to lyse the bacterial cells and release their DNA. For DNA amplification by PCR, 1 μl of this solution was added to a PCR mixture with a total volume of 50 μl.

DNA sequencing and sequence analysis. Double-stranded template sequencing was performed on plasmids with the Thermo Sequenase fluorescently labelled primer cycle kit with 7-deaza-dGTP (Amersham Life Science). Primers labelled with the fluorescent dye IRD-800 at the 5′ end were purchased from MWG Biotech. An automated DNA sequencer, model 4000L (LI-COR Inc., Lincoln, Nebr.), was used for sequencing. Computer analysis of the DNA and amino acid sequences was done with DNASTAR software (DNASTAR Inc., Madison, Wis.). Comparisons of our own sequence data with published sequences in GenBank were performed with the BLAST software via the Internet (http://www.ncbi.nlm.nih.gov/BLAST/ [2]).

Nucleotide sequence accession number. The nucleotide sequence presented in this article (Fig. 1) has been deposited in GenBank under accession no. AJ004950.

RESULTS

Identification of the *clc* element's putative integrase gene. Previously, we have reported chromosomal integration of a 105-kb genetic element (named the *clc* element) at two sites in *P. putida* F1 (25). The two integration sites in F1 were both identified as glycine tRNA structural genes, and each integration appeared to occur at the 3′ end of the target *gly*-tRNA. This observation suggested that a site-specific recombinase was responsible for the chromosomal integrations. Near the right end of the *clc* element (insert of plasmid pRR108 [Table 1]), an ORF was identified by sequencing and its predicted amino acid sequence was homologous to those of site-specific recombinases of the bacteriophage P4 integrase subfamily. The nucleotide sequence of this ORF (tentatively named *int-B13*) had a length of 1,971 bp, corresponding to a coding capacity of 657 amino acids (aa) and a molecular mass of 74 kDa. To confirm the presence and actual size of the *int-B13*-encoded polypeptide, the gene was cloned into pET3c and overexpressed in



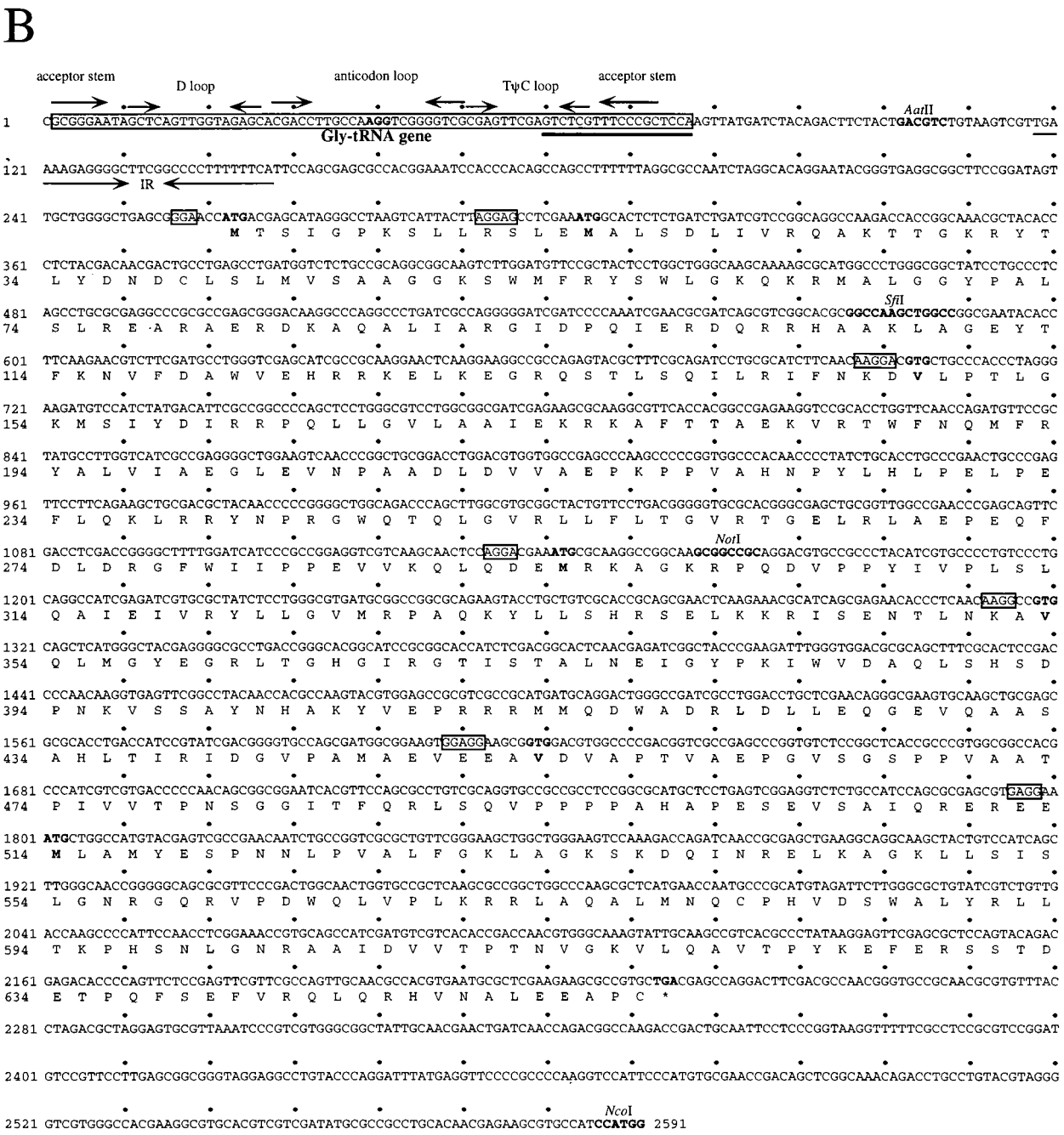


FIG. 1. Physical map of the right end of the integrated *clc* element in *P. putida* F1 harboring the integrase gene *int-B13*. (A) Restriction map of the insert DNA of pRR108 (Table 1), showing some of the important restriction sites. Grey shading indicates DNA which is part of the *clc* element. The hatched box below corresponds to the sequence depicted in panel B (opposite orientation). (B) Nucleotide sequence of the region containing *glyV* and *int-B13*. The sequence of *glyV* is boxed, and the identity segment of 18 bp is underlined. Arrows indicate inverted repeats (IR) within *glyV* and one between the tRNA gene and *int-B13*. The proposed amino acid sequence of Int-B13 and the ribosomal binding sites of putative ORFs are shown.

TABLE 2. PCR primers used in this study

Primer	Nucleotide sequence	Position (5′ end)
RR301	5′ GAG AAC GGA TTC AAC GCC ACC 3′	348 bp left of *attB1-F1*[a]
RR303	5′ ACT GCA GCA GAG CAC GCC GTT CG 3′	523 bp right of *attB1-F1*[a]
RR315	5′ TGC TCT CAG TTC CCG CAT CC 3′	*clc* element, 209 bp from left end, directed inwards[a]
RR316	5′ GAT GAC GTT GTG ACG ACT GC 3′	*clc* element, 178 bp from left end, directed outwards[a]
RR316a	5′ TGC TGG GTG TGG GTC TAT GGA TGC 3′	*clc* element, 147 bp from left end, directed outwards[a]
RR319	5′ TCA GGC TCA GGC AGT CGT TGT CG 3′	*clc* element, 329 bp from right end, directed outwards[a]
RR327	5′ CTC TCA GTT CCC GCA TCC GCT TCC 3′	*clc* element, 212 bp from left end, directed inwards[a]
RR330	5′ CTG AGC GGG ACA TAT GAC GAG CA 3′	*int-B13*, nt 249–271 (Fig. 1), introduction of *Nde*I site at ATG start codon
RR331	5′ GAC CCA GGC ATC GAA GAC GT 3′	*int-B13*, nt 608–627 (Fig. 1)
RR332	5′ CAT GAT TAC GAA TTC GCG AGC TCC 3′	pUC28 polylinker

[a] Relative to the positions of exact junctions between the integrated *clc* element and chromosomal DNA in strain F1.

E. coli BL21(DE3)(pLysS). The DNA sequence of the *int-B13* ORF predicted two possible translational starts close to another (Fig. 1, nt 262 and 304). Plasmid pRR169 contained the *int-B13* gene under the transcriptional control of the T7 promoter and carried a translational fusion from the ATG start codon at nt 262. When an *E. coli* BL21 culture with pRR169 was induced with IPTG and the crude extract was separated by SDS-PAGE, a protein band of the expected size (74 kDa) for Int-B13 was obtained (Fig. 2). However, at this resolution it was not possible to discern whether the ATG at nt 262 or the second ATG codon at nt 304 was actually used (or whether both were used). To confirm the origin of this protein band from the *int-B13* gene, a frameshift mutation was introduced in the unique *Not*I restriction site at nt 1158 in the sequence of *int-B13* (Fig. 1B). Induction of *E. coli* BL21 carrying this plasmid (pRR169ΔNot) produced a protein of 56 kDa on an SDS-PAGE gel (Fig. 2), which was the expected size for the truncated Int-B13. This confirmed the coding capacity of the *int-B13* ORF.

Int-B13 is a member of the integrase family of site-specific recombinases. A clustal alignment (MegAlign, DNASTAR software) of Int-B13's amino acid sequence with the seven most similar integrases is shown in Fig. 3A. The protein sequence with the highest similarity to Int-B13 is the integrase of retronphage ϕR73 (37) with 37% identity and 57% homology in an alignment covering 407 aa. The conserved residues His-396, Arg-399, and Tyr-433 (integrase family positions) found in all site-specific recombinases of the integrase (Int) family, with the exception of pSAM2 Int (3, 8), were also present in the Int-B13 sequence (His-365, Arg-368, and Tyr-401 [Fig. 3A]). The conserved Tyr-342 of λ Int (corresponding to Tyr-401 for Int-B13) has been shown to be the residue forming a covalent bond between the Int protein and the DNA at the attachment site(s) during recombination (23). In accordance with the findings of Abremski and Hoess (1), a second conserved arginine residue was also found in Int-B13, Arg-261 (Fig. 3A). The Int-B13 protein, however, was considerably longer than the phage P4-related integrases, which are all between 385 and 440 aa in length. The C-terminal 220 aa of Int-B13 had high similarity to only one other polypeptide in GenBank, encoded by an ORF present near the *pah* gene cluster for naphthalene degradation found in *Pseudomonas aeruginosa* PaK1 (Fig. 3B). No function has been assigned to the predicted protein from this ORF. Interestingly, the *P. aeruginosa* ORF could be part of a larger ORF, since it is located at the end of the DNA fragment sequenced. The translated sequence upstream of the ORF also matched the Int-B13 protein sequence very well (Fig. 3B).

Attachment sites of Int-B13. Site-specific recombinases like the integrases of bacteriophages P4 and ϕR73 mediate recombination between a phage attachment site, *attP*, and a bacterial chromosomal attachment site, *attB* (9). Likewise, the junction sequences between a prophage (or another type of integrated element) and chromosomal DNA are termed *attL* (left-end junction) and *attR* (right-end junction). During integration, the actual recombination event involving strand exchange occurs within short sequences identical to both *attP* and *attB*, the *att* core or identity segment. Since the putative chromosomal attachment sites *attB1-F1* of *P. putida* F1 (formerly INT1) and *attB2-F1* (INT2) had been identified previously (25), we now determined the *attP* site of the *clc* element itself. The precise sequence of the *attP* site would be evident only from a circular form with a direct junction between the left and right ends, which in the integrated form are 105 kb apart. This junction could be amplified from total DNA of *Pseudomonas* sp. strain B13 by using PCR with primers RR316a (left end) and RR319

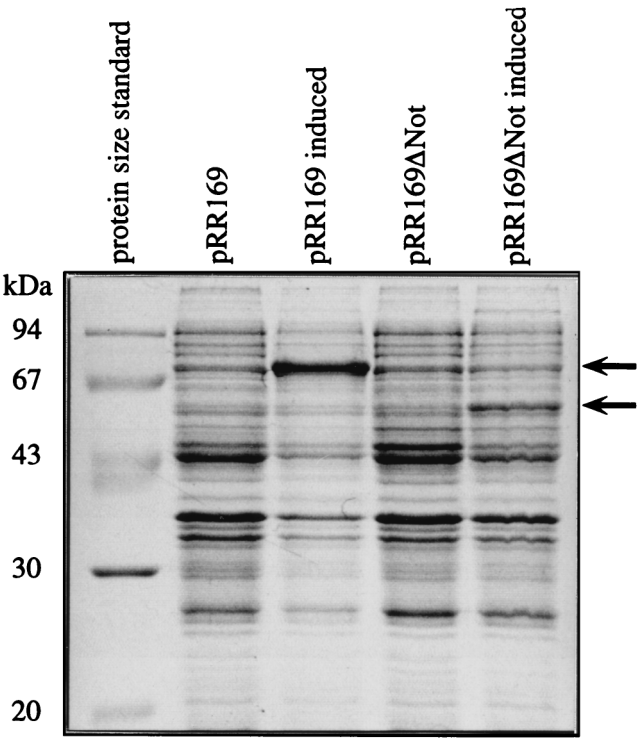


FIG. 2. Overexpression of the Int-B13 polypeptide in *E. coli*. Shown is an SDS-PAGE gel of crude extracts from *E. coli* BL21(DE3)(pLysS) cultures harboring either pRR169 (intact *int-B13*) or pRR169ΔNot (frameshift mutation in *int-B13*), with or without induction with IPTG. The sizes of protein standards are listed. Arrows indicate the positions of full-length Int-B13 (74 kDa) and truncated Int-B13 (56 kDa).

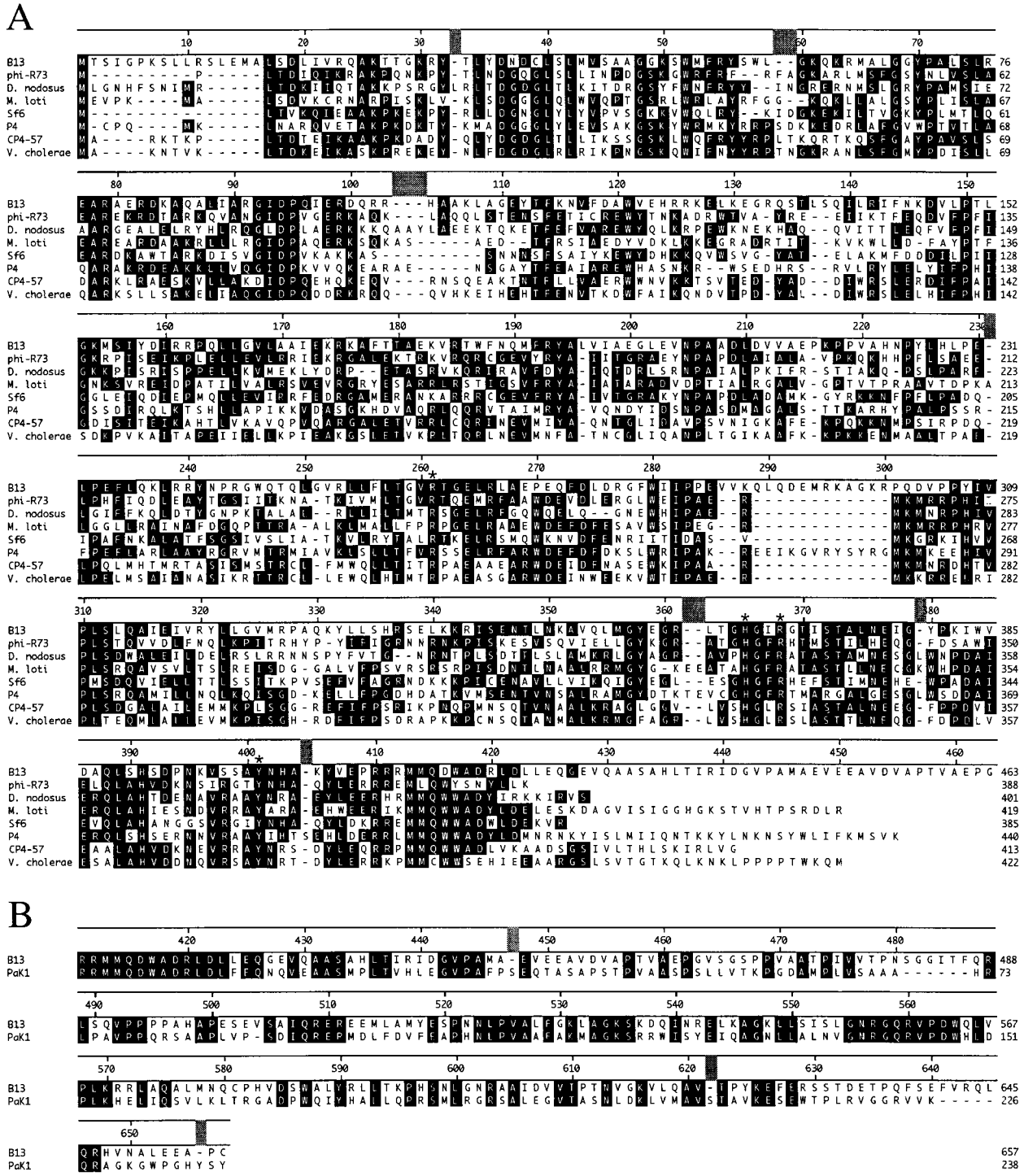
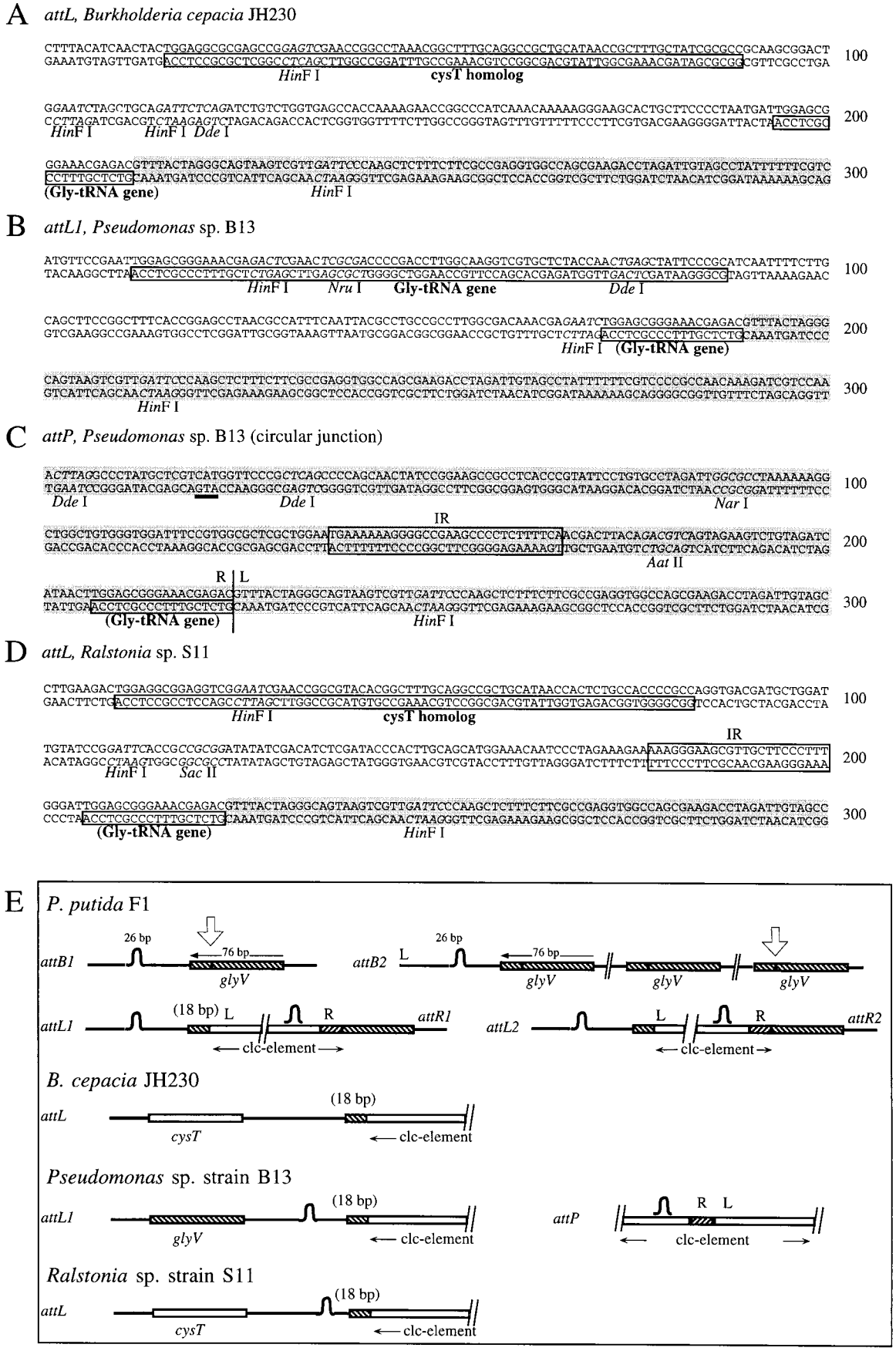


FIG. 3. Amino acid sequence alignment between the Int-B13 protein and homologous sequences identified in the databases. The ruler shown above the sequences is that for Int-B13. (A) Alignment of the N-terminal 460 aa of Int-B13 with the seven most similar integrases. The four conserved amino acid residues presumed to be responsible for catalytic function of integrases are indicated with asterisks. The GenBank accession numbers for the integrase sequences shown are as follows: *clc* element of *Pseudomonas* sp. strain B13, AJ004950; retronphage ϕR73, A42465; *D. nodosus vap* region, L31763; *M. loti* symbiosis island, AF049242; bacteriophage Sf6, P37317; satellite phage P4, RSBPP4; prophage CP4-57, P32053; and *V. cholerae* pathogenicity island, U02372. (B) Alignment between the C-terminal 247 aa of the Int-B13 protein and a translated nucleotide sequence from *P. aeruginosa* PaK1 (GenBank accession no. D84146). The predicted amino acid sequence from strain PaK1 is translated starting from nt 3 of the published sequence and is in frame with the proposed ORF1, which starts at Met-24.

(right end) (Table 2). As expected, the DNA sequence (insert of pRR146) of the amplified fragment contained the left- and right-end sequences and an 18-bp segment identical to the 3′ end of *glyV* (Fig. 4). The 18-bp segment seemed to be the core sequence of the *clc* element's *attP*, since only this segment was 100% identical to part of the chromosomal *attB* sites.

Sequence specificity of Int-B13-mediated integrations in different bacteria. The PCR was used to amplify the junctions



between chromosomal DNA and the left end of the integrated *clc* element (*attL*) in different recipients. The aim was to analyze whether *glyV* would be used as the target for Int-B13-mediated integration in hosts other than the previously analyzed *P. putida* F1 (25). The exact junction would be evident only from the nucleotide sequence at the left end, where the remainder of the *glyV* gene from *attB* is found. The bacteria analyzed for this purpose were *Pseudomonas* sp. strain B13 itself, *Ralstonia* sp. strain S11 (26), and *B. cepacia* JH230 (34). First, we determined by Southern hybridizations of *Nae*I- and *Sph*I-digested total DNAs what fragment sizes were to be expected from inverse PCR (iPCR) and what the copy number of the element in each strain was. From these hybridizations it was evident that the B13 genome contained two separate copies of the *clc* element, whereas *Ralstonia* sp. strain S11 and *B. cepacia* JH230 both contained only one integrated element (results not shown). The left junction of one of the integration sites of the *clc* element in strain B13 was amplified by iPCR. For this purpose, total DNA was digested with *Nae*I, religated, and subjected to PCR with the primers RR315 and RR316. Cloning and sequencing revealed the 3′ end of a glycine tRNA gene directly adjacent to the left end of the *clc* element, as in *P. putida* F1 (Fig. 4). A second, complete Gly-tRNA gene was detected 99 bp to the left of the integration site in strain B13.

The iPCR approach used to amplify *attL* from strain B13 did not yield any product for *Ralstonia* sp. strain S11 and *B. cepacia* JH230. Instead, for strain S11 we successfully used a linker-mediated PCR. Total DNA of this strain was digested with *Sph*I, ligated to a linker (i.e., *Sph*I-digested pUC28 DNA), and subjected to PCR with the primers RR316a and RR332. Sequencing of the cloned PCR product again revealed the 18-bp 3′ end of *glyV* adjacent to the left end of the *clc* element (Fig. 4). Downstream of the partial *glyV* sequence, another putative structural tRNA with 77% identity to *cysT* from *Streptomyces lividans* was found. For *B. cepacia* JH230, an inverse nested PCR was needed for amplification of the left junction. Then total DNA from strain JH230 was digested with *Nae*I, religated, and subjected to two rounds of PCR, first with primers RR315 and RR316 and then with primers RR316a and RR327. This procedure resulted in a PCR product whose junction sequence again contained the 3′ end of *glyV* (Fig. 4). Also here, a putative cysteine tRNA gene (86% similarity to *cysT* from *E. coli* K-12) was located downstream of the integration site.

Next, we examined if the *clc* element would always integrate in the same manner in one recipient strain. Six independently obtained transconjugants of *P. putida* F1 from matings with strain B13 (25) were analyzed for the left junction at the *attB1* site. Total DNA from the transconjugants was amplified in the PCR with primers RR301 and RR316. The cloned PCR products were sequenced and found to be identical (results not shown). In all transconjugants, the left end of the *clc* element had joined the target glycine tRNA gene exactly 18 bp from its 3′ end.

The characterized attachment sites were all identical with respect to the last 18 bp (3′ portion of *glyV*) adjacent to the exact left end of the *clc* element. This indicated that integrations occurred with high accuracy at this position on the 3′ end of the Gly-tRNA gene. In addition to the 18-bp identity segment found in all attachment sites, some other conserved sequences were also observed. The *attP* site of the *clc* element and *attB1* had a sequence identity of 83% in a 92-bp overlap including the 18-bp segment. The *clc* element's *attP* site and the *attB1* and *attB2* target sites from *P. putida* F1 had similar AT-rich inverted repeat sequences close to the target *glyV* gene (25), and the *attL* site of *Ralstonia* sp. strain S11 also contained an inverted repeat sequence (Fig. 4). Such structures are commonly found adjacent to (clusters of) tRNA structural genes, but their function or role in the attachment recognition is unknown.

Site-specific integration mediated by the *int-B13* gene product in *E. coli*. To demonstrate that the *int-B13* gene product was sufficient to promote site-specific recombination between *attP* and *attB*, we looked for integrase activity in *E. coli* DH5α, a strain which by itself is deficient in recombination. We cloned the *int-B13* gene plus the *attP* sequence on one plasmid (as on the circular form of the *clc* element) and looked at integration into a cloned *attB* site, present on a second plasmid. Surprisingly, the *attP* region could not be combined with the *int-B13* gene on a high-copy ColE1-based replicon, but only on the low-copy vector pACYC184 (i.e., pRR172 [Fig. 5]). Even then, *E. coli* DH5α cells harboring plasmid pRR172 had a reduced growth rate. These observations indicated that some kind of detrimental activity was associated with the construct. *E. coli* DH5α cotransformed with the plasmids pRR123 (containing *attB1*) and pRR172 (*attP* plus *int-B13*) showed the formation of a chimeric plasmid in some instances. Restriction enzyme analysis of this chimeric plasmid indicated an integrative recombination between the *attP* site of pRR172 and the *attB1* site of pRR123 (results not shown). DNA sequencing of the chimeric plasmid revealed that the sites of recombination were in fact identical to those previously characterized in the F1 transconjugant RR221 (25). The chimeric plasmid contained one *attR* site with a complete copy of *glyV* and one *attL* site with only the 18-bp 3′ end of *glyV* (Fig. 5).

Plasmid isolations from *E. coli* DH5α (pRR123 plus pRR172) cultivated in the presence of ampicillin and tetracycline usually showed the presence of the two original plasmids, whereas the chimeric plasmid was observed only occasionally. To determine the integration in a more sensitive and statistically reliable manner, we used PCR on individual transformants. Colonies of *E. coli* DH5α cotransformed with pRR123 plus pRR172 were subjected to colony PCR with the primers RR303 and RR319 (Table 2) (Fig. 5), resulting in an 852-bp PCR product specific for the chimeric plasmid (Fig. 5). From each of three independent cotransformations of pRR123 plus pRR172, 10 randomly chosen colonies were analyzed by PCR. This resulted in 66.7% of the colonies being positive for the chimeric plasmid form (Table 3). As controls, we used cotransformations of (i) a plasmid with a frameshift in *int-B13* but otherwise identical to pRR172 (pRR172ΔNot) and a plasmid with the *attB1* site (pRR123) and (ii) pRR123 and a plasmid containing an intact *int-B13* but without the *attP* site (pRR171). In these cotransformations, only 6.7% of the colonies were positive for the chimeric form (Table 3). The results indicated that activity

FIG. 4. Nucleotide sequences for Int-B13 attachment sites. The positions of relevant restriction sites, of large inverted repeats (IR), the Gly-tRNA gene or part of it, and putative other tRNA structural genes are shown. Grey shading indicates DNA that is part of the *clc* element. (A) *attL* from *B. cepacia* JH230. A sequence segment homologous to *cysT* from *E. coli* K-12 (GenBank accession no. X52789) is indicated. (B) *attL1* from *Pseudomonas* sp. strain B13. (C) *attP* from *Pseudomonas* sp. strain B13. The exact border between the *clc* element's left and right ends is indicated, and the proposed ATG start codon for *int-B13* is underlined. (D) *attL* from *Ralstonia* sp. strain S11. A sequence segment homologous to *cysT* from *S. lividans* (GenBank accession no. X52072) is indicated. (E) Summarized overview of the features of the Int-B13 attachment sites in the different hosts. For an explanation, see the text. Vertical arrows point to the insertion sites in *P. putida* F1. Loop structures (not necessarily the exact same sequence) are indicated. Diagram is not drawn to scale.

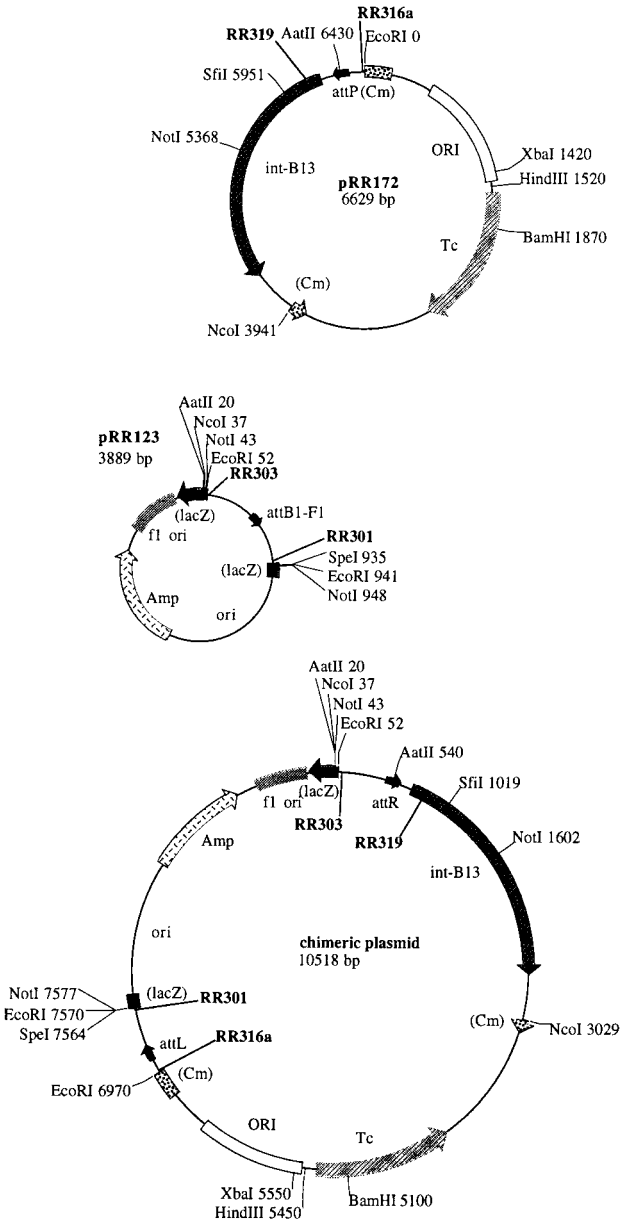


FIG. 5. Circular maps of the plasmids pRR123 and pRR172 (Table 1) and the chimeric plasmid resulting from Int-B13-mediated recombination between these two plasmids in *E. coli* DH5α. The relevant genetic markers, restriction sites, and PCR primer binding sites are shown. The arrows on the attachment sites (*attP*, *attB1-F1*, *attL*, and *attR*) indicate the direction of the glycine tRNA gene, or part of it. Inactivated genes are depicted within parentheses.

of an intact *int-B13* gene was needed for precise integration into *attB* and that *attP* (but not *attR*) was required for this integration to occur. The few positives obtained with pRR171 and pRR172ΔNot could have been PCR artifacts, resulting from hybridizations between DNA transcripts from pRR123 and those from the second plasmid. Since the regions *attP* and *attB1-F1* contain nearly identical sequence segments of 92 bp, transcripts terminating in either region could in the next cycle of annealing hybridize to the wrong template. We cannot exclude the possibility that some integrase activity resulted from pRR171 or pRR172ΔNot, but a chimeric plasmid could not be isolated when either of these plasmids was cotransformed with pRR123.

DISCUSSION

To the right end of the mobile *clc* element from *Pseudomonas* sp. strain B13, we localized functions involved in site-specific chromosomal integration. An ORF (*int-B13*) coding for an integrase of the bacteriophage P4 subfamily started approximately 200 bp from the junction between the element's right end and the chromosomal target, a glycine tRNA structural gene (Fig. 1). The sequence similarity of the 657-aa product of the *int-B13* gene to P4-related integrases and a demonstration of the gene's functionality gave evidence that Int-B13 was responsible for site-specific integrative recombination between the *clc* element's attachment site (*attP*) and chromosomal attachment sites (*attB* sites). Based on these results, we speculate that the *int-B13* gene is also responsible for site-specific chromosomal integration of the complete *clc* element.

The *clc* element's integrase showed significant amino acid sequence homology to integrases from bacteriophages like ϕR73, P4, and Sf6 (13, 17, 24, 37) (Fig. 3B). A high degree of amino acid sequence homology was also found between Int-B13 and the integrase IntS from the 500-kb symbiosis island of *Mesorhizobium loti* (36) and between Int-B13 and the integrase from the *vap* region of *D. nodosus* (11, 12). The majority of these P4-type integrases mediate site-specific integrative recombination involving tRNA structural genes. For instance, retronphage ϕR73 integrates into a *sel*-tRNA gene (37), satellite phage P4 integrates into a *leu*-tRNA (24), the symbiosis island from *M. loti* uses a *phe*-tRNA as a target site (36), and the *vap* region from *D. nodosus* seems to integrate into a *ser*-tRNA gene (11, 12). Similar to the observations for the *clc* element, these integrases mediate insertions into the 3′ ends of their target tRNA genes. Upon integration, the 3′ portion of the tRNA gene at *attB* is replaced by an identical segment carried on *attP*. For the *clc* element, this identity segment had a length of 18 bp. The exact reconstruction of the gene sequence of the target tRNA's 3′ end is an important feature in maintaining its essential function (9). Reiter et al. (28) pointed out that the identity segments of many elements inserting into tRNA genes extend from the anticodon loop through the 3′ end. However, the identity segments of the *att* sites from the *clc* element (18 bp) (Fig. 1B), phage P4 (20 bp), the *vap* region of *D. nodosus* (19 bp), and the *M. loti* symbiosis island (17 bp) are shorter, extending from the TψC loop through the 3′ end. Regions of dyad symmetry characteristic of tRNA genes are supposed to serve as integrase binding sites (28). However, this can be true only for *attB* or *attL* DNA containing the complete tRNA gene and not for the corresponding *attP* site which contains only the 3′ portion of the sequence.

The complete functional Int-B13 protein had a considerably higher molecular mass than other known P4-type integrases. Even so, several smaller ORFs were found in frame with the

TABLE 3. Colony PCR detection of a chimeric plasmid after cotransformations into *E. coli* DH5α

Plasmids	*n*[a]	No. of positives[b]	% Positives[c]
pRR123 plus pRR171	30	2	6.7 ± 5.8
pRR123 plus pRR172	30	20	66.7 ± 15.3
pRR123 plus pRR172ΔNot	30	2	6.7 ± 11.5

[a] Total number of tested colonies (from three independent cotransformations).
[b] Colony PCR yielded a product of 852 bp with the primer RR303 plus RR319, an annealing temperature of 58°C, an extension time of 50 s, and 35 cycles of amplification.
[c] Mean ± standard deviation.

largest coding region (Fig. 1B). Although not investigated, the translational start of Int-B13 may be ATG at nt 304 rather than that at nt 262, due to a better ribosome binding site. Other downstream translational starts may result in truncated forms of Int-B13 lacking part of the N-terminal domain. For the conjugative transposon Tn*916*, truncated integrase proteins are thought to be involved in regulating recombinational activity (32) by interacting with the full-length integrase protein or the attachment sites. Since Int-B13 is so much longer than other P4-related integrases and the C-terminal region is not homologous to the site-specific recombinases, the ORF starting at Val-453 could have a different role. Perhaps this part codes for the excionase function, which is typically clustered together with an integrase. The excionase stimulates excisive recombination and is usually a small protein of 60 to 120 aa (5, 30). The C-terminal region of Int-B13 did not show homology with known excionases, but this type of protein normally has very little homology (7, 30). The only sequence with significant homology to the C-terminal domain of Int-B13 was an ORF originating in *P. aeruginosa* PaK1 (38). The ORF is located upstream of an NAH7-like *pah* gene cluster, putatively encoding naphthalene degradation. Interestingly, a translation of the (published) nucleotide sequence upstream of the ORF also revealed the RRMMQDWADRLDL residue motif, which forms the last conserved region in the C termini of the P4-related integrases (Fig. 3). Therefore, we suspect that the proposed ORF represents the C-terminal domain of a larger ORF, similar to Int-B13. No function has yet been assigned to the ORF flanking the *pah* cluster, but the entire gene cluster is thought to be part of a mobile element (39).

The previously isolated and characterized plasmid pB13 carrying the *clc* genes in *Pseudomonas* sp. strain B13 (10) seems to be identical to our integrated *clc* element (25). The fact that other research groups have been unable to isolate plasmid DNA from strain B13 (22, 41) was probably due to the *clc* element's integration into the chromosome. Our observations at the moment indicate that the extrachromosomal circular form of the element is abundant only in strain B13 in stationary phase during growth on 3CBA (results not shown). Interestingly, in most other transconjugants analyzed so far, none or very little of the circular form can be detected by PCR amplification (results not shown). How excision and transfer of the *clc* element are regulated will therefore have to be studied more extensively.

It is becoming clear that a new form of mobile genetic elements exists, which we propose to call "gene islands" (after the use of the terms pathogenicity island and symbiosis island). Such elements harbor integrases related to those of bacteriophages and may have been evolutionarily derived from those. Examples include insertional plasmids from actinomycetes (5–8), some of the pathogenicity islands from gram-negative pathogens (11, 12, 18), and a recently discovered 500-kb transferable region (symbiosis island) from *M. loti* (36). This symbiosis island seems to carry all the genetic information required for nodule formation, symbiotic nitrogen fixation, and synthesis of three vitamins. For the first time, our work demonstrated that a bacteriophage-related integrase was associated with horizontal transfer of genes coding for degradation of xenobiotics, and the *clc* element could therefore be considered a "degradation island." Apparently, bacteriophage-related integrases using tRNA structural genes as (chromosomal) insertion sites are involved in horizontal transfer of very diverse genetic determinants, not only those of bacterial virulence (11). This class of integrating elements may have been underestimated and could have greater evolutionary importance than previously thought.

REFERENCES

1. **Abremski, K. E., and R. H. Hoess.** 1992. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. Protein Eng. **5:**87–91.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Argos, P., A. Landy, K. Abremski, J. B. Egan, E. Haggard-Ljungquist, R. H. Hoess, M. L. Kahn, B. Kalionis, S. V. Narayana, L. S. Pierson, N. Sternberg, and J. M. Leong.** 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. **5:**433–440.
4. **Benes, V., Z. Hostomsky, L. Arnold, and V. Paces.** 1993. M13 and pUC vectors with new unique restriction sites for cloning. Gene **130:**151–152.
5. **Boccard, F., T. Smokvina, J. L. Pernodet, A. Friedmann, and M. Guerineau.** 1989. The integrated conjugative plasmid pSAM2 of *Streptomyces ambofaciens* is related to temperate bacteriophages. EMBO J. **8:**973–980.
6. **Brasch, M. A., and S. N. Cohen.** 1993. Excisive recombination of the SLP1 element in *Streptomyces lividans* is mediated by Int and enhanced by Xis. J. Bacteriol. **175:**3075–3082.
7. **Brasch, M. A., G. S. Pettis, S. C. Lee, and S. N. Cohen.** 1993. Localization and nucleotide sequences of genes mediating site-specific recombination of the SLP1 element in *Streptomyces lividans*. J. Bacteriol. **175:**3067–3074.
8. **Brown, D. P., K. B. Idler, and L. Katz.** 1990. Characterization of the genetic elements required for site-specific integration of plasmid pSE211 in *Saccharopolyspora erythraea*. J. Bacteriol. **172:**1877–1888.
9. **Campbell, A. M.** 1992. Chromosomal insertion sites for phages and plasmids. J. Bacteriol. **174:**7495–7499.
10. **Chatterjee, D. K., and A. M. Chakrabarty.** 1983. Genetic homology between independently isolated chlorobenzoate-degradative plasmids. J. Bacteriol. **153:**532–534.
11. **Cheetham, B. F., and M. E. Katz.** 1995. A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. Mol. Microbiol. **18:**201–208.
12. **Cheetham, B. F., D. B. Tattersall, G. A. Bloomfield, J. I. Rood, and M. E. Katz.** 1995. Identification of a gene encoding a bacteriophage-related integrase in a *vap* region of the *Dichelobacter nodosus* genome. Gene **162:**53–58.
13. **Clark, C. A., J. Beltrame, and P. A. Manning.** 1991. The *oac* gene encoding a lipopolysaccharide O-antigen acetylase maps adjacent to the integrase-encoding gene on the genome of *Shigella flexneri* bacteriophage Sf6. Gene **107:**43–52.
14. **Dorn, E., M. Hellwig, W. Reineke, and H.-J. Knackmuss.** 1974. Isolation and characterization of a 3-chlorobenzoate degrading pseudomonad. Arch. Microbiol. **99:**61–70.
15. **Frantz, B., and A. M. Chakrabarty.** 1987. Organization and nucleotide sequence determination of a gene cluster involved in 3-chlorocatechol degradation. Proc. Natl. Acad. Sci. USA **84:**4460–4464.
16. **Herrero, M., V. de Lorenzo, and K. N. Timmis.** 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. J. Bacteriol. **172:**6557–6567.
17. **Inouye, S., M. G. Sunshine, E. W. Six, and M. Inouye.** 1991. Retronphage ϕR73: an *E. coli* phage that contains a retroelement and integrates into a tRNA gene. Science **252:**969–971.
18. **Kovach, M. E., M. D. Shaffer, and K. M. Peterson.** 1996. A putative integrase gene defines the distal end of a large cluster of ToxR-regulated colonization genes in *Vibrio cholerae*. Microbiology **142:**2165–2174.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
20. **Marmur, J.** 1961. A procedure for the isolation of deoxyribo-nucleic acid from micro-organisms. J. Mol. Biol. **3:**208–218.
21. **McShan, W. M., Y. F. Tang, and J. J. Ferretti.** 1997. Bacteriophage T12 of *Streptococcus pyogenes* integrates into the gene encoding a serine tRNA. Mol. Microbiol. **23:**719–728.
22. **Oltmanns, R. H., H. G. Rast, and W. Reineke.** 1988. Degradation of 1,4-dichlorobenzene by enriched and constructed bacteria. Appl. Microbiol. Biotechnol. **28:**609–616.
23. **Pargellis, C. A., S. E. Nunes-Düby, L. M. de Vargas, and A. Landy.** 1988. Suicide recombination substrates yield covalent lambda integrase-DNA complexes and lead to identification of the active site tyrosine. J. Biol. Chem. **263:**7678–7685.
24. **Pierson, L. S., and M. L. Kahn.** 1987. Integration of satellite bacteriophage P4 in *Escherichia coli*. DNA sequences of the phage and host regions involved in site-specific recombination. J. Mol. Biol. **196:**487–496.
25. **Ravatn, R., S. Studer, D. Springael, A. J. B. Zehnder, and J. R. van der Meer.** 1998. Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. strain B13. J. Bacteriol. **180:**4360–4369.
26. **Ravatn, R., A. J. B. Zehnder, and J. R. van der Meer.** 1998. Low-frequency horizontal transfer of an element containing the chlorocatechol degradation genes from *Pseudomonas* sp. strain B13 to *Pseudomonas putida* F1 and to

indigenous bacteria in laboratory-scale activated-sludge microcosms. Appl. Environ. Microbiol. **64:**2126–2132.
27. **Reineke, W., S. W. Wessels, M. A. Rubio, J. Latorre, U. Schwien, E. Schmidt, M. Schlömann, and H. J. Knackmuss.** 1982. Degradation of monochlorinated aromatics following transfer of genes encoding chlorocatechol catabolism. FEMS Microbiol. Lett. **14:**291–294.
28. **Reiter, W. D., P. Palm, and S. Yeats.** 1989. Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res. **17:**1907–1914.
29. **Rosenberg, A. H., B. N. Lade, D. S. Chui, S. W. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56:**125–135.
30. **Salyers, A. A., N. B. Shoemaker, A. M. Stevens, and L. Y. Li.** 1995. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol. Rev. **59:**579–590.
31. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Scott, J. R., and G. G. Churchward.** 1995. Conjugative transposition. Annu. Rev. Microbiol. **49:**367–397.
33. **Shoemaker, N. B., G. R. Wang, and A. A. Salyers.** 1996. The *Bacteroides* mobilizable insertion element, NBU1, integrates into the 3′ end of a Leu-tRNA gene and has an integrase that is a member of the lambda integrase family. J. Bacteriol. **178:**3594–3600.
34. **Springael, D., and R. Ravatn.** Unpublished results.
35. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
36. **Sullivan, J. T., and C. W. Ronson.** 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc. Natl. Acad. Sci. USA **95:**5145–5149.
37. **Sun, J., M. Inouye, and S. Inouye.** 1991. Association of a retroelement with a P4-like cryptic prophage (retronphage ϕR73) integrated into the selenocystyl tRNA gene of *Escherichia coli*. J. Bacteriol. **173:**4171–4181.
38. **Takizawa, N.** 1996. Direct submission to the databases (GenBank accession no. D84146).
39. **Takizawa, N.** Personal communication.
40. **Thiem, S. M., M. L. Krumme, R. L. Smith, and J. M. Tiedje.** 1994. Use of molecular techniques to evaluate the survival of a microorganism injected into an aquifer. Appl. Environ. Microbiol. **60:**1059–1067.
41. **Zhou, J. Z., and J. M. Tiedje.** 1995. Gene transfer from a bacterium injected into an aquifer to an indigenous bacterium. Mol. Ecol. **4:**613–618.