The site-specific recombinase encoded by pinD in Shigella dysenteriae is due to the presence of a defective Mu prophage

Akira Tominaga

Tel: +81 086 251 7864. Fax: +81 086 254 4648. e-mail: atomi@cc.okayama-u.ac.jp

INTRODUCTION

Inversional switching systems in prokaryotes, which control alternative expression of two sets of genes, are made up of an invertible DNA segment, its flanking two crossover sites (inv) and a site-specific recombinase gene. Four systems have previously been characterized in detail: the Salmonella typhimurium H segment-hin gene (H-hin) (Silverman et al., 1979; Szekely & Simon, 1981; Zieg & Simon, 1980), phage Mu G-gin (Koch et al., 1987), phage P1 C-cin (Hiestand-Nauer & Iida, 1983; Iida, 1984; Iida et al., 1982) and Escherichia coli e14 P-pin (Plasterk et al., 1983; Plasterk & van de Putte, 1985). In Salm. typhimurium, H inversion concerns flagellar phase variation, and in both phases inversion of the segments mediates reciprocal expression of two sets of tail fibre genes. Phenotypic change caused by P inversion is unknown. In contrast to the difference in biological function, these recombinases are functionally interchangeable with one another (Kamp & Kahmann, 1981; Kutsukake & Iino, 1980; Momota & Enomoto, 1986; van de Putte et al., 1984).

Previously, I reported two recombinase genes from strains of the genus Shigella (Tominaga et al., 1991). These, pinB and pinD, are derived from Shigella boydii, and Shigella dysenteriae, respectively. Afterwards, the strain of S. boydii used (Tominaga et al., 1991) was proved to be Shigella sonnei. Although phenotypic changes in S. sonnei and S. dysenteriae are unknown, the difference of inversional activity between pinB and pinD for known invertible segments suggests that pinB and pinD are different. The invertible B segment and pinB gene analysed in nucleotide level are highly homologous to phage Mu G-gin system (Tominaga et al., 1991). On the other hand, it is unclear whether pinD is different from known recombinase genes, since the switching system of S. dysenteriae has not been identified.

In this report, to elucidate the differences in function and structure between pinD and known recombinase genes, the nucleotide sequences of the pinD gene and its flanking regions from S. dysenteriae were determined. The result indicates that the genetic organization of

Abbreviation: IRR, inverted right repeat.

The DDBJ accession number for the sequence reported in this paper is D63765.
pinD flanking regions is identical to that of the G segment and β region of Mu. Moreover, Southern analysis and functional analysis indicate that the pinD system is the remnant of a Mu-like prophage.

METHODS

Bacterial strains, phage and plasmids. S. dysenteriae Sh16 (Scott, 1968) was used as DNA donor for Southern hybridization and cloning the pinD gene (Tominaga et al., 1991). E. coli K-12 strains used were as follows. EJ2282 [ΔhisC purB51 ΔpinE recA56 hisDR] (Tominaga et al., 1991) and MV1184 (Vieira & Messing, 1987), which was found to be defective in pinE function, were used as the hosts for recombinant plasmids. MucTs (Tominaga et al., 1986) was used as probe for hybridization.

Plasmid pTY501 carrying a 6.75 kb SalI fragment containing the H-fljB region fixed at the ‘off’ state was used for the detection of H inversion as described below. Plasmids containing the inverted segment fixed in the (-) orientation were constructed as follows. A 3.1 kb HindIII fragment containing the P segment of e14 was subcloned from the plasmid pHA201 (Tominaga et al., 1991) into the SalI site of pUC118 (Vieira & Messing, 1987), yielding pPSM62. To construct plasmid pCSM103, a 6.6 kb BamHI–MluI fragment containing the C segment of phage P1 was subcloned from the plasmid pAT101 (Tominaga et al., 1986) into pUC118. To construct plasmid pHSM64, a 3.3 kb PvuII fragment from the pTY109 containing the H segment (Okazaki et al., 1993) was subcloned into the SalI site of pUC118. A 4.2 kb BamHI–SalI fragment from pTSB918 containing the B segment (Tominaga et al., 1991) was subcloned into the BamHI and SalI sites of pUC118, to give pBSM73. A 6.6 kb PstI–AvaiI fragment containing the G segment of phage Mu (Koch et al., 1987) was cloned into the PstI and AvaiI sites of pUC119 (Vieira & Messing, 1987), to give pGSM97. Plasmid pTSD23 containing pinD is a pACYC177 (Chang & Cohen, 1978) -derivative that carries a 1.6 kb PvuII fragment from a deletion plasmid for expression. pTSD29 is a pUC119-derivative that carries a 1.3 kb PstI fragment containing pinD and a part of P segment fixed in the (-) orientation (probe PINE) was subcloned from pHA201 into the PstI site of pUC118. A 1.0 kb (Fig. la, position 490–1485) fragment containing pinD was used as probe PIND. Southern analysis was performed by blotting agarose gels onto nylon membrane filters (Amersham) using a vacuum blotter. The restriction fragments used for probes were prepared as described before (Tominaga et al., 1991) with digoxigenin-labelled dUTP with a non-radioactive DNA-labelling kit (Boehringer). Hybridization was done by standard methods (Maniatis et al., 1982), and hybrid bands were detected by the digoxigenin-ELISA method using the Boehringer detection kit.

RESULTS AND DISCUSSION

Cloning and expression of the S. dysenteriae pinD gene in E. coli

A site-specific recombinase gene pinD in S. dysenteriae strain Sh16 was found previously through its ability to mediate H segment inversion (Tominaga et al., 1991). Digests with BamHI, HindIII or SalI of total Sh16 DNA were hybridized with the probe containing the B segment and pinB gene (B-pinB) from S. sonnei. BamHI or SalI digest showed one hybrid band of approximately 21 kb and HindIII digest showed two bands of 2.1 and 12.5 kb (data not shown). The size of hybrid bands with BamHI or SalI digest suggests that a region containing the pinD gene is different in restriction sites from the corresponding region of S. sonnei since the 5.9 kb BamHI fragment (B-pinB) contains no HindIII sites and a unique SalI site (Tominaga et al., 1991).

For cloning the pinD gene, SalI fragments were inserted into pBR322 and introduced into the non-motile strain EJ2286 [ΔhisC/pTY501(H–) fljB+]. (Tominaga et al., 1991). One of the recombinant plasmids, pTSD608,
which conferred motility on the host by causing inversion of the H segment, was isolated and contained a 10.1 kb insert (Fig. 1b). To localize the pinD gene, this 10.1 kb SalI fragment, whose size was different from that expected from Southern hybridization, was subdivided into several smaller plasmids. Using the inversion activity for the H segment as an assay, a 3.0 kb HindIII-PvuII fragment (pTSD611) was found to be the smallest fragment which carried the pinD gene (Fig. 1b). An invertible segment was expected to lie in either side of pinD by analogy with other switching systems. However, no inversion region was detected in this fragment when pTSD608 prepared from successive cultures was digested with restriction enzymes (Fig. 1b).

In a previous report (Tominaga et al., 1991), pinD mediated inversion of C, H and P segments with different frequencies to those for genes intrinsic to each segment. To compare the inversional activity of pinD for different segments, pTSD23 containing pinD was introduced into MV1184 along with compatible plasmids carrying one of five segments fixed in the (−) orientation: B, C, G, H or P. After approximately 30 generations, all plasmids digested with appropriate restriction enzymes produced two sets of fragments specific to both orientations resulting from the inversion of segments (data not shown). In all cases, the inversion was equilibrated in both orientations, but no inversion was detected even after longer incubation when pTSD23 was absent. This result indicates that PinD has qualitatively the same activity for the recombination of all invertible segments examined in E. coli K-12. The quantitative differences in activities of recombinases observed previously (Tominaga et al., 1991) may have been due to different levels of expression of pinD or availability of other accessory factors (Glasgow et al., 1989; Johnson et al., 1987; Kahmann et al., 1985).

Nucleotide sequences of the pinD gene and its flanking regions

The nucleotide sequence of the 3.0 kb HindIII-PvuII fragment bearing the pinD gene was determined (Fig. 1a). Three complete and two incomplete ORFs were inferred (Fig. 1b). A complete ORF, starting with GTG instead of ATG as in gin and pinB (Plasterk et al., 1983; Tominaga et al., 1991) and preceded by the Shine-Dalgarno (SD) sequence and Prihnow box, was identified as the pinD gene. It consists of 591 bp, encoding a 197 amino acid polypeptide with a molecular mass of 22,307 Da. The predicted amino acid sequence of PinD, except for four C-terminal amino acids, is highly homologous to that of Gin (193 amino acids); 98.9% (191 of 193) amino acids are identical. Two amino acid residues are different in their polarity (Asp-48 to Ala and Ala-184 to Thr), but they are in the non-conserved region among known recombinases (Glasgow et al., 1989). In the upstream region of the pinD gene, a 38 bp sequence, which contains a 26 bp inv sequence for invertible recombination, was identical to the inverted right repeat (IRR) of the B segment, and the incomplete ORF1 (598 bp, 199 amino acids) including part of the IRR sequence was inferred to be part of the invertible D segment, highly similar to that of the variable region S'v (199 of 318 amino acids) in the G segment (Kahmann & Kamp, 1987).

A short sequence (positions 1–227) in the S' end of the fragment was identical to the S' end sequence of the insertion element IS600 (Matsutani et al., 1987). Another putative IS600 element (Fig. 1b) was discovered by sequencing (data not shown) and so the D segment is disrupted by insertion of two IS600 elements. Two complete ORFs, ORF2 and 3, with the same transcriptional direction as pinD, are highly homologous to the com and mom genes, respectively, on the β region of Mu. ORF2, partially overlapping with ORF3, was preceded by a perfect SD sequence and a possible promoter (Fig. 1a). Amino acid sequence homologies between ORFs and three corresponding genes are shown in Fig. 1b. Although the mom gene is followed by the attR sequence that is required for transposition, a sequence corresponding to attR was not found in the downstream region of ORF3. ORF4 was found by database searching to be 64.1% (66 of 103) and 76.8% (43 of 56) homologous to the incomplete hypothetical ORFs Aorf3 and orfx, which are present at flanking regions of plastid ATP synthase operons of the red alga Antithamnion sp. (Kostrzewa & Zetsche, 1992) and Galdieria sulphuraria (Kostrzewa & Zetsche, 1993), respectively. Since D-pinD is very similar to G-gin and B-pinB, a comparison of nucleotide sequences among three systems was carried out. Although pinD also showed highest homology for gin (567 of 591 identical bases), a 50 bp sequence located between the IRR and the pinD gene was most homologous to the corresponding region of B-pinB (44 of 50 identical bases). Among upstream regions of the IRR sequence the highest homology was observed between B and G segments (558 of 578 identical bases). This result suggests that these three systems are closely related to one another and had exchanged horizontally during evolution to generate chimeras.

Southern analysis

Many IS sequences are detected in Shigella dysenteriae and are shown to be responsible for deletions or rearrangements of flanking DNA (Galas & Chandler, 1989). Two copies of IS600 were contained in the cloned SalI fragment, and no invertible region was detected in this fragment, though the IRR sequence was present near pinD. To discover whether the region containing the invertible D segment and pinD has undergone rearrangements and whether a Mu-like genome besides the G segment and β region is present in the chromosome, hybridization analyses were carried out (Fig. 2). Sh. dysenteriae total DNA was digested with PvuII and hybridized with two kinds of Pin probes and six kinds of Mu probes that comprise almost all of the Mu genome (see Methods). Probe PIND containing pinD showed a 4.2 kb main band and a 7.2 kb weak band (Fig. 2, lane 4). The size of this main band was consistent with that of the PvuII fragment containing the D-pinD in pTSD608. In ad-
Fig. 1. Nucleotide sequences and comparison of genetic organization of the flanking regions of pinD with those of gin.
(a) Sequence of the 2979 bp HindIII-PvuII fragment is numbered from the HindIII site. The right inverted repeat sequence (IRR), which is identical to the IRR of the G segment, is boxed. The deduced amino acid sequences of five potential ORFs are shown: PinD, ORF2 and ORF3 are complete ORFs, and ORF1 and ORF4 are incomplete ORFs; their translation start and stop sites are indicated by arrows and asterisks, respectively. ORF1 is considered to start from the glycine in IRR. The Shine-Dalgarno (SD) sequence and the consensus promoter sequences (-10 and -35) are underlined. The underlined sequence in the 5' end is the 3' end sequence of attR.
(b) Restriction map of pTSD608 that carries a 10-1 kb SalI fragment and diagram of ORFs. ORFs, and the number of amino acids in them, are indicated by open arrows. Genetic organization of the G segment is in the (+) orientation and attR is the right end of Mu. Percentage amino acid sequence homologies between pTSD611 and Mu are indicated. Restriction cleavage sites: Av, AvaI; Hc, HindIII; M, MluI; Pv, PvuII; S, SalI.
Fig. 2. Southern hybridization analysis of Shigella dysenteriae DNA with probes specific for site-specific recombinase genes and Mu genome. Genome (lanes 1, 2, 4 and 5) or plasmid pTSD608 (lane 3) DNA was digested with PvuII, and processed as described in Methods and hybridized with probes: lane 1, probe MU6 containing the G segment of Mu; lanes 2 and 3, probe MU1 containing the 5.1 kb left end of Mu; lane 4, probe PIND; lane 5, probe PINE.

Fig. 3. Assay for the DNA modification. DNA samples of pTSD29 carrying ORF3 were prepared before (lanes 1 and 2) and after (lanes 3–5) induction of the ORF3 product. Three micrograms of DNA was used for PvuII digestion, which divided the plasmid into two fragments, 1.1 and 2.8 kb, and electrophoresed in a 10% agarose gel: lane 1, sample prepared before induction without PvuII digestion; lane 2, before induction; lanes 3, 4 and 5: 40, 80 and 120 min after induction, respectively.

Four kinds of Mu probes, MU2, 3, 4 and 5, which cover the central 24 kb region encoding most of structural genes, showed no hybrid band (data not shown). Two kinds of Mu probes, MU1 and MU6, showed positive bands. Probe MU1 containing the left end to kil gene region showed a 4.2 kb band (Fig. 2, lane 2), indicating the presence of a homologous region corresponding to the left end region of Mu. Probe MU6 containing tail genes, the G segment, and a part of gin gene showed a 4.2 kb main band and three weak bands, of 3.5, 6.5 and 7.2 kb (Fig. 2, lane 1), suggesting that the 4.2 and 7.2 kb fragments are the same fragments detected by probes PIND and PINE, respectively, and that Gin gene is more similar to D-pinD than to pinE. Other weak bands might have been derived from sequences homologous to some tail genes. Since the size of the genomic fragment detected by the probe MU1 is the same as that of fragment containing D-pinD, it was predicted that a region homologous to the left end of Mu was present in the cloned SauI fragment. To confirm this possibility, probe MU1 was hybridized with a PvuII digest of pTSD608, and the 4.2 kb PvuII fragment containing pinD was indeed detected (Fig. 2, lane 3). This Mu-homologous region is inferred in a 0.2 kb left end of the 4.2 kb PvuII fragment (Fig. 1b), since the region to the left of this fragment was not detected by MU1. These results indicate that regions homologous to the right and left ends of the Mu genome are close in the 4.2 kb PvuII fragment from Shigella dysenteriae, suggesting that regions homologous to both ends of Mu are remnants generated by deletion or rearrangement of an internal region of Mu-like prophage. The structure of a remnant containing a recombinase gene and one inverted repeat sequence could be an intermediate to create inversion systems in host bacteria by the appearance of secondary inverted repeat sequences (Rozsa et al., 1995). The discrepancy of the size between the cloned fragment and that detected by Southern analysis is unknown, but the strain used for the Southern analysis might have contained a minor subpopulation of cells with the DNA in the inverted orientation and it may have been this DNA which was isolated in pTSD608.

ORF3 encodes a mom homologue

Since ORF3 was significantly homologous to mom and contains only four altered amino acids apart from four extra amino acids, the function of ORF3 was examined. Mom acts to modify intracellular DNA including host DNA and co-resident plasmid DNA besides Mu DNA (Toussaint, 1976), and mom-specific modification affects digestion with Hgal, PvuII and SalI, but not BamHI and HindIII (Kahmann & Hattman, 1987). To control the expression of ORF3, plasmid pTSD29 carrying a lacZ–ORF3 fusion was constructed. To determine whether the ORF3 product modifies DNA or not, pTSD29 DNA was prepared before and after induction of ORF3 and cleaved with PvuII. Before induction, the PvuII digestion was not restricted, whereas, at least 40 min after induction the digestion was severely restricted (Fig. 3). In contrast, the cleavage with HindIII was complete in spite of ORF3 induction (data not shown),
suggested that ORF3 has the same sequence specificity to modify DNA as Mu mom.

Two inversion systems, B-pinB and D-pinD, found as a part of a defective prophage in Shigella spp. are highly homologous to that of Mu. Moreover, Southern analysis with probe PINE suggested that an e14-like element is present in *S. dysenteriae*. Although the biological function of these systems has not been determined as in P-pinE on *E. coli* e14, reproductive fitness has been reported in *E. coli* lysogenized with various phage (Edlin et al., 1975, 1977). Thus, the cryptic prophage containing the inversion system in *Shigella* spp. may produce a phenotypic change in undetectable ways under the laboratory conditions and confer a selective advantage to bacteria under some growth conditions.

REFERENCES


Two inversion systems, suggesting that ORF3 has the same sequence specificity to modify DNA as Mu mom.


Received 16 December 1996; accepted 20 January 1997.