JOURNAL OF BACTERIOLOGY, July 1996, p. 4004–4011 Vol. 178, No. 14
0021-9193/96/$04.00+0

Integration of Bacteriophage Mx8 into the *Myxococcus xanthus* Chromosome Causes a Structural Alteration at the C-Terminal Region of the IntP Protein

NOBUKI TOJO,[1] KAZUTSUKA SANMIYA,[1]† HIROYUKI SUGAWARA,[1]‡
SUMIKO INOUYE,[2]* AND TERUYA KOMANO[1]

Department of Biology, Tokyo Metropolitan University, Minamiohsawa, Hachioji-shi, Tokyo 192-03, Japan,[1] *and Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854*[2]

Received 11 January 1996/Accepted 28 April 1996

Mx8 is a generalized transducing phage that infects *Myxococcus xanthus* cells. This phage is lysogenized in *M. xanthus* cells by the integration of its DNA into the host chromosome through site-specific recombination. Here, we characterize the mechanism of Mx8 integration into the *M. xanthus* chromosome. The Mx8 attachment site, *attP*, the *M. xanthus* chromosome attachment site, *attB*, and two phage-host junctions, *attL* and *attR*, were cloned and sequenced. Sequence alignments of *attP*, *attB*, *attL*, and *attR* sites revealed a 29-bp segment that is absolutely conserved in all four sequences. The *intP* gene of Mx8 was found to encode a basic protein that has 533 amino acids and that carries two domains conserved in site-specific recombinases of the integrase family. Surprisingly, the *attP* site was located within the coding sequence of the *intP* gene. Hence, the integration of Mx8 into the *M. xanthus* chromosome results in the conversion of the *intP* gene to a new gene designated *intR*. As a result of this conversion, the 112-residue C-terminal sequence of the *intP* protein is replaced with a 13-residue sequence. A 3-base deletion within the C-terminal region had no effect on Mx8 integration into the chromosome, while a frameshift mutation with the addition of 1 base at the same site blocked integration activity. This result indicates that the C-terminal region is required for the enzymatic function of the *intP* product.

Myxococcus xanthus is a unique gram-negative bacterium living in soil. *M. xanthus* cells can undergo multicellular development involving cell-to-cell interactions (for a review, see reference 8). Upon nutritional starvation on a solid surface, cells aggregate to form mounds called fruiting bodies within which rod-shaped cells are converted into spherical or ovoid myxospores.

Several bacteriophages that infect *M. xanthus* cells are known (17). Myxophage Mx8 is a generalized transducing phage of *M. xanthus* (22). Purified phage particles have 56-kb linear double-stranded DNA molecules with an average terminal redundancy of 4.3 kb (31). Restriction analyses showed that Mx8 phage DNA is circularly permuted (see Fig. 1A). This phage can be lysogenized in *M. xanthus* cells by integrating its DNA into the host chromosome through site-specific recombination between the *attP* site on the phage DNA and the *attB* site on the host chromosome (4, 25). This recombination system has been used to introduce recombinant plasmids into the *M. xanthus* chromosome, since various plasmids containing a fragment of Mx8 DNA have been shown to stably integrate into the chromosomal *attB* site (15, 29, 31). In spite of the effectiveness and wide utilization of the Mx8 *attP*-mediated integration of plasmids, the integration mechanism is not well understood at present.

To reveal the mechanism of Mx8 site-specific recombination in *M. xanthus*, we analyzed the *intP-attP* region of Mx8 phage. Comparison of *attP*, *attB*, *attL*, and *attR* sequences revealed a 29-bp segment that is absolutely conserved. The Mx8 *intP* gene was shown to encode a basic protein with 533 amino acids that carries two domains similar in structure to those of site-specific recombinases of the integrase family. We also demonstrated that the Mx8 *attP* site is located within the coding sequence of the Mx8 *intP* gene.

MATERIALS AND METHODS

Bacteria, phage, and plasmids. *M. xanthus* DZF1 *sglA* (16) was used. *Escherichia coli* JM83 Δ(*lac-proAB*) *rpsL thi ara* φ80 d*lacZ*Δ*M15* (39) was used as a host for transformation and P1 transduction. *E. coli* JM109 *recA1* Δ(*lac-proAB*) *endA1 gyrA96 thi hsdR17 supE44 relA1*/F′ *traD36 proAB lacI*^{q}*ZΔM15* (39) was also used to construct various plasmids and to generate single-stranded DNA with phage M13 K07 (38). *E. coli* LE329 (DE3) (28) was used as a host for T7 promoter-mediated gene expression (32). P1*clr100* Cm was used to introduce various plasmids into *M. xanthus* cells from *E. coli* cells (30). pHSG399 (34), pUC19 (39), pUC118 (38), pUC119 (38), pP1inc (36), pMC1403 (6), and pET11a (32) were used as vectors for cloning and sequencing. pNT001 was constructed by replacing the 3.4-kb *Nsp*V-*Bal*I segment of pMC1403 DNA with a *Bgl*II linker. pNT002 is a derivative of pNT001 and was obtained by altering the multicloning site to GAATTCCCGGGGATCCGTCGATCGACCTGCAGAT CCC instead of GAATTCCCGGGGATCCC as in pMC1403. The last three nucleotides, CCC, of the pNT002 cloning site correspond to the eighth codon of the *lacZ* gene.

Culture conditions. *M. xanthus* cells were grown in Casitone-yeast extract (CYE) medium (16) at 30°C. Solid medium contained 1.5% Bacto Agar (Difco). Kanamycin sulfate (40 μg/ml) was added for the selection of kanamycin-resistant (Km^r) *M. xanthus* cells.

E. coli cells were grown in Luria broth or 2× YT medium (26) at 37°C. Ampicillin (100 μg/ml), chloramphenicol (25 μg/ml), tetracycline (6 μg/ml), and kanamycin sulfate (50 μg/ml) were added when necessary.

DNA manipulation and sequencing. Preparation of the chromosomal DNA and plasmid DNA, transformation, Southern hybridization, and other methods of DNA manipulation were performed as described previously (26). The DNA sequences were determined by the dideoxy chain termination method (27) in both strands, with overlapping fragments being used and single-stranded DNA being used as a template. An Applied Biosystems 373A DNA sequencing system was used according to the instructions of the manufacturer.

* Corresponding author. Mailing address: Department of Biochemistry, Robert Wood Johnson Medical School, 675 Hoes Ln., Piscataway, NJ 08854. Phone: (908) 235-4161. Fax: (908) 235-4559. Electronic mail address: sinouye@rwja.umdnj.edu.

† Present address: BioScience Center, Nagoya University, Nagoya 464-01, Japan.

‡ Present address: R & D Center, Terumo Co., Ashigarakamigun, Kanagawa 259-01, Japan.

Cloning of the *attP-intP*, *attL*, *attR*, and *attB* regions. A 2.9-kb *Sma*I fragment of the Mx8 genomic DNA *attP-intP* region was cloned into the *Sma*I site of pHSG399 to give pMP001. An integrative plasmid, pMP002 (see Fig. 2A), was constructed by ligating the 2.9-kb *Sma*I fragment and a 1.3-kb DNA fragment containing the Km^r gene from Tn*5* into the *Sma*I site of pP1inc, a vector carrying the P1 phage incompatibility region (P1*inc*), to allow the P1 transduction. pMP002 was introduced from *E. coli* into *M. xanthus* cells by P1 transduction (30) to give its integration into the chromosomal *attB* site of *M. xanthus*.

To clone the *attL* and *attR* sites, the chromosomal DNA of pMP002-integrated *M. xanthus* cells was isolated, digested with *Eco*RI or *Bam*HI, and self-ligated (9). The ligation mixture was used to transform *E. coli* JM109 cells for ampicillin resistance. pMP003, a 10.5-kb plasmid containing the *attL* region, and pMP004, an 11.2-kb plasmid containing the *attR* region, were isolated from the *Eco*RI- and *Bam*HI-digested chromosome DNA, respectively.

To clone the *attB* site, a 2.6-kb *Bam*HI-*Kpn*I fragment located near the *attB* site on the chromosome was isolated from pMP003 DNA, treated with DNA polymerase I Klenow fragment (Klenow enzyme), and inserted into the *Hin*cII site of pUC19. A 6.3-kb *Bam*HI fragment containing the P1*inc* and Km^r genes was inserted into the *Bam*HI site of the resulting plasmid to give pMP005 (see Fig. 2C). pMP005 was introduced into *M. xanthus* DZF1 by P1 transduction, and Km^r colonies were isolated. From the Km^r *M. xanthus* cells, the chromosomal DNA was isolated, digested with *Bam*HI, self-ligated, and used to transform *E. coli* JM109 cells for ampicillin resistance. pMP006, a 7.1-kb plasmid containing the 4.4-kb *Bam*HI fragment carrying the *M. xanthus attB* site (see Fig. 2A), was obtained.

Construction of deletion and frameshift mutations in the *intP* gene and assay of their integration ability. To construct plasmids with deletions from the left end of the Mx8 sequence (pMP007 through pMP012 [see Fig. 5]), pMP001 DNA was first digested at the unique *Bam*HI site in the vector and then digested completely or partially at various restriction sites within the insert. After treatment with the Klenow enzyme, the resulting fragments were ligated with the Km^r DNA fragment. For deletions from the right end of the Mx8 sequence (pMP013 through pMP016), pMP001 DNA was digested at the *Kpn*I site in the vector and at various restriction sites in the insert, treated with the Klenow enzyme, and religated. To generate frameshift mutations (pMP017 through pMP021), pMP001 DNA was digested with an appropriate restriction enzyme, treated with the Klenow enzyme, and self-ligated. The DNA junctions produced in various deletion and frameshift mutations were confirmed by either restriction enzyme analyses or DNA sequencing. All plasmids contained the Km^r gene at the left end of Mx8 DNA in the same direction as that of the *intP* gene (see Fig. 5). All the plasmids carrying the mutations described above were converted into the pMP002 configuration as described earlier, yielding pMP007 to pMP021.

The final plasmids were transferred from *E. coli* to *M. xanthus* cells by P1 transduction to test their ability to integrate into the *M. xanthus* chromosome via site-specific recombination.

Construction of *intP-lacZ* fusion genes and β-galactosidase assay in *M. xanthus*. Two types of *intP-lacZ* fusion genes were constructed (6). In the case of the construction of plasmids pMP022 to pMP027, 5′-end-truncated fragments of various sizes (see Fig. 6 [thin lines from pMP022 to pMP027]) were first generated by digesting pMP001 DNA with various restriction enzymes, and then the fragments were treated with the Klenow enzyme and digested with *Stu*I. The isolated fragments were then inserted at the *Sma*I site of pNT002 so that the *intP* gene was fused at the glutamic acid residue at position 65 with the *lacZ* gene in the same reading frame. In other constructions (pMP028 to pMP030), the fusion was carried out downstream of the *attP* site. For this purpose, the *Not*I site in the *intP* gene of pMP001 DNA (see Fig. 6) was first converted to an *Nae*I site with the Klenow enzyme. Subsequently, the 2.66-kb *Sma*I-*Nae*I fragments of the resulting plasmid were inserted into the *Sma*I site of pNT001 to create the *intP-lacZ* fusion (see Fig. 6 [pMP028, -029, and -030]). In both constructions, pMP001 derivatives with frameshift mutations were also used. In all constructs (pMP022 to pMP030), a 7.8-kb *Eco*RI fragment containing the P1*inc* region, the Tc^r gene of pBR322, and a part of the *M. xanthus lonD* coding sequence (37) was inserted into the *Eco*RI site of the pNT001 and pNT002 derivatives carrying the *intP-lacZ* fusion genes. The *lonD* gene was used for the integration site for all the constructs described above. Since the plasmids lack the *intP* function, they are integrated into the *lonD* locus on the *M. xanthus* chromosome by homologous recombination after being introduced by P1 transduction. The chromosome structure of each integrant was confirmed by Southern hybridization analysis. The β-galactosidase activity of *M. xanthus* cells carrying *intP-lacZ* fusion genes was measured with *o*-nitrophenyl-β-D-galactopyranoside being used as the substrate and was expressed as nanomoles of *o*-nitrophenol per minute per milligram of protein as described previously (21).

Expression of *intP* and *intR* genes in *E. coli*. The *intP* and *intR* genes were inserted into a T7 expression vector, pET11a, and overexpressed in *E. coli* cells (32). At first, the sequence corresponding to the N-terminal region of the *intP* and *intR* genes was amplified by PCR with pMP001 DNA being used as the template. PCR was carried out with a forward primer (5′-CTCATATGGGTA ACGTCTATC-3′) containing an ATG initiation codon at the *Nde*I site to replace putative initiation codon GTG of the *intP* and *intR* genes and a reverse primer (5′-CTGGATCCACGCGTCAGCACACC-3′) corresponding to the sequence around the *Mlu*I site of the *intP* and *intR* genes. This primer also creates a *Bam*HI site downstream of the *Mlu*I site. The PCR product was digested with *Bam*HI and inserted into the *Sma*I-*Bam*HI site of pUC19 to give pMP031. To generate expression plasmids with the entire *intP* and *intR* genes, pMP031 DNA was digested with *Mlu*I and *Bam*HI and ligated with 1.4- and 1.2-kb *Mlu*I-*Bam*HI fragments containing the sequences corresponding to the C-terminal regions of the *intP* and *intR* genes to give pMP032 and pMP033, respectively. The *Nde*I-*Bam*HI fragments of pMP032 and pMP033 DNA were then inserted at the *Nde*I-*Bam*HI sites of pET11a to give pMP034 and pMP035, respectively.

For the expression of *intP* and *intR* genes, *E. coli* LE329 (DE3) cells harboring pMP034 or pMP035 were grown to an A_{620} of 0.5 and treated with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 30 min at 37°C to induce T7 RNA polymerase expression. Rifampin (final concentration, 200 μg/ml) was added into the culture, and incubation was continued for an additional 30 min. Then, 10 μCi of [^{35}S]methionine was added to 2 ml of the cultures, and the cells were labeled for 10 min. Protein patterns were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 17.5% gel. Proteins were visualized by both staining with Coomassie brilliant blue and autoradiography.

Materials. Restriction enzymes, the DNA polymerase I Klenow fragment, oligonucleotide linkers, the DNA ligation kit, and the nick translation kit were purchased from either New England Biolabs, Takara Shuzo, or Nippon Gene. [α-^{32}P]dCTP and [^{35}S]methionine were from ICN Biomedicals.

Nucleotide sequence accession numbers. The entire nucleotide sequence data for *attB*, *attL*, *intP-attP*, and *attR* described in this paper will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases under accession numbers D26557, D26558, D26559, and D26560, respectively.

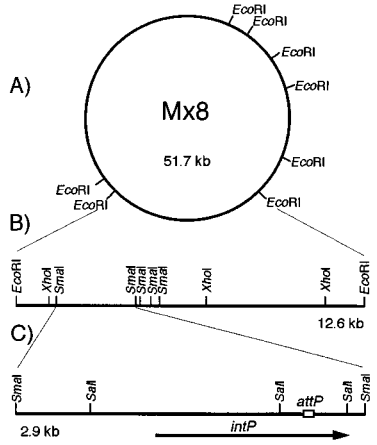

FIG. 1. (A) Circular restriction map of the 51.7-kb Mx8 genome (31). (B) Restriction map of the 12.6-kb *Eco*RI fragment. Only *Sma*I sites located within the 5.6-kb *Xho*I fragment are indicated. (C) Restriction map of the 2.9-kb *Sma*I fragment. The open bar on the map indicates the location of the *attP* site. The arrow below the map represents the *intP* coding region and its orientation.

RESULTS

Localization of the Mx8 attachment site. Previously, it was reported that plasmids containing the 12.6-kb *Eco*RI fragment (Fig. 1B) of Mx8 DNA could integrate into a unique site on the *M. xanthus* chromosome via site-specific recombination (29, 31). Later, the internal 2.9-kb *Sma*I fragment (Fig. 1C) was identified as a fragment which is still capable of integrating itself into the *M. xanthus* chromosome (15).

Hence, we constructed an integrative plasmid, pMP002, by using the 2.9-kb *Sma*I fragment for in situ cloning of the *attL* and *attR* sites (Fig. 2A). This plasmid was introduced into *M. xanthus* DZF1 cells from *E. coli* cells by P1 transduction (30). Many kanamycin-resistant transductants were obtained. These transductants were likely to be produced by site-specific recombination between the *attP* site on pMP002 and the *attB* site on the *M. xanthus* chromosome as shown in Fig. 2B. Chromosomal DNAs of a few independent transductants were isolated and subjected to Southern hybridization analysis with pUC19 DNA being used as a probe. The results indicated that in all cases, one copy of pMP002 DNA was integrated into the

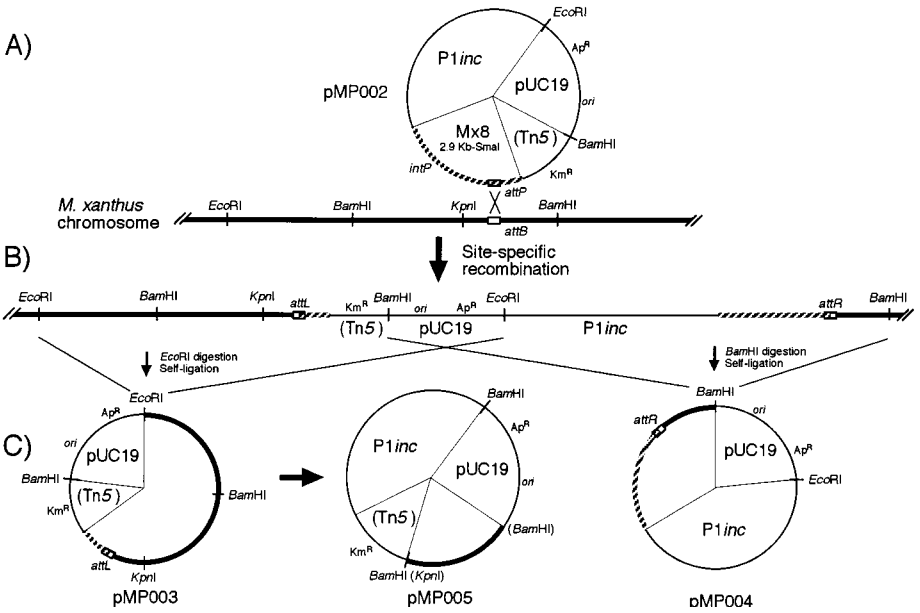

FIG. 2. Schematic diagram of cloning of the Mx8 *attL*, *attR*, and *attB* sites. (A) Structures of pMP002 and *M. xanthus* chromosome DNA around the *attB* site. Hatched line, the 2.9-kb *Sma*I fragment of Mx8 containing the *attP* site (thick hatched bar) and the *intP* gene; bold line, *M. xanthus* chromosomal DNA with the *attB* site (open bar); thin line, the vector sequence. (B) Structure of the *M. xanthus* chromosome with pMP002 integration. The *attL* and *attR* sites are indicated by bars. (C) Structures of pMP003, pMP004, and pMP005. pMP005 was used for cloning of the *attB* site. Only the relevant sites are shown for *Kpn*I. Restriction sites in parentheses were treated with the Klenow enzyme.

unique site, *attB*, on the *M. xanthus* chromosome (data not shown). Furthermore, the approximate locations of the *Eco*RI and *Bam*HI sites adjacent to the *attB* site on the chromosome were deduced (Fig. 2B).

Isolation of *attP*, *attL*, *attR*, and *attB* sites and identification of the core sequence. To clone the *attL* and *attR* sites, the chromosomal DNA from one pMP002-integrated *M. xanthus* strain was digested with *Eco*RI or *Bam*HI. Since both the origin and the ampicillin-resistant gene in the plasmids used are derived from pUC9 and flanked by a *Bam*HI site and an *Eco*RI site (Fig. 2B), self-ligation of the chromosomal DNA digests either by *Eco*RI or *Bam*HI generates ampicillin-resistant plasmids containing either *attL* (by *Eco*RI digestion) or *attR* (by *Bam*HI digestion), thus generating pMP003 and pMP004, respectively. Approximately 50 ampicillin-resistant colonies were obtained for each digestion.

In the case of *Eco*RI digestion, pMP003, a 10.5-kb plasmid containing a 6.5-kb *attL* host-phage junction region, was obtained (Fig. 2C). In the case of *Bam*HI digestion, pMP004, an 11.2-kb plasmid containing a 3.6-kb *attR* region, was obtained.

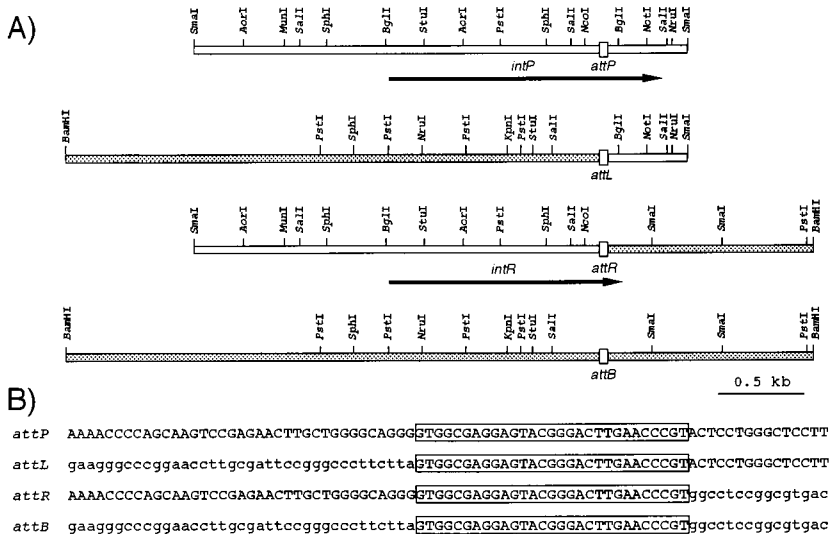

FIG. 3. (A) Restriction maps around the *attP*, *attL*, *attR*, and *attB* sites. Restriction sites in the 2.9-kb *Sma*I fragment for *attP* and *intP*, the 3.6-kb *Bam*HI-*Sma*I fragment for *attL*, the 3.6-kb *Sma*I-*Bam*HI fragment for *attR* and *intR*, and the 4.4-kb *Bam*HI fragment for *attB* are shown. The location of the 29-bp core sequence is shown by an open bar on the map. The locations and orientations of the *intP* and *intR* genes are shown by arrows. (B) Alignment of the *attP*, *attL*, *attR*, and *attB* sequences. The 29-bp core sequences are boxed. Chromosomal DNA sequences are indicated by lowercase letters.

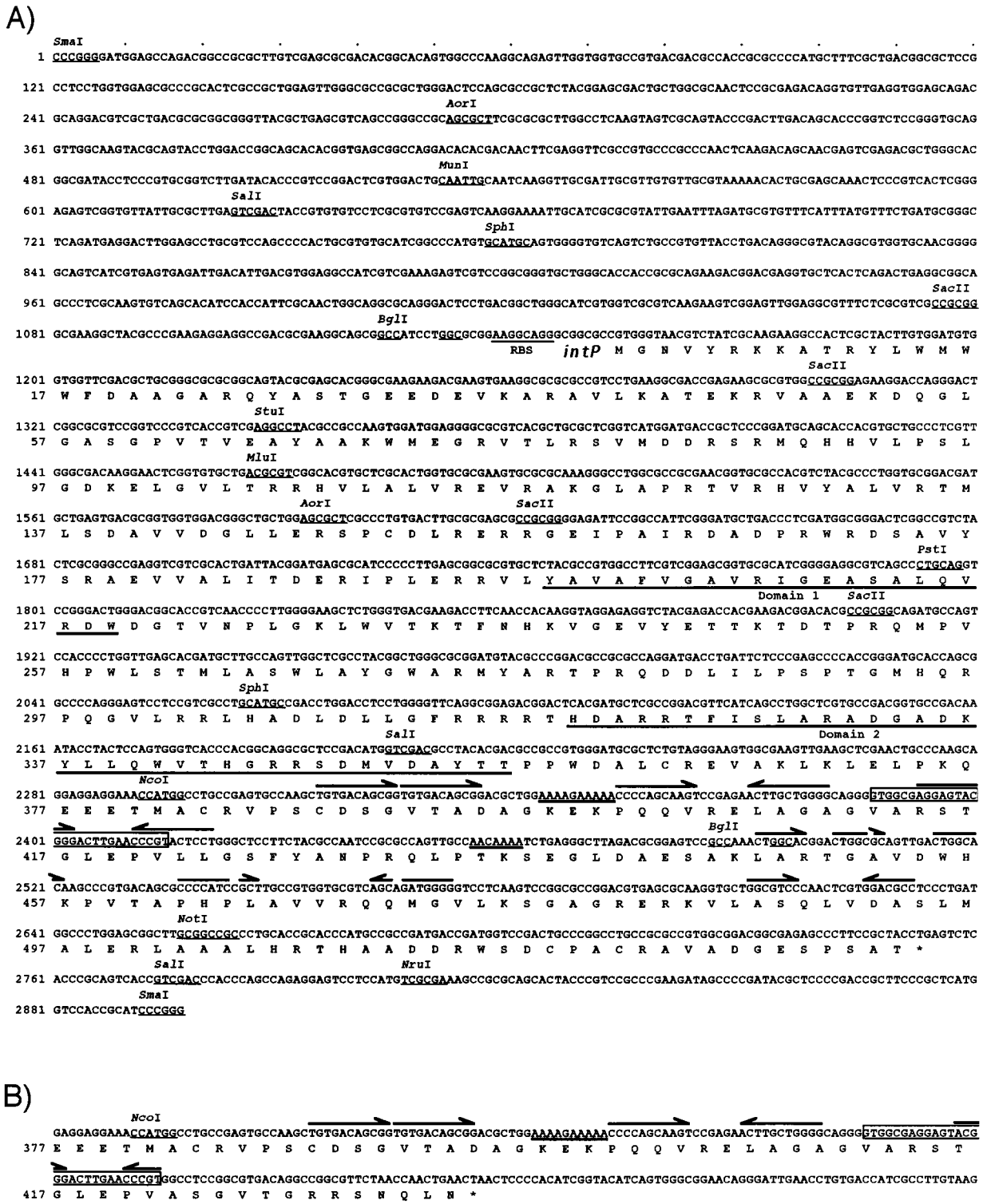

FIG. 4. (A) Nucleotide sequence of the 2,897-bp *Sma*I fragment of Mx8 and the deduced amino acid sequence of the *intP* gene. (B) Nucleotide sequence around the *attR* site and deduced amino acid sequence of the *intR* C-terminal region. A putative ribosome-binding site (RBS), direct and inverted repeats (arrows), *attP* and *attR* core sequences (boxes), relevant restriction sites (underlines), and A-clusters (underlines) are indicated. The amino acid sequences of domains 1 and 2 of the IntP protein, conserved in site-specific recombinases of the Int family (Fig. 9), are also indicated.

From the restriction enzyme analyses of pMP002, pMP003, and pMP004 DNA, the recombination sites in *attP*, *attL*, and *attR* were located within the 0.6-kb *Sal*I fragment of pMP002 DNA, the 0.7-kb *Sal*I fragment of pMP003 DNA, and the 0.4-kb *Sal*I-*Sma*I fragment of pMP004 DNA, respectively (Fig. 3A). These fragments were subcloned and sequenced (Fig. 3B).

Next, we also cloned the *attB* region by integrating pMP005

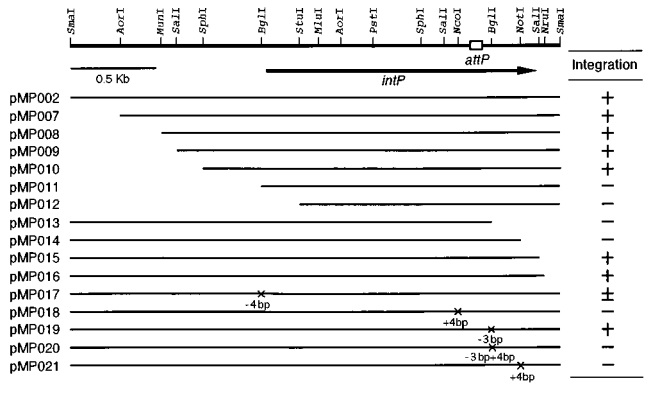

FIG. 5. Integration ability of pMP002 and its derivative plasmids into the *M. xanthus* chromosomal *attB* site. At the top, a restriction map of the 2.9-kb *Sma*I fragment of Mx8 is shown together with the *attP* core sequence (open bar). A long arrow represents the coding sequence and orientation of the *intP* gene. Solid lines below the map indicate DNA portions present in various plasmids. Crosses on the line indicate the mutations constructed by modifying various restriction sites. Integration of these plasmids into the *attB* site was determined by the formation of *M. xanthus* Km^r colonies after P1 transduction. +, high frequency; ±, low frequency; −, no transduction.

(Fig. 2C), which contained the 2.6-kb *Bam*HI-*Kpn*I fragment of pMP003 DNA, as well as P1*inc* and Km^r DNA, into the *M. xanthus* chromosome via homologous recombination. By religating the *Bam*HI digest of the pMP005-integrated *M. xanthus* chromosomal DNA, we obtained pMP006, which carried the 4.4-kb *Bam*HI fragment (Fig. 2A) encompassing the *attB* site. The restriction enzyme analyses of pMP006 DNA indicated that the *attB* site was located within the 0.6-kb *Sal*I-*Sma*I fragment (Fig. 3A). The 0.6-kb *Sal*I-*Sma*I fragment was subcloned and sequenced (Fig. 3B).

Figure 3B shows the alignment of the *attP*, *attL*, *attR*, and *attB* sequences. Each sequence was found to contain an identical 29-bp sequence designated as the core sequence. Site-specific recombination between the phage *attP* and host *attB* sites is considered to occur within the core sequence on the basis of the sequences flanking the core sequence in *attL*, *attR*, *attB*, and *attP*, as in the case of λ phage integration (4). The 29-bp Mx8 *attP* core sequence was located near the right end of the 2.9-kb *Sma*I fragment of Mx8 DNA (Fig. 1C).

Nucleotide sequence of Mx8 *intP-attP* region. Since pMP002 carrying the 2.9-kb *Sma*I fragment of Mx8 DNA was inserted into the chromosomal *attB* site of *M. xanthus* at a high frequency, the *Sma*I fragment is considered to bear all information required for site-specific recombination. Hence, the entire nucleotide sequence of the 2,897-bp *Sma*I fragment of Mx8 was determined (Fig. 4A). Analysis of the sequence revealed a long open reading frame (ORF), tentatively designated the *intP* gene, whose sequence shows significant homology to those of several phage and plasmid site-specific recombinases belonging to the integrase family (see Fig. 9). As shown later, translation of the *intP* gene was postulated to start with the GTG codon at position 1154 and end with the TGA codon at position 2753. A purine-rich Shine-Dalgarno-like sequence (AAGGCAGG) separated by 8 bp from the initiation codon was found. Thus, the *intP* gene was deduced to encode a protein with 533 amino acids and a calculated molecular weight of 59,400. The calculated isoelectric point of the IntP protein is 9.9. The codon usage in the *intP* gene is similar to that in other *M. xanthus* genes, and the codon used in ORF has a high G+C content at the third position (12). There is another ORF with 268 amino acid residues from the GTG codon at position 355 to the TAA termination codon at position 1159 whose sequence does not show any significant homology to those of known proteins as determined by a search with the BLAST program.

To our surprise, the *attP* core sequence described above was located within the *intP* coding sequence. Thus, the integration of Mx8 phage or its derivative plasmids into the *attB* site of the host chromosome results in a replacement of the C-terminal 112-amino-acid segment of the *intP* gene from the *attP* site (Fig. 4A) with the C-terminal 13-amino-acid segment from the *attR* site (Fig. 4B). The new ORF, designated the *intR* gene, now encodes a protein with 434 amino acids and a calculated molecular weight of 48,863. The calculated isoelectric point of the IntR protein is 10.1.

The neighboring region of the Mx8 *attP* sequence contains several direct and inverted repeats, located asymmetrically, like the *attP* region of various temperate phages (7, 23, 35) (Fig. 4A). The 29-bp *attP* core sequence overlaps with 11-bp inverted repeats (Fig. 4A). These sequences might be important for the integration of the phage DNA and could present recognition sites for integrase and another DNA-binding protein(s) (7, 23, 24, 35). In addition, A-cluster sequences, known to cause a structural change in DNA (13), were found at nucleotide positions 2344 to 2353 and 2455 to 2461. They might also play important roles in the recombination event.

Analysis of integration with deletion and frameshift mutations of the *intP* gene. The finding that the *attP* core sequence is located within the coding region of the *intP* gene raises an interesting question about the role of convertible C-terminal segments for IntP and IntR proteins. To determine whether the entire region of the *intP* gene is necessary for the site-specific integration of Mx8 into the *attB* site on the *M. xanthus* chromosome, we carried out deletion and frameshift analysis of the DNA segment in pMP002. All the mutants used are shown in Fig. 5. It was found that the plasmids from pMP007 through pMP010 carrying deletions at the region upstream of the *intP* gene were able to integrate themselves at the *M. xanthus* chromosomal *attB* site, since many Km^r colonies were obtained, while pMP011 and pMP012 were not able to integrate (Fig. 5). In the case of the deletions at the downstream region of the 29-bp core sequence, pMP013 and pMP014 could

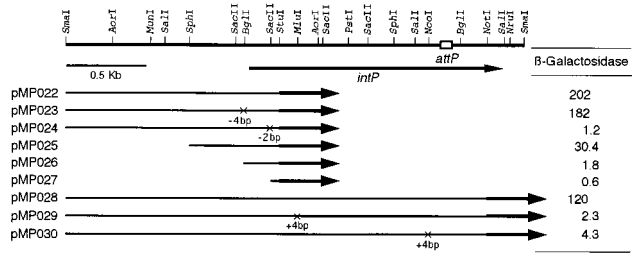

FIG. 6. Expression of *intP-lacZ* fusion genes during vegetative growth of *M. xanthus*. Two types of *intP-lacZ* fusion genes together with various lengths of the upstream control sequence of Mx8 were constructed and introduced into a P1-transducing phage as described in the text. These plasmids were then introduced into *M. xanthus* cells by P1 transduction. In these experiments, the entire plasmid DNA was integrated into the *lonD* locus (37) on the *M. xanthus* chromosome by single homologous recombination. At the top, a restriction map of the 2.9-kb *Sma*I fragment of Mx8 is shown. A long arrow represents the coding sequence and orientation of the *intP* gene. In the various plasmids, a thin line and a bold arrow represent the Mx8 sequence and the *E. coli lacZ* gene, respectively. Crosses on the line indicate the mutations constructed by modifying various restriction sites. In these plasmids, the Km^r gene is located at the left end of the Mx8 sequence in the rightward direction. The β-galactosidase activities of *M. xanthus* cells carrying the *intP-lacZ* protein fusion gene were determined and are indicated in nanomoles of *o*-nitrophenol per minute per milligram of protein.

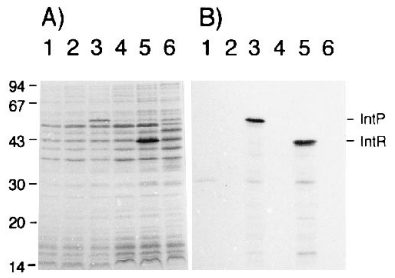

FIG. 7. Expression of *intP* and *intR* genes in *E. coli*. *E. coli* LE329 (DE3) cells carrying pET11a (lanes 1 and 2), pMP034 (lanes 3 and 4), and pMP035 (lanes 5 and 6) were treated (lanes 1, 3, and 5) or not (lanes 2, 4, and 6) with IPTG and then were treated with rifampin. The cells were then labeled with [^{35}S]methionine. Total cellular proteins were analyzed by SDS-PAGE. (A) Coomassie brilliant blue staining. (B) Autoradiogram. The locations of molecular mass markers (in kilodaltons) are indicated on the left. The locations of the IntP and IntR proteins are indicated on the right.

not integrate into the *M. xanthus* chromosome, while pMP015 and pMP016 could. These results indicate that the minimum DNA region required for Mx8 site-specific integration is located within the 2.0-kb *Sph*I-*Sal*I segment which contains the entire *intP* gene together with the *attP* site (Fig. 5). This finding could be explained by any of three possibilities: the DNA sequence after the *attP* core sequence was required (i) for the *attP* function, (ii) for encoding the IntP protein, or (iii) for both. In fact, it is known that the 240-bp *attP* sequence is required in *cis* for integration of the λ phage into the *E. coli* chromosome (35).

Hence, we constructed a number of frameshift mutations within the 2.0-kb *Sph*I-*Sal*I fragment to test these possibilities. pMP017, pMP018, pMP019, pMP020, and pMP021 carried a 4-bp deletion at the *Bgl*I site (position 1123 in Fig. 4A), a 4-bp insertion at the *Nco*I site (position 2292 in Fig. 4A), a 3-bp deletion at the *Bgl*I site (position 2486 in Fig. 4A), a replacement of AAC by TGCA at the *Bgl*I site (position 2486 in Fig. 4A), and a 4-bp insertion at the *Not*I site (position 2657 in Fig. 4A), respectively (Fig. 5). pMP019 and pMP017 could integrate themselves into the *M. xanthus* chromosomal *attB* site at a normal and a reduced frequency, respectively, while pMP018, pMP020, and pMP021 could not.

These results indicate that the entire *intP* region, including the C-terminal segment after the *attP* core sequence, is indispensable for Mx8 site-specific integration into the *M. xanthus* chromosome. Especially telling is the finding that the 3-base deletion at the *Bgl*I site (position 2486) did not abolish the integration activity while the introduction of a frameshift with the addition of a base at the same point blocked the activity. This result strongly supports our prediction. In addition, some of the sequence upstream beyond the putative translation initiation site was found to also be necessary for integration. This sequence is likely to be required for the expression of the *intP* gene.

Expression of the *intP-lacZ* fusion gene in *M. xanthus*. That pMP016 carrying a 4-bp deletion at the *Bgl*I site (position 1123) still had an integration activity suggests that the translation initiation site of the *intP* gene is located downstream of the *Bgl*I site. To determine the initiation site of the *intP* ORF and to verify the requirement for the upstream regions, a series of *intP-lacZ* protein fusion genes were constructed (Fig. 6). Plasmids pMP022 through pMP027 carried *intP-lacZ* fusion genes in which the *intP* gene was fused at the 65th codon to the *E. coli lacZ* gene together with various lengths of the upstream region as described in Materials and Methods. These plasmids were integrated into the *lonD* locus (37) on the *M. xanthus* chromosome by homologous recombination.

M. xanthus cells carrying pMP022 expressed about 200 U of β-galactosidase activity per mg of protein (Fig. 6), indicating that the Mx8 *intP* gene is expressed during vegetative growth of *M. xanthus* and that the initiation codon of the *intP* gene is located upstream of the *Stu*I site (position 1347). The frameshift mutation at the *Bgl*I site in pPM023 did not affect β-galactosidase expression, while the frameshift mutation at the *Sac*II site (position 1299) in pPM024 did, indicating that the initiation codon of the *intP* gene is located between the *Bgl*I and *Sac*II sites. Within this region, there are the following four possible initiation codons: GTG at position 1154, ATG at position 1196, GTG at position 1256, and GTG at position 1295 (Fig. 4A). However, we tentatively assigned the GTG codon at base 1154 as the initiation codon for the *intP* gene, because it accompanies a putative purine-rich Shine-Dalgarno-like sequence as indicated in Fig. 4A (underlined). *M. xanthus* cells carrying pMP025 produced β-galactosidase (30 U/mg of protein) while those carrying pMP026 or pMP027 did not, suggesting that the 0.35-kb *Sph*I-*Bgl*I upstream fragment is also required for expression of the *intP* gene.

To confirm that translation of the *intP* gene continues beyond the *attP* core sequence, another type of *intP-lacZ* protein fusion gene was constructed (Fig. 6). Plasmids pMP028, pMP029, and pMP030 carried the *intP-lacZ* fusion gene in which the *intP* gene was fused to the *lacZ* gene at the 503th codon at the *Not*I site near the C-terminal region of the *intP* gene as described in Materials and Methods. These plasmids were also integrated into the *lonD* locus on the *M. xanthus* chromosome. As shown in Fig. 6, *M. xanthus* cells carrying pMP028 produced β-galactosidase at about 120 U/mg of protein. On the other hand, frameshift mutations at the *Mlu*I and *Nco*I sites (positions 1466 and 2292, respectively) in pMP029 and pMP030 abolished β-galactosidase expression. These results clearly demonstrate that *intP* translation continues beyond the *attP* site.

Expression of *intP* and *intR* genes in *E. coli*. To demonstrate that the *intP* and *intR* gene products have different molecular weights, both *intP* and *intR* were expressed in *E. coli* with a T7 expression system (32). After an *Nde*I site was created at the initiation codons of the *intP* and *intR* genes, their coding sequences were inserted into a T7 expression vector, pET11a, to give pMP034 and pMP035, respectively, as described in Materials and Methods.

The *E. coli* LE329 (DE3) cells harboring pMP034 and pMP035 were grown in the presence and absence of IPTG.

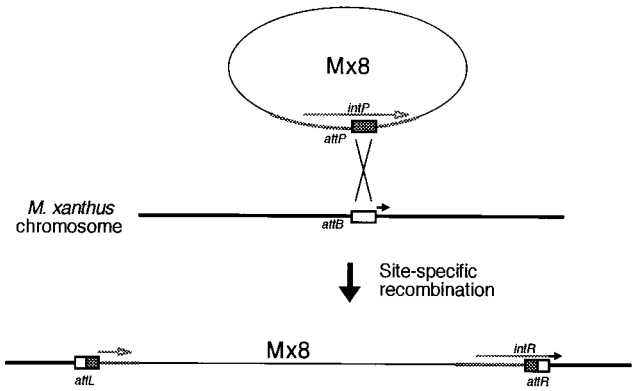

FIG. 8. Schematic illustration for the integration of Mx8 phage by site-specific recombination.

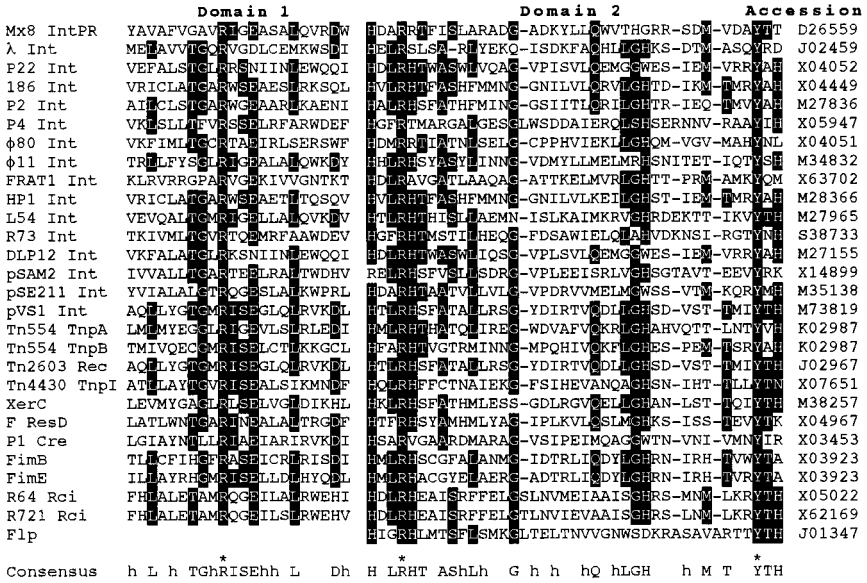

FIG. 9. Alignment of the amino acid sequences of two conserved domains among the site-specific recombinases of the integrase family. Amino acid residues identical in more than 11 proteins in this alignment are printed in reverse type. Accession numbers for international nucleotide sequence databases are indicated. A consensus sequence is shown at the bottom, with three completely conserved amino acids being indicated with asterisks. "h" represents hydrophobic amino acid residues.

The cells were then labeled with [^{35}S]methionine in the presence of rifampin as described in Materials and Methods. The total cellular proteins were then analyzed by SDS-PAGE (Fig. 7). Proteins which were induced in the presence of IPTG are even visible in the protein patterns produced by staining with Coomassie brilliant blue (lanes 3 and 5 in Fig. 7A). These protein bands are clearly detected by autoradiography (lanes 3 and 5 in Fig. 7B). The major band, which has a molecular mass of 57 kDa and which is the product of *intP*, was observed for cells harboring pMP034. On the other hand, the major product from pMP035 has an apparent molecular mass of 47 kDa and is the product of *intR*. These molecular masses agree well with the calculated molecular masses of IntP (59 kDa) and IntR (49 kDa).

DISCUSSION

In the present study, the DNA sequences of *attP*, *attL*, *attR*, and *attB* required for site-specific recombination of myxophage Mx8 were determined, and it was found that they share a 29-bp core sequence. We also identified the *intP* gene of Mx8, which encodes a basic protein with 533 amino acids. The striking feature of the Mx8 integration system is that the Mx8 *attP* site is located within the *intP* coding sequence. Thus, integration of Mx8 into the *M. xanthus* chromosomal *attB* site results in a replacement of the 112-amino-acid C-terminal segment of the *intP* gene with a sequence consisting of 13 residues to create the *intR* gene. *E. coli* cells harboring the *intP*$^+$ plasmid, pMP034, produced a 57-kDa protein, while those with the *intR*$^+$ plasmid, pMP035, produced a 47-kDa protein. Conversion of the *intP* gene into the *intR* gene by Mx8 integration into the *M. xanthus* chromosome is shown schematically in Fig. 8.

There are a number of site-specific inversion systems in which one of recombination sites is located within the coding sequence of the target genes. In Mu and P1 phages, the inversion of the G and C DNA segments controls the alternative expression of S genes for Mu and 19 genes for P1 with different C-terminal segments, which determines the host specificity of both phages (3, 14, 18). In the shufflon of plasmid R64, the multiple inversion of four DNA segments selects one of seven different *pilV* genes with different C-terminal segments, which determines recipient specificity in liquid mating of R64 (19, 20). In H and type 1 fimbria systems, the inversion of a DNA segment changes the direction of a promoter located within the invertible DNA segment to turn on or turn off the expression of target genes (1, 40).

Various temperate phages and integrative plasmids are integrated into and excised from the chromosomal *attB* sites of host cells. The *attB* sites are often found within tRNA genes and intergenic spaces (5). When an *attB* site is located within a host gene, the 3′ end of the gene is duplicated in the lysogenized phage to preserve the structure of the host gene. For example, retrophage ϕR73 is integrated into the selenocystyl-tRNA gene of the host chromosome (33), and the 29-bp sequence present in the *attP* site of ϕR73 is identical to the 29-bp sequence in the tRNA gene.

In both inversion and integration-resolution systems, site-specific recombination genes are usually located near the recombination sites. In H inversion, the *hin* gene is located within the invertible H segment (40). It should be noted that in all the site-specific recombination systems known so far, the recombination sites are always located outside the coding sequence of the recombinase genes. To our knowledge, the Mx8 integration system is the first example in which the recombination site is located within the coding sequence of the site-specific recombinase gene.

Two major groups of site-specific recombinases are known. One group is the Din family, including Hin, Gin, Cin, and Pin invertases and resolvases of Tn*3*, Tn*21*, Tn*501*, and Tn*1000* (10). Another group is the Int family, including various Int proteins of phages (λ, P22, 186, P2, P4, ϕ80, ϕ11, FRAT1, HP1, L54, R73, and DLP12) and integrative plasmids (pSAM2, pSE211, and pVS1), resolvases of the *E. coli* chromosome (XerC), F and P1 plasmids (ResD and Cre, respectively) and various transposons (Tn*554*, Tn*2603*, and Tn*4430*), and invertases of the *E. coli* chromosome (FimB and FimE) and bacte-

rial and yeast plasmids (Rci and Flp, respectively) (Fig. 9). Although site-specific recombinases in the Int family are diversely differentiated during evolution, they share two conserved domains (2). Figure 9 shows the alignment of amino acid sequences of two conserved domains among various site-specific recombinases from the Int family. The alignment indicates that the Mx8 IntP (or IntR) protein belongs to the Int family. The constant region of the IntP protein shows the highest similarity to the Int protein of mycobacteriophage FRAT1 (23% identity over 317 amino acids) (11). In all site-specific recombinases in the Int family, three amino acids (R in domain 1 and R and Y in domain 2) are completely conserved (Fig. 9). The conserved tyrosine residue in domain 2 is known to make a phosphodiester bond with the 3′ end of the recombining DNA strand as a reaction intermediate. The Mx8 IntP protein carries an additional C-terminal domain with about 100 amino acids compared with other Int proteins. A 3-base deletion within the C-terminal region had no effect on Mx8 integration into the chromosome, while a frameshift mutation by the addition of 1 base at the same side blocked integration activity. This result indicates that the C-terminal region is required for the enzymatic function of the *intP* product.

In the DNA sequence shown in Fig. 4, there is another ORF starting with the GTG codon at position 355 and ending with the TAA codon at position 1159 and overlapping the *intP-intR* gene by 5 nucleotides. Whether this ORF functions as an *xis* gene remains to be determined.

ACKNOWLEDGMENTS

We thank M. Inouye for discussions and a critical reading of the manuscript and T. Furuichi for plasmids.

A part of this work was supported by NIH grant GM26843 (to S.I.) and JSPS Research Fellowship for N.T.

REFERENCES

1. **Abraham, J. M., C. S. Freitag, J. R. Clements, and B. I. Eisenstein.** 1985. An invertible element of DNA controls phase variation of type I fimbriae of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **82:**5724–5727.
2. **Argos, P., A. Landy, K. Abremski, J. B. Egan, E. Haggard-Ljungquist, R. H. Hoess, M. L. Kahn, B. Kallionis, S. V. L. Narayana, L. S. Pierson III, N. Sternberg, and J. M. Leong.** 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. **5:**433–440.
3. **Bukhari, A. I., and L. Ambrosio.** 1978. The invertible segment of bacteriophage Mu DNA determines the absorption properties of Mu particles. Nature (London) **271:**575–577.
4. **Campbell, A. M.** 1962. Episomes. Adv. Genet. **11:**101–145.
5. **Campbell, A. M.** 1992. Chromosomal insertion sites for phages and plasmids. J. Bacteriol. **174:**7495–7499.
6. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-glucosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translation initiation signals. J. Bacteriol. **143:**971–980.
7. **Craig, N. L.** 1988. The mechanism of conservative site-specific recombination. Annu. Rev. Genet. **22:**77–105.
8. **Dworkin, M., and D. Kaiser.** 1993. Myxobacteria II. American Society for Microbiology, Washington, D.C.
9. **Furuichi, T., M. Inouye, and S. Inouye.** 1985. Novel one-step cloning vector with a transposable element: application to the *Myxococcus xanthus* genome. J. Bacteriol. **164:**270–275.
10. **Glasgow, A. C., K. T. Hughes, and M. I. Simon.** 1989. Bacterial DNA inversion systems, p. 637–659. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
11. **Haeseleer, F., J. F. Pollet, A. Bollen, and P. Jacobs.** 1992. Molecular cloning and sequencing of the attachment site and integrase gene of the temperate mycobacteriophage FRAT1. Nucleic Acids Res. **20:**1420.
12. **Hagen, T. J., and L. J. Shimkets.** 1990. Nucleotide sequence and transcriptional products of the *csg* locus of *Myxococcus xanthus*. J. Bacteriol. **172:**15–23.
13. **Hagerman, P. J.** 1990. Sequence-directed curvature of DNA. Annu. Rev. Biochem. **59:**755–781.
14. **Hiestand-Nauer, R., and S. Iida.** 1983. Sequence of the site-specific recombinase gene *cin* and of its substrates serving in the inversion of the C segment of bacteriophage P1. EMBO J. **2:**1733–1740.
15. **Hsu, M.-Y., S. Inouye, and M. Inouye.** 1989. Structural requirements of the RNA precursor for the biosynthesis of the branched RNA-linked multicopy single-stranded DNA of *Myxococcus xanthus*. J. Biol. Chem. **264:**6214–6219.
16. **Inouye, M., S. Inouye, and D. R. Zusman.** 1979. Gene expression during development of *Myxococcus xanthus*: pattern of protein synthesis. Dev. Biol. **68:**579–591.
17. **Kaiser, D.** 1984. Genetics of myxobacteria, p. 163–184. *In* E. Rosenberg (ed.), Myxobacteria. American Society for Microbiology, Washington, D.C.
18. **Kamp, D., R. Kahman, D. Zipser, T. R. Broker, and L. T. Chow.** 1978. Inversion of the G DNA segment of phage Mu controls phage infectivity. Nature (London) **271:**577–580.
19. **Komano, T., S.-R. Kim, and T. Yoshida.** 1995. Mating variation by DNA inversions of shufflon in plasmid R64. Adv. Biophys. **31:**181–193.
20. **Komano, T., S.-R. Kim, T. Yoshida, and T. Nisioka.** 1994. DNA rearrangement of the shufflon determines recipient specificity in liquid mating of IncI1 plasmid R64. J. Mol. Biol. **243:**6–9.
21. **Kroos, L., A. Kuspa, and D. Kaiser.** 1986. A global analysis of developmentally regulated genes in *Myxococcus xanthus*. Dev. Biol. **117:**252–266.
22. **Martin, S., E. Sodergren, T. Masuda, and D. Kaiser.** 1978. Systematic isolation of transducing phages for *Myxococcus xanthus*. Virology **88:**44–53.
23. **Mizuuchi, K., R. Weisber, L. Enquist, M. Mizuuchi, M. Buracrynska, C. Fiedler, P.-L. Hsu, W. Ross, and A. Landy.** 1980. Structure and function of the phage λ *att* site: size, Int-binding sites, and location of the crossover point. Cold Spring Harbor Symp. Quant. Biol. **45:**429–437.
24. **Moitoso de Vargas, L., S. Kim, and A. Landy.** 1989. DNA looping generated by DNA bending protein IHF and two domains of lambda integrase. Science **244:**1457–1461.
25. **Orndorff, P., E. Stellwag, T. Starich, M. Dworkin, and J. Zissler.** 1983. Genetic and physical characterization of lysogeny by bacteriophage MX8 in *Myxococcus xanthus*. J. Bacteriol. **154:**772–779.
26. **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.
27. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
28. **Shimamoto, T., M. Inouye, and S. Inouye.** 1995. The formation of the 2′,5′-phosphodiester linkage in the cDNA priming reaction by bacterial reverse transcriptase in a cell-free system. J. Biol. Chem. **270:**581–588.
29. **Shimkets, L. J., and S. J. Asher.** 1988. Use of recombination techniques to examine the structure of *csg* locus of *Myxococcus xanthus*. Mol. Gen. Genet. **211:**63–71.
30. **Shimkets, L. J., R. E. Gill, and D. Kaiser.** 1983. Developmental cell interactions in *Myxococcus xanthus* and the *spoC* locus. Proc. Natl. Acad. Sci. USA **80:**1406–1410.
31. **Stellwag, E., J. M. Fink, and J. Zissler.** 1985. Physical characterization of the genome of the *Myxococcus xanthus* bacteriophage MX-8. Mol. Gen. Genet. **199:**123–132.
32. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
33. **Sun, J., M. Inouye, and S. Inouye.** 1991. Association of a retroelement with a P4-like prophage (retrophage φR73) integrated into the selenocystyl tRNA gene of *Escherichia coli*. J. Bacteriol. **173:**4171–4181.
34. **Takeshita, S., M. Sato, M. Toba, W. Masahashi, and T. Hashimoto-Goto.** 1987. High-copy-number and low-copy-number plasmid vectors for *lacZ* α-complementation and chloramphenicol- or kanamycin-resistance selection. Gene **62:**63–74.
35. **Thompson, J. F., and A. Landy.** 1989. Regulation of bacteriophage lambda site-specific recombination, p. 1–22. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
36. **Tojo, N., S. Inouye, and T. Komano.** 1993. Cloning and nucleotide sequence of the *Myxococcus xanthus lon* gene: indispensability of *lon* for vegetative growth. J. Bacteriol. **175:**2271–2277.
37. **Tojo, N., S. Inouye, and T. Komano.** 1993. The *lonD* gene is homologous to the *lon* gene encoding an ATP-dependent protease and is essential for the development of *Myxococcus xanthus*. J. Bacteriol. **175:**4545–4549.
38. **Vieira, J., and J. Messing.** 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153:**3–11.
39. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
40. **Zieg, J., and M. Simon.** 1980. Analysis of the nucleotide sequence of an invertible controlling element. Proc. Natl. Acad. Sci. USA **77:**4196–4200.