Bacterial Insertion Sequences

E. Ohtsubo and Y. Sekine

1 Introduction

While DNA has a property of being fundamentally stable as invariable genetic information, studies on gene expression and gene organization have revealed that the genome is often subject to dynamic changes. Some of these changes are brought about by mobile genetic elements which have been found in prokaryotic and eukaryotic genomes so far studied. Insertion sequences (ISs) are bacterial mobile DNA elements which cause various kinds of genome rearrangements, such as deletions, inversions, duplications, and replicon fusions, by their ability to transpose. These were discovered during investigation of mutations that are highly polar in the galactose and lactose operons of *Escherichia coli* K-12 (Jordan et al. 1968; Malamy 1966, 1970; Shapiro 1969) and in the early genes of bacteriophage λ (Brachet et al. 1970). Many of these mutations were shown by

Institute of Molecular and Cellular Biosciences, The University of Tokyo, Bunkyo-ku, Yayoi 1-1-1, Tokyo 113, Japan

H. Saedler et al. (eds.), *Transposable Elements*

electron microscope heteroduplex analysis to be insertions of distinct segments of DNA which are hence called insertion sequences (FIANDT et al. 1972; HIRSCH et al. 1972; MALAMY et al. 1972). It was later shown that the transcription of flanking genes can originate from promoters located within an IS or from hybrid promoters created by the insertion event or by the IS-mediated genome rearrangements. An important note here is that the finding of IS elements as mobile elements to new loci to turn genes either off or on would re-evaluate the controlling elements described in maize by MCCLINTOCK (1956, 1965) (see STARLINGER and SAEDLER 1976).

In addition to the IS elements found in early investigations (IS1, IS2, IS3, and IS4), numerous others have been identified in the genomes, plasmids, and bacteriophages of a wide range of bacterial genera and species. They range in length from 800 to 2500 base pairs (bp) and can be found in the genomes of many different bacteria at multiplicities between a few and a few hundred per genome. They are often associated with genes responsible for resistance to antibiotics, heavy metal ions, etc. as components of transposons that frequently appear on natural plasmids.

IS elements contain one or more open reading frames encoding an enzyme, transposase, which is required for transposition. The termini of the majority of known IS elements carry inverted repeats (IRs) of about 10–40 bp, which are also required for transposition. These terminal repeats serve as recognition sites for transposase during the transposition process. IS elements can move to new sites by mechanisms largely independent of the homology-dependent recombination pathway. Upon insertion, these elements generate small, directly repeated duplications of the target DNA at the insertion point. This is presumably due to the staggered cleavage of target DNA by transposases. Many elements appear to induce a duplication of a fixed number, ranging from 2 to 13 bp, but some show variations in target duplication length. In a variety of DNA rearrangements, a target sequence at the illegitimate recombination junction is also duplicated.

In this chapter, we describe various IS elements and their characteristics, focusing on two elements, IS1 and IS3, which we have been studying for the past few years. We will present in the first section the finding that the majority of IS elements found in gram-negative and gram-positive bacteria actually belong to several families, which are classified based on structures, mechanisms of transposition and gene expression, and homologous genes encoding transposases. It is remarkable that the range of hosts for some families is found to be extremely broad, even including fungi, plants, invertebrate and vertebrate metazoa, and ciliated protozoa. It is particularly interesting that the most conserved transposase domain of several IS families is shared by retrovirus and retrotransposon integrases. Moreover, both prokaryotic and eukaryotic members of some families require translational frameshifting to produce the active transposase or integrase which apparently promotes the cleavage and integration of the elements. In the second and final sections, therefore, we will present a mechanism of the expression of transposases of several elements by translational frameshifting

and the generation of possible transposition intermediates by the action of transposases, to show that some IS elements use mechanisms in both gene expression and transposition similar to those used by retroelements. For the most comprehensive recent review, see GALAS and CHANDLER (1989), who have described the current understanding of the nature, occurrence, and genetic activities of various IS elements in detail.

2 Families of Insertion Sequences

Until recently, IS elements were considered to be a heterogeneous class of bacterial mobile DNA elements, and affiliations were restricted to highly similar ISs originating from related hosts. With the increasing number of known ISs, more distant relationships were pointed out, leading to the emergence of families of elements with conserved transposase domains, common structural features of similar functional properties. The majority of the IS elements so far identified may belong to the following families (see Table 1).

2.1 IS1 Family

IS1 (FIANDT et al. 1972; HIRSCH et al. 1972; MALAMY et al. 1972) is present in various copy numbers in chromosomes as well as in plasmids of bacteria belonging to Enterobacteriaceae (NYMAN et al. 1981; RAMIREZ et al. 1992; LAWRENCE et al. 1992; BISERCIC and OCHMAN 1993; MATSUTANI and OHTSUBO 1993). IS1 is the element that appears as spontaneous insertion mutations in various genes much more frequently than other ISs (for recent examples, see TOBA and HASHIMOTO 1992; RODRIGUEZ et al. 1992; OU et al. 1992; SKALITER et al. 1992). IS1 from resistance plasmid R100 is 768 bp, the smallest known, and has imperfect IRs (IRL and IRR) of about 30 bp at its termini (see Fig. 1) (OHTSUBO and OHTSUBO 1978). The majority of sequenced IS1 insertions generate target duplication of 9 bp, as previously identified (GRINDLEY 1978; CALOS et al. 1978). However, duplications of 7, 8, 10, 11, and 14 bp have been also observed (see GALAS and CHANDLER 1989 for references). IS1 is involved in various kinds of genomic rearrangements including co-integration between two replicons to form a characteristic co-integrate with two copies of IS1 (see Fig. 2) (IIDA and ARBER 1980; E. OHTSUBO et al. 1980). The co-integration event generates target duplication at the recombination junctions (Fig. 2) (E. OHTSUBO et al. 1980). Note that many IS elements including IS10R which transpose in a nonreplicative manner (see N. KLECKNER et al., this volume) do not generate the co-integrate.

IS1 contains two open reading frames, called insA and insB, which are essential for transposition (E. OHTSUBO et al. 1981; Y. MACHIDA et al. 1982, 1984; JAKOWEC et al. 1988). insA and insB code for putative proteins of 91 and 125 amino

Table 1. A list of insertion sequences and related elements identified in each family

Family	ISs and related elements
IS1	IS1 (R100) or IS1R, IS1A~IS1G, IS1(SD), IS1(νξ), IS1(SS), IS1(SF), IS1SFO, IS1s1~3, IS1Efe, IS1Ehe, IS1Sfl, IS1Sso, IS1Sdy
IS3	IS2, IS3, IS21, IS26, IS51, IS120, IS136, IS150, IS240, IS426, IS476, IS481, IS600, IS629, IS861, IS904, IS911, IS986, IS3411, IS6110, ISR1, ISL1, ISS1W, [IS232, IS981, IS987, IS1076L, IS1076R, IS1138, IS1533] (HTLV1, HTLVII, HIV1, Human-D, HumERVKA, Visna, FIV, EIAV, BLV, RSV, Mouse IAP, MMTV, MoMuLV, BaEV, SNV, HSpuENV, *Drosophila* 412, *Drosophila* 17.6, *Drosophila* 1731, Gypsy, Copia, *Bombyx* Mag, Tobacco Tnt, *Arabidopsis* Ta1-3, *Trichplusia ni* Ted, Ty1-17, Ty3-1, Ty3-2, S. *pombe* Tf1)
IS4	Group A: IS4, IS10R, IS50R, IS186A, IS186B, IS231A~IS231F, IS231V, IS231W, IS421, IS701, IS942, IS1151, IS5377, ISH26, ISH27-1~ISH27-3, ISH51-3, [IS231G, IS231H]
	Group B: IS5, IS102, IS112, IS402, IS427, IS493, IS702, IS869, IS903, IS1031A, IS1031C, IS1031D, IS1096, IS1106, ISH1, ISH11, ISH28, ISRm4, ISTUB, ISVM5-2, Tn4811 [IS6501, IS31831]
IS630-Tc1	IS630, IS895, IS1066, [IS870]
	(Tc1, IpTc1, CbTc2, Tc3, CbTc, Tec1, Tec2, TBE1, *Mariner*, CpMar, Tes1, Bari, Minos, RSa, RiATL, DmUhu, DmHB1, DmHB2)
Unusual ISs	IS110, IS116, IS117, IS492, IS900, IS901, IS902, IS1000, IS1111a, IS1533
IS1071 and Tn3	IS1071
	Tn3 family transposons: γδ, Tn3, Tn21, Tn501, Tn917, Tn2501, Tn3926, Tn4430, Tn4556
IS30	IS30, IS1086, IS4351, [ISAS2]
IS15 (or IS6)	IS15Δ, IS15(P-21), IS15R, IS26, IS46, IS176, IS240A, IS240B, IS257L, IS257R1, IS257R2, IS431L, IS431R, IS946, IS6100, ISS1N, ISS1S, ISS1T, ISS1W, [IS904]
IS91	IS91, IS801
IS256	[IS256, IS406, IS905, IS1081, IS6120, IST2, ISRm3]
not classified	[IS53, IS66, IS200, IS298, IS407, IS466, IS511, IS866, IS891, IS892, IS986, IS1016, IS1131, IS1133, IS1136, ISC1217, ISAE1, ISAS1, IS-PA-4, IS-PA-5]

ISs and related elements in eukaryotes (in parentheses) are listed whose transposase/integrase sequence relations have been studied. For IS1 family members, see OHTSUBO et al. (1984), UMEDA and OHTSUBO (1991), MILLS et al. (1992) and LAWRENCE et al. (1992); for IS3 family members, see FAYET et al. (1990), KHAN et al. (1991), KULKOSKY et al. (1992) and REZSÖHAZY et al. (1993a); for IS4 family members, see REZSÖHAZY et al. (1993a); for IS630-Tc family members, see HENIKOFF (1992) and DOAK et al. (1994); for unusual IS members, see LENICH and GLASGOW (1994); for IS1071 and Tn3 family members, see NAKATSU et al. (1991) and MAEKAWA and OHTSUBO (1994); for IS30 family members, see DONG et al. (1992) REZSÖHAZY et al. (1993a); for IS6 (IS15) family members, see KATO et al. (1994) and REZSÖHAZY et al. (1993a). Note that IS4 family elements are separated into two groups, A and B (see text). The other IS elements, including those which may or may not show homology with the family members, are shown in square brackets without references.

acids (aa), respectively, and the two frames are 47 bp apart (see Fig. 1). There is an open reading frame of 126 bp, named B' frame, in the region extending from the 5'-end of insB. This B' frame overlaps insA in the –1 reading frame (Fig. 1). As will be described in the next section in detail, IS1 uses translational frameshifting to produce a transframe protein from two out-of-phase reading frames, insA and B'-insB (SEKINE and OHTSUBO 1989; LUTHI et al. 1990; ESCOUBAS et al. 1991). The frameshifting event in the –1 direction occurs at a run of six adenines present in the

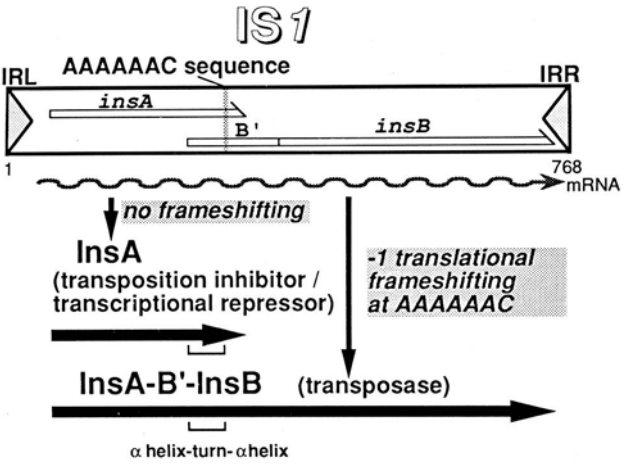

Fig. 1. Structure and gene expression of IS1

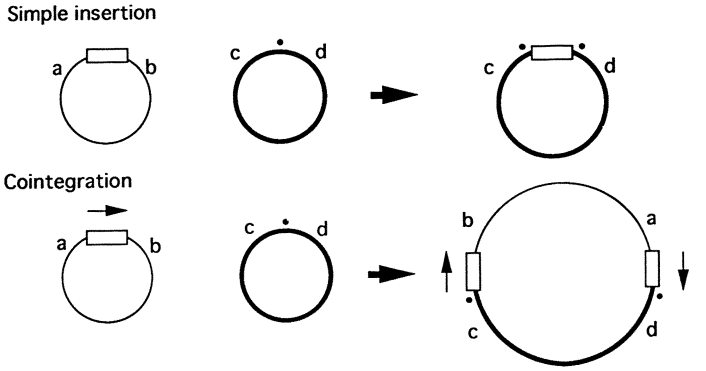

Fig. 2. Simple insertion and cointegration mediated by an IS element. Donor and target plasmids are indicated by *thin* and *thick lines*, respectively. *Open boxes* represent IS elements. Note that the co-integrate formed contains two copies of IS elements, whose orientations are shown by *thin arrows*. A *dot* on the target plasmid represents the target site which is duplicated upon simple insertion or co-integration. ***a, b, c,*** and ***d*** indicate DNA sequences on donor and target plasmid

overlapping region between insA and B' and produces the InsA-B'-InsB fusion protein, i.e., IS1 transposase (SEKINE et al. 1992). Unless frameshifting occurs, only the InsA protein is produced (see Fig. 1).

E. coli K-12 strains contain various number of copies of IS1 in their chromosomes. For W3110, the reference strain of the Kohara phage library (KOHARA et al. 1987), seven copies of IS1 were identified and six were mapped at 0.4, 6.3, 6.5, 22.3, 75.6, and 87.5 min (UMEDA and OHTSUBO 1989) and a possible seventh at 7 min (VAN HOVE et al. 1990) (IS1A-IS1G in Table 1). Recently, the eighth copy of IS1 identified was mapped at 49.6 min (ZUBER and SCHUMANN 1993). BIRKENBIHL and

Vielmetter (1989) also found seven copies of IS1 and located six copies on the physical map (although not to Kohara phages). Their nucleotide sequences revealed that there are actually four kinds with a sequence difference of 1~10% by base substitutions as compared with that of IS1 in plasmid R100 (Umeda and Ohtsubo 1991). Only two kinds of sequences have been found to be identical to the insertions in the *E. coli* K-12 genes (Umeda and Ohtsubo 1991), such as the *lacI* gene (Johnsrud 1979), the *lacZ* gene (Malamy et al. 1985), and the *unc* gene cluster (Kanazawa et al. 1984). The IS1-mediated rearrangements by deletion, tandem duplication, transposition, and amplification of a chromosomal DNA segment, are shown to have occurred in the *E. coli* chromosome (see Umeda and Ohtsubo 1989, 1991).

IS1 elements have been isolated also from the chromosomes of natural strains of *E. coli*, *Shigella* strains, *Salmonella typhimurium,* and other related enteric bacteria (see Table 1) (Ohtsubo et al. 1984; Mills et al. 1992; Lawrence et al. 1992; Bisercic and Ochman 1993). Their nucleotide sequences were very similar to those identified in *E. coli* K-12, except for one in *S. dysenteriae,* called IS1 (νξ) (H. Ohtsubo et al. 1981), which shows sequence difference of about 45% as compared with that of IS1 in R100. The similarity of the IS1 sequences suggests that they have been transferred horizontally among Enterobacteriaceae (H. Ohtsubo et al. 1981; Nyman et al. 1981; Ramirez et al. 1992; Lawrence et al. 1992; Bisercic and Ochman 1993; Matsutani and Ohtsubo 1993). Almost all the IS1 family elements contain open reading frames corresponding to insA and B'-insB, suggesting that they produce their transposases by translational frameshifting (Sekine and Ohtsubo 1989; Lawrence et al. 1992). An IS1-like element of only 116 bp, with IRs and some internal regions of IS1, was found to be widespread in Enterobacteriaceae (Goussad et al. 1991).

2.2 IS3 Family

IS3 (Fiandt et al. 1972; Malamy et al. 1972) is an insertion element present in the *E. coli* K-12 chromosome (Deonier et al. 1979; Birkenbihl and Vielmetter 1989; Umeda and Ohtsubo 1989) and in plasmid F (Hu et al. 1975; Yoshioka et al. 1990). IS3 is also present in various copy numbers in chromosomes of bacteria belonging to Enterobacteriaceae (Matsutani et al. 1987; Lawrence et al. 1992; see Galas and Chandler 1989). This element (1258 bp in length) has imperfect terminal IRs (IRL and IRR) of 39 bp (see Fig. 3) and generates target duplication of 3 bp upon insertion (Sommer et al. 1979; Timmerman and Tu 1985; Yoshioka et al. 1987; Spielmann-Ryser et al. 1991; Lawrence et al. 1992). Unlike IS1, IS3 does not mediate co-integration and is thus supposed to transpose in a nonreplicative manner (Spielmann-Ryser et al. 1991; Sekine et al. 1994). IS3 codes for two open reading frames, orfA and orfB. A reading frame (B') extending upstream from the initiation codon ATG of orfB overlaps orfA in the –1 frame (Fig. 3). Like IS1, a transframe protein of 317 aa, i.e., IS3 transposase, is produced by –1 translational frameshifting at an AAAAG sequence present in the overlapping region between

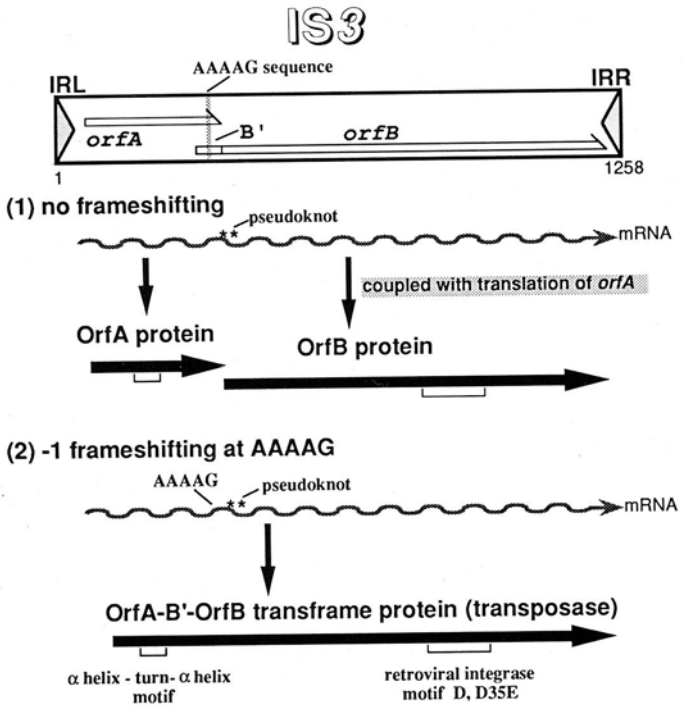

Fig. 3. Structure and gene expression of IS3

the two orfs (Fig. 3; see next section). Unlike IS1, however, both OrfA (99 aa) and OrfB (288 aa) proteins are produced from orfA and orfB, respectively, unless frameshifting occurs (Fig. 3; SEKINE et al. 1994).

There exists a group of IS elements, called the IS3 family (SCHWARTZ et al. 1988), which are structurally related to IS3. This is a large group (FAYET et al. 1990) including IS2, IS150, and IS911 (see Table 1). They are present not only in gram-negative bacteria but also in gram-positive bacteria (see GALAS and CHANDLER 1989). Several copies of IS2 and IS3 (BIRKENBIHL and VIELMETTER 1989; UMEDA and OHTSUBO 1989) and one copy of IS150 (BIRKENBIHL and VIELMETTER 1989) have been mapped on the *E. coli* K-12 chromosomes. IS2 and IS3 are importantly involved in both integration of plasmid F to form Hfr strains and excision from the Hfr strains to form F-prime factors (UMEDA and OHTSUBO 1989). A particular pair of IS3 is involved in inversion of a chromosomal segment (KOMADA et al. 1991; AJDIC et al. 1991). Almost all members of this group have two open reading frames corresponding to those in IS3. Although the predicted amino acid sequences of orfAs show little similarity, those of orfBs are significantly related (SCHWARTZ et al. 1988; FAYET et al. 1990; PRÈRE et al. 1990). Transposases of IS911 and IS150 have been shown to be produced from the two open reading frames by translational frameshifting in the –1 direction (POLARD et al. 1991; VÖGELE et al. 1991), as will be described below in detail. The other IS3 family elements have frameshifting

signals similar to those of IS3, IS911, and IS150, suggesting that they also employ frameshifting during expression of their transposase genes (SEKINE and OHTSUBO 1991; LAWRENCE et al. 1992; CHANDLER and FAYET 1993).

Interestingly, the most conserved IS3 transposase domain is shared by retrovirus and retrotransposon integrases (FAYET et al. 1990; KHAN et al. 1991; KULKOSKY et al. 1992; REZSÖHAZY et al. 1993a). In retroviral/retrotransposon integrase proteins of several organisms (see Table 1), sequence comparison of their deduced aa sequences reveals strong conservation of a constellation of aa characterized by two invariant Asp (D) residues and a Glu (E) residue, which is referred to as the D, D(35)E region (KHAN et al. 1991; KULKOSKY et al. 1992). The same constellation is in the transposases of the IS3 family elements. The invariant acidic D and E residues in Rous sarcoma virus integrase (KULKOSKY et al. 1992) and human immunodeficiency virus integrase (KULKOSKY et al. 1992; VAN GENT et al. 1992; ENGELMAN and CRAIGIE 1992; DRELICH et al. 1992) are actually shown to be important in site-specific cleavage and integration of viral DNA. The invariant D and E residues are proposed to participate in coordination of the metal cofactor (Mn^{2+} or Mg^{2+}) required for the catalytic activities of integrases. A metal-DNA complex may be necessary to position both LTR and target DNA substrates for nucleophilic attack during the cleavage and joining reactions. The conservation of this region suggests that the component residues are involved in DNA recognition, cutting, and joining, which are properties shared among integrases of divergent origin. It is interesting that a comparison of terminal IRs of IS3 family elements shows a high frequency at the tips of these elements of the sequence 5'-TG...CA-3'; the same sequence is found at the proviral DNA ends of retroviruses and is part of the *cis*-acting region required for integration (FAYET et al. 1990; KHAN et al. 1991). It should be noted here that transposons such as Tn7, Tn552, Tn5090, and Mu, with the ends beginning by 5'-TG, also produce transposases with the D, D(35)E region (RÅDSTRÖM et al. 1994; see also N. Craig, this volume).

2.3 IS4 Family

IS4 (FIANDT et al. 1972), which is 1426 bp with IRs of 18 bp, generates direct repeats of the target sequence of 11~13 bp upon insertion (KLAER et al. 1981). Unlike IS1 and IS3, IS4 displays one long open reading frame encoding a putative transposase of 442 aa and, in fact, produces it (TRINKS et al. 1981). A unique IS4 copy has been mapped at about 97 min in the *E. coli* K-12 chromosome (KLAER and STARLINGER 1980). MAHILLON et al. (1985) have described that IS231 is similar to IS4 in overall structure and in fact shares homology in transposase of IS4. They further noticed that IS231 is similar to not only IS4, but also other IS elements such as IS10R, IS50R, IS5, IS903, and ISH1, thus defining a new family of IS elements (see Table 1). IS231 itself defines a family of eight ISs, 1.7–2.0 kb in length, originating from the gram-positive entomopathogen *Bacillus thuringiensis* (MAHILLON et al. 1985, 1987; REZSÖHAZY et al. 1992, 1993a). These elements are delimited by 20-bp IRs. Six of them (IS231A~IS231F) display one long open

reading frame encoding a 477/478-aa transposase. The other two (IS231V and IS231W) show two slightly overlapping open reading frames (ORFA and ORFB) on the same DNA strand. It was speculated that in these two elements +1 (or +2) translational frameshifting could lead to the synthesis of a single 472 aa transposase (REZSÖHAZY et al. 1993b). Altogether, the eight IS231 transposases share 40% sequence identity, with five regions displaying more than 60% identity. One of these regions, designated C1, corresponds to a conserved 60 aa C-terminal box, which is shared by IS4, IS10R, IS50R, and ISH1 transposases (MAHILLON et al. 1985).

This IS4 family now comprises more than 40 ISs from widely different origins (REZSÖHAZY et al. 1993a). Members of this family all display the conserved C1 region. Moreover, they also share a second region (designated N3) highly conserved in the N-terminal half of the IS231 transposases (REZSÖHAZY et al. 1993a). The pattern of sequence similarities within these conserved regions and their relative location within the transposase define two groups within the IS4 family, the IS4 and IS5 groups (see groups A and B in Table 1). These relationships are reinforced by sequence conservation found within the IR sequences of these elements, in which the three external nucleotides, defining the limits of the elements, are rather group specific, 5'-CAT... for group A and either 5'-GGC... or 5'-GAG... for group B (REZSÖHAZY et al. 1993a). Interestingly, the N3 and C1 regions of the IS4 family may correspond to the integrase domain shared by retroelements and the IS3 family members (REZSÖHAZY et al. 1993a). This relationship might indicate a common step in the transposition mechanism of these otherwise unrelated mobile genetic elements. It is interesting to point out that the IS elements in group A include IS10R and IS50R, which transpose in a nonreplicative manner (see N. Kleckner et al., this volume), whereas those in group B include IS102 and IS903, which can form co-integrates (H. OHTSUBO et al. 1980b; GRINDLEY and JOYCE 1981). IS5 in group B is present in the *E. coli* K-12 chromosome in more copies than the others, and their locations have been determined (BIRKENBIHL and VIELMETTER 1989; UMEDA and OHTSUBO 1990a). They are involved in rearrangements of the chromosome by inversion, tandem duplication, etc. (UMEDA and OHTSUBO 1990a).

2.4 IS630-Tc1 Family

IS630 is a 1153-bp element with terminal IRs of 28 bp and has an open reading frame encoding a putative 343-aa protein (MATSUTANI et al. 1987; TENZEN et al. 1990). IS630, which was identified in *Shigella sonnei,* has been shown to transpose in *E. coli* K-12 specifically to the dinucleotide 5'-TA-3' in the core of at least 4-bp palindromic sequences, such as CTAG, TTAA, and ATAT (TENZEN et al. 1990), in which the CTAG sequences are used as hot spots for transposition (TENZEN and OHTSUBO 1991). IS630, like IS3, does not mediate co-integration, indicating that it transposes in a nonreplicative manner.

In prokaryotes, *Pseudomonas* IS1066 (VAN DER MEER et al. 1991), *Anabaena* IS895 (ALAM et al. 1991), and *Agrobacterium* IS870 (FURNIER et al. 1993) were identified to be related to IS630. In eukaryotes, however, many transposable elements are known to transpose into the dinucleotide TA. Examples are transposable elements, Tc1 and Tc3 of *Caenorhabditis elegans,* Tc2 and Tc6 of *C. elegans,* Tc1-like elements found in arthropods and vertebrates (see R.H.A. Plasterk, this volume), Tec1 and Tec2 of the ciliate *Euplotes crassus* (JARACZEWSKI and JAHN 1993), and *pogo* of *Drosophila* (TUDOR et al. 1992). HENIKOFF (1992) used blocks of aligned protein segments derived from the Tc1 family members to search a nucleotide sequence databank and detected the relatives of Tc1, IS630, and Tc1-like elements in arthropods and vertebrates. DOAK et al. (1994) have reported that the transposable elements TBE1, Tec1, and Tec2 of hypotrichous ciliated protozoa appear to encode a protein that belongs to the IS630-Tc1 family of transposases (see Table 1). DOAK et al. (1994) noted that most family members transpose into the dinucleotide target, TA, and that members with eukaryotic hosts have a tendency for somatic excision that is carried to an extreme by the ciliate elements. Alignments including the additional members, and also *mariner* elements, show that transposases of this family share strongly conserved residues in a large C-terminal portion, including a fully conserved dipeptide, DE, and a block consisting of a fully conserved D residue and highly conserved E residue, separated by 35 (or 34) residues (D35E). This D35E motif likely is homologous to the D35E motif of the family of retroviral-retrotransposon integrases and IS3-like transposases. The homologous relations suggest that the two families share homologous catalytic transposase domains and that members of both families may share a common transposition mechanism.

Recently, an element named Tnr1 (235 bp in length) with terminal inverted repeats of 75 bp has been identified in rice (UMEDA et al. 1991; TENZEN et al. 1994). Because of its small size, Tnr1 is supposed to be a defective form of an autonomous element capable of transposing by itself. Comparison of nucleotide sequences of the regions with or without a Tnr1 member revealed that Tnr1 transposes to 5'-PuTAPy-3' duplicating TA (TENZEN et al. 1994), suggesting that Tnr1 may be a member of the IS630-Tc1 family. A family of elements named *stowaway,* homologous to Tnr1, are associated with the genes of both monocotyledonous and dicotyledonous plants (BUREAU and WESSLER 1994). The finding of these elements suggests that the range of hosts for the aggregate IS630-Tc1 family is extremely broad, including bacteria, fungi, plants, invertebrate and vertebrate metazoa, and ciliated protozoa.

2.5 A Family of Unusual IS Elements

LENICH and GLASGOW (1994) reported that the predicted amino acid sequence of Piv, an essential protein in site-specific DNA inversion of the pilin segment in *Moraxella lacunata,* shows significant homology with the transposases of a family of IS elements (see Table 1) which were previously proposed to be a group

(Hoover et al. 1992; Kunze et al. 1991; Leskiw et al. 1990; Moss et al. 1992; Stark et al. 1992). These proteins contain four regions where highly conserved aa are clustered, representing aa motifs or domains that are important for transposition. Many of the IS elements in this family are unusual in structure and behavior, and can be subgrouped on the basis of the absence or presence of terminal IRs and of whether insertion results in target site duplications. One subgroup is composed of IS elements found in the gram-positive species *Streptomyces clavuligerus* (IS116), *Mycobacterium paratuberculosis* (IS900), *M. avium* (IS901), and *M. avium* subsp. *silvaticum* (IS902) (Kunze et al. 1991; Leskiw et al. 1990; Moss et al. 1992; Green et al. 1989). These IS elements do not have terminal IRs and do not generate a short repeat at the site of insertion: IS110 and IS117 from *Streptomyces coelicolor* may be part of this subgroup though they have imperfect terminal IRs, and it is uncertain whether IS110 generates target site duplications (Bruton and Chater 1987; Henderson et al. 1989). IS1000 from *Thermus thermophilis* and IS111 a from *Coxiella burnetii* also give no sequence duplications at the site of insertion, but they do have terminal IRs (Ashby and Berquist 1990; Hoover et al. 1992). IS492 from *Pseudomonas atlantica* generates a 5-bp repeat at its target sites but has no terminal IRs (Bartlett and Silverman 1989). IS1533 has imperfect terminal IRs and may generate a 2-bp duplication upon insertion (Zuerner 1994). No common features in the ends of the IS elements or in their target sites have been noted for this family of transposable elements.

2.6 IS1071, a Member of Tn3 Family

Nakatsu et al. (1991) described the structure of a 17-kb transposon Tn5271, which resides in the plasmids or chromosome of an *Alcaligenes sp.* strain. The transposon is flanked by a directly repeated sequence of 3201 bp, designated IS1071. The IRs of IS1071 and the derived aa sequence of the single open reading frame within IS1071 are related to the IRs and transposase proteins of the Tn3 family transposons (see Table 1) (Nakatsu et al. 1991; Maekawa and Ohtsubo 1994). Tn3 family transposons generally code for two genes (tnpA and tnpR) that are necessary for its transposition: tnpA encodes the transposase, which catalyzes the first step of transposition that is the formation of a co-integrate; tnpR encodes the resolvase/repressor, which acts at the res sites to resolve the co-integrate molecule with two copies of Tn3 into two products, each with a single copy of Tn3 (for a recent review, see Sherratt 1989). The absence of tnpR within IS1071 suggests that this element is capable of determining the first step only. This was confirmed by observations on the IS1071-dependent formation of stable co-integrates in a recombination-deficient *E. coli* K-12 strain (Nakatsu et al. 1991). The existence of IS1071 may support an evolutionary scheme in which the Tn3 family transposons descended from simpler insertion sequences.

2.7 Others

Dong et al. (1992) reported that there is a family of IS elements related to IS30 encoding a single open reading frame for a protein (see Table 1). Locations of IS30 on the *E. coli* K-12 chromosome have been determined (Birkenbihl and Vielmetter 1989; Umeda and Ohtsubo 1990b). Rezsöhazy et al. (1993a) noted that members of this family display the conserved regions, which correspond to the integrase domain shared by retroelements and the IS3 family members. Interestingly, the transposase domain found in these members has both the length feature of amino acid spacer characteristic of the IS3 family and the amino acid conservation specific of the C1 domain of the IS4-related elements, indicating a common step in the transposition mechanism of the IS elements belonging to the IS3, IS4, and IS30 families. There is another family, called the IS15 (or IS6) family (see Table 1) (Martin et al. 1990; Kato et al. 1994). Some members of this family have been identified as those of the IS3 family (see Table 1). Rezsöhazy et al. (1993a) noted that five members (IS15Δ, IS240A, IS431L, IS946, and ISS1S), however, display the conserved regions corresponding to the integrase domain but are distinct from the IS3 family elements. These suggest that the IS15 (IS6) family members may be divided into at least two groups.

There exist IS elements which seem not to belong to any of the families described above. The most interesting of them is *E. coli* IS91, which specifically inserts at CAAG or GAAC of target and does not duplicate any sequence upon insertion (Mendiola and de la Cruz 1989). The related IS element to IS91 is *Pseudomonas* IS801 (Romantschuk et al. 1991; Mendiola et al. 1992). IS91/IS801 transposase is interestingly related to the rolling-circle-type replication proteins of the pUB110 family of plasmids which produce a single-strand nick in a specific site, suggesting that transposition of IS91/IS801 involves single-strand nicking by the transposases (Mendiola and de la Cruz 1992). The other elements, which may or may not show homology with the IS elements described above, are listed without references in Table 1. These include IS256 and its relatives, forming a family (see Table 1). It is not our aim to describe characteristics of all of the unclassified elements in detail here. With the increasing number of new ISs, distant relationships will be further pointed out, leading to more or fewer families of elements.

3 Translational Frameshifting in Production of Transposases Encoded by IS Elements

IS1 and IS3 family elements use translational frameshifting in the –1 direction to produce their transposases. In this section we will describe this event in some detail and compare it with the frameshifting event well known in retroelements as producing a transframe polyprotein from which integrase is derived by processing.

3.1 IS1 Transposase

As described in the previous section, translational frameshifting occurs in the –1 direction from the 3'-end region of insA to B'-insB in IS1 to produce the InsA-B'-InsB transframe protein, i.e., transposase (see Fig. 1) (SEKINE and OHTSUBO 1989; ESCOUBAS et al. 1991). This finding could explain many things that cannot be readily accounted for in the expression of insA and the (B')-insB frame: The protein product from insB or B'-insB was not detected, but the insA product was (ARMSTRONG et al. 1986; ZERBIB et al. 1987; ESCOUBAS et al. 1991); two IS1 mutants defective in insA and insB, respectively, do not complement each other in restoring their co-integration ability; a 5-bp insertion in the B' frame results in loss of the ability of IS1 to mediate co-integration (MACHIDA et al. 1982).

The frameshifting in the –1 direction occurs within the run of six adenines in the sequence 5'-TTAAAAAACTC-3' at nucleotide (nt) position 305~315 in the overlapping region (Fig. 1) and produces transposase with a polypeptide segment Leu-Lys-Lys-Leu at residues 84~87 (SEKINE and OHTSUBO 1989; SEKINE et al. 1992). An IS1 mutant with a single base insertion in the run of adenines which produces the transframe protein with the segment Leu-Lys-Lys(or Arg)-Leu without frameshifting could efficiently mediate co-integration and adjacent deletion (SEKINE and OHTSUBO 1989; SEKINE et al. 1992). Substitution mutations at each of three (2nd, 3rd and 4th) adenine residues, which comprise a codon for Lys in insA, caused serious negative effects in frameshifting, but those introduced in the region flanking the run of adenines did not. These indicate that the AAA codon for Lys is the site of frameshifting and that $tRNA^{Lys}$ thus plays an important role in frameshifting (SEKINE and OHTSUBO 1992).

In many genetic systems to be described below, secondary structures downstream of the frameshift site are supposed to elevate opportunities for a change of reading frames. Several possible secondary structures in the region downstream of the run of adenines are, however, not required for frameshifting in IS1, but the termination codon of insA located at 17 bp downstream of the run of adenines plays an important role in enhancement of frameshifting (SEKINE et al. 1992; SEKINE and OHTSUBO 1992). Enhancement of –1 frameshifting by a termination codon immediately downstream of the frameshift site has been reported also in an artificial context (WEISS et al. 1987). The efficiency of frameshifting in IS1 is only 0.2~0.3%, however, which is very low compared with that of other genetic systems (see below). IS1 may adopt the low level of frameshifting, which results in production of a low amount of transposase, to avoid deleterious rearrangement of the host chromosome containing IS1 (SEKINE and OHTSUBO 1989; ESCOUBAS et al. 1991; SEKINE et al. 1992).

The InsA protein has the carboxyl-terminal region containing an α-helix-turn-α-helix motif (see Fig. 1) (ZERBIB et al. 1987; SEKINE and OHTSUBO 1991), which is observed in many DNA binding proteins (PABO and SAUER 1984). It has in fact been shown that InsA specifically binds to both IRL and IRR (ZERBIB et al. 1987, 1990b; SEKINO et al. 1995). IRL contains a promoter for the transcript coding for InsA and transposase, whereas IRR contains a promoter used for synthesizing RNA being

oriented opposite to insA and insB (Chan and Lebowitz 1982; C. Machida et al. 1984). Since the InsA-binding regions overlap the –35 region of the promoters, InsA could inhibit the transcription by interacting with RNA polymerase at IRL and IRR. Inhibition of transcription from IRL by InsA has actually been shown in vivo (Machida and Machida 1989; Zerbib et al. 1990a; Escoubas et al. 1991). The transposase (InsA-B'-InsB) protein, which includes the same DNA-binding motif, binds to the DNA segment with or without the IR sequence but preferentially to that with IR (Sekino et al. 1995). The nonspecific DNA-binding ability of transposase may be involved in recognition of the target DNA, an important process of transposition of IS1. It is speculated that IS1 transposase consists of at least two domains, the N-terminal half, which almost entirely overlaps InsA, and the C-terminal half, which almost entirely overlaps B'-InsB. The frameshifting event adds the latter domain to the former to give the transposase activity recognizing IRs and the target sequence to initiate the transposition reaction.

3.2 IS3 Transposase

As described also in the previous section, IS3 produces its transposase by –1 frameshifting at an AAAAG sequence present in the overlapping region between orfA and B'-orfB (see Fig. 3). Amino acid sequencing analysis has shown that the sequence 5'-CAAAAGGC-3' at nt position 327–334 in the overlapping region encodes the transframe polypeptide segment Gln-Lys-Gly, suggesting that frameshifting occurs at either codon CAA or codon AAG (Sekine et al. 1994). Mutational analysis indicates that AAG is the site of frameshifting (Sekine et al. 1994). This suggests that $tRNA^{Lys}$ recognizing this codon plays an important role in –1 frameshifting. This frameshifting requires a pseudoknot structure in the region downstream of the AAAAG sequence (see Fig. 3) (Sekine et al.1994). A mutant IS3 with a single base insertion in the sequence, which results in in-frame alignment of orfA and B'-orfB, mediated adjacent deletion to produce various miniplasmids at a very high frequency (Sekine and Ohtsubo 1991).

The IS3 family members IS150 and IS911 also produce their transposases by –1 frameshifting at a motif AAAAAAG present in the overlapping region between two open reading frames. In IS150 (Vögele et al. 1991), the sequence 5'-CUAAAAAAGCU-3' at nt position 528–538 encodes Leu-Lys-Lys-Ala, indicating that a frameshifting occurs at either codon CUA for Leu or one of the consecutive codons, AAA and AAG, for Lys in the 0-frame. This frameshifting requires a single stem-loop structure in the region downstream of the frameshifting site (Vögele et al. 1991). In IS911 (Polard et al. 1991), the sequence 5'-TTAAAAAAGGC-3' at nt position 322–332 was preliminarily shown to encode Leu-Lys-Lys-Gly, indicating that a frameshifting occurs at either codon TTA for Leu or one of the consecutive codons, AAA and AAG, for Lys in the 0-frame.

Unlike IS1, however, IS3, IS150, and IS911 encode two proteins, OrfA and OrfB, in addition to the transposase protein, from two open reading frames unless frameshifting occurs. In IS3, OrfB is produced in a manner dependent on translation of orfA, whose termination codon overlaps the ATG codon of orfB. In

other words, OrfB is produced due to translational coupling between orfA and orfB (see Fig. 3). The pseudoknot structure required for frameshifting is important also for translational coupling in production of OrfB. Presumably, the initiation codon for orfB and its upstream sequence within the pseudoknot structure are occluded, unless ribosomes proceed toward the termination codon of orfA, allowing exposure of the region essential for translation of orfB (SEKINE et al. 1994). IS150 and IS911, however, produce OrfB using the mechanism different from that in IS3: In IS150, the OrfB protein, whose coding region begins from an ATG codon which is in phase with orfA and is located upstream of the AAAAAAG motif, is produced in a manner absolutely dependent on frameshifting at the motif (VÖGELE et al. 1991); in IS911, OrfB is produced by utilizing an unusual initiation codon AUU in phase with orfB located upstream of the AAAAAAG motif (POLARD et al. 1991).

Both OrfA and OrfB proteins are not required for transposition of IS3 (SEKINE et al. 1994). The OrfA protein contains an α-helix-turn-α-helix DNA-binding motif in the middle (PRÈRE et al. 1990; SEKINE and OHTSUBO 1991). This motif is also present in the transposase protein (see Fig. 3). It is therefore possible that OrfA competes with transposase to bind to the terminal IRs (IRL and IRR) and could thus become a transposition inhibitor. Since IRL contains a possible promoter sequence for transcription of the IS3-coded genes, OrfA could also inhibit transcription from this promoter, as has been shown in IS1. In fact, our recent results indicate that OrfA inhibits IS3-mediated deletion and that OrfB enhances the inhibitory effect by OrfA, while OrfB alone has no effect (Y. Sekine, K. Izumi, E. Ohtsubo, unpublished results). It has been reported that OrfA encoded by IS911 is, however, not a transposition inhibitor but stimulates intermolecular transposition (POLARD et al. 1992).

As described in the first section, the orfB frame of IS3 family members contains a conserved aa sequence motif found in retroviral integrase. The actual function of the OrfB proteins is not known at present. In the transframe proteins, the conserved region may function in catalysis of the transposition reaction and be directed to its site of action at the IS ends using the N-terminal portion encoded by orfA (SCHWARTZ et al. 1988; PRÈRE et al. 1990; SEKINE and OHTSUBO 1991). The occurrence of frameshifting suggests the two-domain structure in the transposase proteins.

3.3 Other Transframe Proteins Including Retroviral Polyproteins

In prokaryote, a chromosomal gene *dnaX* encoding the τ subunit of DNA polymerase III holoenzyme uses the –1 frameshifting to produce the transframe protein that is γ subunit (BLINKOWA and WALKER 1990; FLOWER and MCHENRY 1990; TSUCHIHASHI and KORNBERG 1990). The sequence 5'-GCAAAAAAGAG-3' in *dnaX* encodes Ala-Lys-Lys-Glu (TSUCHIHASHI and KORNBERG 1990), indicating that a frameshifting occurs at GCA for Ala or one of the consecutive codons, AAA and AAG, for Lys in the 0-frame. Mutational analysis of the region containing the AAAAAAG motif indicates that codon AAG is the site of frameshifting (TSUCHIHASHI and BROWN 1992). It is therefore likely that in IS150 and IS911, which use the same AAAAAAG motif for frameshifting, codon AAG is the site of frameshifting.

Translational frameshifting in the –1 direction has been reported in retroviruses (for a review, see VARMUS and BROWN 1989) and other eukaryotic viruses (BRIERLEY et al. 1989; DINMAN et al. 1991). These viruses have a conserved heptanucleotide motif at which frameshifting occurs as proposed in the simultaneous slippage model (JACKS et al. 1988). In mouse mammary tumor virus (MMTV), a –1 frameshifting occurs at the gag-pro overlapping region with the sequence AAAAAAC (JACKS et al. 1987; MOORE et al. 1987) identical to that of IS1 to produce its transposase, such that the last codon in gag (0 frame) is codon AAC (HIZI et al. 1987), which is downstream by one codon compared with the site in IS1. This difference between IS1 and MMTV [or other retroviruses which are suggested to use the AAAAAAC motif for frameshifting (RICE et al. 1985; SHIMOTOHNO et al. 1985)] might be due to the structural or functional differences between prokaryotic molecules participating in the translational process and those of eukaryotes.

The efficiency of frameshifting is 5~25% in retroviruses (for a review, see VARMUS and BROWN 1989), 40~50% in *dnaX* (FLOWER and MCHENRY 1990; TSUCHIHASHI and KORNBERG 1990), 30~40% in IS150 (VÖGELE et al. 1991), about 15% in IS911 (POLARD et al. 1991), and about 6% in IS3 (SEKINE et al. 1994). Pseudoknot structures have been shown to have an important role in frameshifting in IBV (BRIERLEY et al. 1989, 1991), MMTV (CHAMORRO et al.1992), the yeast double-stranded RNA virus L-A (DINMAN et al. 1991), and IS3 (SEKINE et al. 1994). A single stem-loop structure(s) downstream of the frameshift site has been demonstrated to stimulate frameshifting in *dnaX* (FLOWER and MCHENRY 1990; TSUCHIHASHI and KORNBERG 1990) and IS150 (VÖGELE et al. 1991). Such structures in mRNA are supposed to cause translating ribosomes to pause at the frameshift site, thereby providing tRNA on the ribosomes with elevated opportunities for a change of reading frames (JACKS et al. 1988; BRIERLEY et al. 1989).

As described here, frameshifting is the event that allows synthesis of several different proteins from a single template, and this leads to the merit of storing as much information as possible in a limited amount of DNA and RNA. It is also conceivable that frameshifting may be an exquisite strategy by which some fine regulations are carried out. Some IS elements, including IS231V and IS231W in the IS4 family, which have a set of smaller ORFs assumed to encode transposase may also employ frameshifting as translation trickery during expression of transposase genes.

4 Possible Intermediate Molecules of Transposition of IS Elements

The transposition reaction has not been reproduced in vitro in any IS elements except IS10R (see N. KLECKNER et al., this volume). Transposases from only a few IS elements have been isolated but not characterized for their enzymatic properties other than DNA binding to the terminal IRs. Therefore, the molecular

mechanism of transposition of the IS element is not yet well understood. Some IS elements or their related elements in eukaryotes, however, have been observed to generate characteristic circular and linear molecules, whose structures are similar to those identified as intermediates of transposition in the transposons, including those described in the chapters by N. KLECKNER et al. and N. CRAIG, this volume, and retroviruses. Here, we will describe IS elements which produce possible transposition intermediates and discuss the possibility that the IS3 family elements and retroelements may share a common transposition mechanism.

4.1 IS Circles

Circular forms of excised DNA have been found in several sequence-specific DNA rearrangement processes as either an intermediate or an end product. In transposons, circular products have been observed for Tn10 (IS10R) and conjugative transposon Tn916. IS10R generates two types of circular molecules upon the action of transposase (see N. Kleckner et al., this volume). One is a circle, a possible end product, formed by transposition of the element into itself. The second structure is a noncovalently closed structure that consists of a protein-DNA complex formed between the transposase and the ends of the element and can thus be converted to a linear form. Tn916 generates a covalently closed circular intermediate (CAPARON and SCOTT 1989). The junction of the transposon termini in the Tn916 circular intermediates is a heteroduplex that contains extra nucleotides derived from adjacent chromosomal sequences from each end of the integration site. This circular product is capable of reintegration (SCOTT et al. 1988).

POLARD et al. (1992) have reported that IS911, an IS3 family element, generates minicircles consisting of the entire sequence of IS911 and a 3-bp sequence intervening between the IRs. The 3-bp sequence is the same as the direct repeats of a target sequence flanking IS911. The minicircles are assumed to be formed by joining one end of IS911 with the overhanging 3-bp sequence on the other side of IS911. SEKINE et al. (1994) have reported that the IS3 mutant, which overproduces transposase without frameshifting, efficiently generates IS3 circles similar to the IS911 circles, in addition to miniplasmids formed by the IS3-mediated deletion from the parental plasmid. The IS3 circles have the intervening 3-bp sequence identical to either one of the sequences flanking IS3 in the parental plasmid or its miniplasmid derivatives (see Fig. 4) (SEKINE et al. 1994). Though it is not clear at present whether the IS3 circles participate as substrates in transposition or not, POLARD et al. (1992) have suggested that the IS911 circles are not the obligatory transposition intermediates. It has been reported, however, that two copies of IS elements separated by a few base pairs are active in transposition of IS21 (REIMMANN et al. 1989), IS3 (SPIELMANN-RYSER et al. 1991), and IS30 (OLASZ et al. 1993). It should be noted that the active junction composed of two IRs in the tandem repeats of IS elements resembles the IS3 (IS911) circle junction. As will

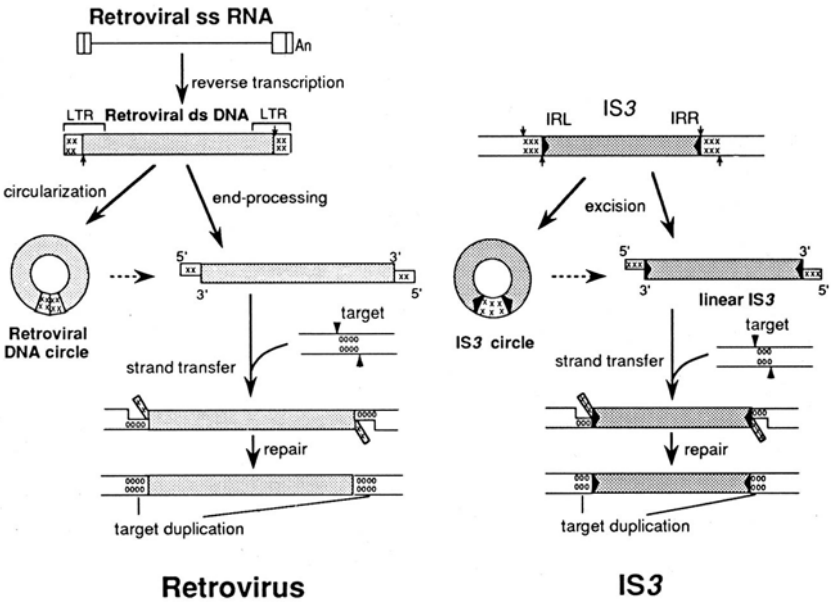

Fig. 4. Retrovirus integration and IS3 transposition

be described below, the IS3 transposase generates linear molecules with 3-nt overhangs at the 5'ends of IS3 which are possible intermediates of IS3 transposition. It is possible that the IS3 circles are converted to such linear molecules (see Fig. 4).

Similarly to IS911 and IS3, an IS1 mutant overproducing transposase generates characteristic circles consisting of the entire IS1 sequence and an intervening sequence, mostly 6–9 bp in length, between the IS1 ends (Y. Sekine, N. Eisaki, K. Kobayashi, E. Ohtsubo, unpublished results). The intervening sequences are derived from either one of the sequences flanking IS1 in the parental plasmid. Such IS1 circles may be formed during a process(es) of transposition or co-integration mediated by IS1.

Retroviruses are known to generate circular DNAs with two LTR sequences after reverse transcription from viral RNA (see Fig. 4). The circular molecules of retroviruses are considered not to be the transposition intermediates during integration of viral DNA (see Brown et al. 1987; Fujiwara and Mizuuchi 1988), but can be converted to linear molecules by integrase, which introduces site-specific staggered breaks at the junction between the two LTRs (see Fig. 4) (Grandgenett et al. 1986; Terry et al. 1988).

In the eukaryotic transposable elements related to IS630, circular forms of the *C. elegans* Tc1 element and *E. crassus* Tec elements have been detected (Rose and Snutch 1984; Ruan and Emmons 1984; Jaraczewski and Jahn 1993). In these circles, the inverted repeats of each element are joined in head-to-head orientation (see R.H.A. Plasterk, this volume). In Tec circles, the inverted repeat

junctions consist of both copies of the target site duplication surrounding ten additional bases. An unusual nuclease-sensitive conformation exists at a circular junction that may be the result of the heteroduplex DNA like that seen in Tn916 circles (JARACZEWSKI and JAHN 1993). To date, there is no direct evidence for reinsertion of the circular molecules, although the frequency of transposition correlates with the presence of circular forms.

4.2 IS Linears

In transposons Tn10 (IS10R) and Tn7, which transpose in a nonreplicative manner, double-strand breaks occur at both end regions of the elements to given an intermediate linear molecule, which is subsequently inserted into the target site (see N. Kleckner et al., this volume). IS3 does not mediate co-integration and is thus supposed to transpose in a nonreplicative manner (SEKINE et al. 1994). This leads us to propose that the IS3 transposase may excise the IS3 sequence from the donor molecules in the IS3 transposition reaction. We have, in fact, found linear molecules of IS3 with 3-nt overhangs at its 5'ends which were generated in addition to the IS3 circles from the plasmid carrying the IS3 mutant overproducing transposase (SEKINE et al. 1995). The nucleotide sequences of the overhangs are the same as those flanking IS3 in the parental plasmid, implying that the linear IS3 molecules are excised from the parental plasmid DNA or from IS3 circles by staggered double-strand breaks at the end regions of IS3 (Fig. 4).

IS3 generates both circular and linear molecules, while IS10R (Tn10) and Tn7 generate linear molecules but not circles. Retroviruses are, however, known to generate double-stranded linear DNA molecules in addition to circular DNAs (see Fig. 4). In this respect, IS3 resembles retroviruses. Note here that, as described in the earlier section, the most conserved IS3 transposase domain is shared by retrotransposon and retrovirus integrases. The conservation of this region suggests that the component residues are involved in DNA recognition, cutting, and joining, since these properties are shared among these proteins of divergent origin. In retroviruses, 2 nt from each 3' end of the linear viral DNA are removed by integrase to produce 5'-protruding ends (CRAIGIE et al. 1990; KATZ et al. 1990; KATZMAN et al. 1989; SHERMAN and FYFE 1990), and the 3' ends of the linear molecules are subsequently joined to the 5' ends generated at a target site (Fig. 4) (FUJIWARA and MIZUUCHI 1988; BROWN et al. 1989). Thus in IS3, the 3'-OH of the linear IS3 molecule is likely to be joined to 5'-P of the target DNA, which is supposed to be exposed by 3-bp staggered breaks, and subsequently the 3-nt gap on the opposite strand is repaired to convert the gap to a duplex form and to remove the 3-nt donor sequence attached to the 5' end of the linear IS3 molecule (see Fig. 4).

As described in this section, the molecular mechanism of transposition of the IS elements including IS3 is poorly understood compared with those of some transposons and retroviruses. Further investigation by reproducing in vitro the

transposition reaction of a representative IS element(s) in each family and by purification and characterization of its transposase is needed.

Acknowledgments. We would like to thank Dr. H. Ohtsubo for stimulating discussions, critical comments on the manuscript, and encouragement in preparing the manuscript.

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

References

Ajdic D, Jovanovic G, Glisin V, Hejna J, Savic DJ (1991) Nucleotide sequence analysis of the inversion termini located within IS3 element α3β3 and β5α5 of *Escherichia coli.* J Bacteriol 173: 906–909

Alam J, Vrba JM, Cai Y, Martin JA, Weislo LJ, Curtis SE (1991) Characterization of the IS895 family of insertion sequences from the Cyanobacterium *Anabaena* sp. strain PCC 7120. J Bacteriol 173: 5778–5783

Armstrong KA, Ohtsubo H, Bauer WR, Yoshioka Y, Miyazaki C, Maeda Y, Ohtsubo E (1986) Characterization of the gene products produced in minicells by pSM1, a derivative of R100. Mol Gen Genet 205:56–65

Ashby MK, Berquist PL (1990) Cloning and sequence of IS1000, a putative insertion from *Thermus thermophilus* HB8. Plasmid 24: 1–11

Bartlett DH, Silverman M (1989) Nucleotide sequence of IS492, a novel insertion sequence causing variation in extracellular polysaccharide production in the marine bacterium, *Pseudomonas atlantica.* J Bacteriol 171: 1763–1766

Birkenbihl RP, Vielmetter W (1989) Complete maps of IS1, IS2, IS3, IS4, IS5, IS30 and IS150 locations in *Escherichia coli* K12. Mol Gen Genet 220: 147–153

Bisercic M, Ochman H (1993) The ancestry of insertion sequences common to *Escherichia coli* and *Salmonella typhimurium.* J. Bacteriol 175: 7863–7868

Blinkowa AL, Walker JR (1990) Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ subunit from within the τ subunit reading frame. Nucleic Acids Res 18: 1725–1729

Brachet P, Eisen H, Rambach A (1970) Mutations of coliphage lambda affecting the expression of replicative functions O and P. Mol Gen Genet 108: 266–276

Brierley I, Digard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frame-shifting signal: requirement for an RNA pseudoknot. Cell 57: 537–547

Brierley I, Rolley NJ, Jenner AJ, Inglis SC (1991) Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. J Mol Biol 220: 889–902

Brown PO, Bowerman B, Varmus HE, Bishop JM (1987) Correct integration of retroviral DNA in vitro. Cell 49: 347–356

Brown PO, Bowerman B, Varmus HE, Bishop JM (1989) Retroviral integration: structure of the initial covalent product and its presursor, and a role for the viral IN protein. Proc Natl Acad Sci USA 86: 2525–2529

Bruton CJ, Chater KF (1987) Nucleotide sequence of IS110, an insertion sequence of *Streptomyces coelicolor* A3(2). Nucleic Acids Res 15: 7053–7065

Bureau TE, Wessler SR (1994) *Stowaway:* a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell 6: 907–916

Calos MP, Johnsrud L, Miller JH (1978) DNA sequence at the integration sites of the insertion element IS1. Cell 13: 411–418

Caparon MG, Scott JR (1989) Excision and insertion of the conjugative transposon Tn916 involves a novel mechanism. Cell 59: 1027–1934

Chamorro M, Parkin N, and Varmus HE (1992) An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. Proc Natl Acad Sci USA 89: 713–717

Chan PT, Lebowitz J (1982) Mapping RNA polymerase binding sites in R12 derived plasmids carrying the replication-incompatibility region and the insertion element IS1. Nucleic Acids Res 10: 7295–7311

Chandler M, Fayet O (1993) Translational frameshifting in the control of transposition in bacteria. Mol Microbiol 7: 497–503

Craigie R, Fujiwara T, Bushman F (1990) The IN protein of Moloney murine leukemia virus processes the viral DNA ends and accomplishes their integration in vitro. Cell 62: 829–837

Deonier RC, Hadley RG, Hu M (1979) Enumeration and identification of IS3 elements in *Escherichia coli* strains. J Bacteriol 137: 1421–1424

Dinman JD, Icho T, Wickner RB (1991) A –1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. Proc Natl Acad Sci USA 88: 174–178

Doak TG, Doerder FP, Jahn CL, Herrick G (1994) A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. Proc Natl Acad Sci USA 91: 942–946

Dong Q, Sadouk A, van der Lelie D, Taghavi S, Ferhat A, Nuyten JM, Borremans B, Mergeay M, Toussaint A (1992) Cloning and sequencing of IS1086, an *Alcaligenes eutrophus* insertion element related to IS30 and IS4351. J Bacteriol 174: 8133–8138

Drelich M, Wilhelm R, Mous J (1992) Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. Virology 188: 459–468

Engelman A, Craigie R (1992) Identification of conserved amino acid residues critical for human immunodeficiency virus type 1 integrase function in vitro. J Virol 66: 6361–6369

Escoubas JM, Prère MF, Fayet O, Salvignol I, Galas D, Zerbib D, Chandler M (1991) Translational control of transposition activity of the bacterial insertion sequence IS1. EMBO J 10: 705–712

Fayet O, Ramond P, Polard P, Prère MF, Chandler M (1990) Functional similarities between retroviruses and the IS3 family of bacterial insertion sequences? Mol Microbiol 4: 1771–1777

Fiandt M, Szybalski W, Malamy MH (1972) Polar mutations in lac, gal, and phage λ consist of a few IS-DNA sequences inserted with either orientation. Mol Gen Genet 119: 223–231

Flower AM, McHenry CS (1990) The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. Proc Natl Acad Sci USA 87: 3713–3717

Fournier P, Paulus F, Otten L (1993) IS870 requires a 5'-CTAG-3' target sequence to generate the stop codon for its large ORF1. J Bacteriol 175: 3151–3160

Fujiwara T, Mizuuchi K (1988) Retroviral DNA integration: structure of an integration intermediate. Cell 54: 497–504

Galas DJ, Chandler M (1989) Bacterial insertion sequences. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington DC, pp 109–162

Goussard S, Sougakoff W, Mabilat C, Bauernfeind A, Courvalin (1991) An IS1-like element is responsible for high-level synthesis of extended-spectrum β-lactamase TEM-6 in Enterobacteriaceae. J Gen Microbiol 137: 2681–2687

Grandgenett DP, Vora AC, Swanstrom R, Olsen JC (1986) Nuclease mechanism of the avian retrovirus pp32 endonuclease. J Virol 58: 970–974

Green EP, Tizard MLV, Moss MT, Thompson J, Winterborne DJ, McFadden JJ, Hermon-Taylor J (1989) Sequence and characteristics of IS900, an insertion element identified in human Crohn's disease isolate of *Mycobacterium paratuberculosis*. Nucleic Acids Res 17: 9063–9073

Grindley NDF (1978) IS1 insertion generates duplication of a nine base pair sequence at its target site. Cell 13: 419–426

Grindley NDF, Joyce CM (1981) Genetic DNA sequence analysis of the kanamycin resistance transposon Tn903. Proc Natl Acad Sci USA 77: 7176–7180

Henderson DJ, Lydiate DJ, Hopwood DA (1989) Structural and functional analysis of the minicircle, a transposable element of *Streptomyces coelicolor* A3(2). Mol Microbiol 3: 1307–1318

Henikoff S (1992) Detection of *Caenorhabditis* transposon homologs in diverse organisms. New Biol 4: 382–388

Hirsch H-J, Starlinger P, Brachet P (1972) Two kinds of insertions in bacterial genes. Mol Gen Genet 119: 191–206

Hizi A, Henderson LE, Copeland TD, Sowder RC, Hixson CV, Oroszlan S (1987) Characterization of mouse mammary tumor virus gag-pro gene products and the ribosomal frameshift site by protein sequencing. Proc Natl Acad Sci USA 84: 7041–7045

Hoover TA, Vodkin MH, Williams JC (1992) A *Coxiella burnetii* repeated DNA element resembling a bacterial insertion sequence. J Bacteriol 174: 5540–5548

Hu S, Ptashne K, Cohen SN, Davidson N (1975) αβ sequence of F is IS3. J Bacteriol 123: 687–692

Iida S, Arber W (1980) On the role of IS1 in the formation of hybrids between the bacteriophage P1 and the R plasmid NR1. Mol Gen Genet 177: 261–270

Jacks T, Townsley K, Varmus HE, Majors J (1987) Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. Proc Natl Acad Sci USA 84: 4298–4302

Jacks T, Madhani HD, Masiarz FR, Varmus HE (1988) Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell 55: 447–458

Jakowec M, Prentki P, Chandler M, Galas DJ (1988) Mutational analysis of the open reading frames in the transposable element IS1. Genetics 120: 47–55

Jaraczewski JW, Jahn CL (1993) Elimination of Tec elements involves a novel excision process. Genes Dev 7: 95–105

Johnsrud L (1979) DNA sequence of the transposable element IS1. Mol Gen Genet 169: 213–218

Jordan E, Saedler H, Starlinger P (1968) 0-zero and strong polar mutations in the gal operon are insertions. Mol Gen Genet 102: 353–363

Kanazawa H, Kiyasu T, Noumi T, Futai M, Yamaguchi K (1984) Insertion of transposable elements in the promoter proximal region of the gene cluster for *Escherichia coli* H^+-ATPase: 8 base pair repeat generated by insertion of IS1. Mol Gen Genet 194: 179–187

Kato K, Ohtsuki K, Mitsuda H, Yomo T, Negoro S (1994) Insertion sequence IS6100 on plasmid pOAD2, which degrades nylon oligomers. J Bacteriol 176: 1197–1200

Katz RA, Merkel G, Kulkosky J, Leis J, Skalka AM (1990) The avian retroviral IN protein is both necessary and sufficient for integrative recombination in vitro. Cell 63: 87–95

Katzman M, Katz RA, Skalka AM, Leis J (1989) The avain retroviral integration protein cleaves the terminal sequences of linear viral DNA at the in vivo sites of integration. J Virol 63: 5319–5327

Khan E, Mack JPG, Katz RA, Kulkosky J, Skalka AM (1991) Retroviral integrase domains: DNA binding and the recognition of LTR sequences. Nucleic Acids Res 19: 851–860

Klaer R, Starlinger P (1980) IS4 at its chromosomal site in *E. coli* K-12. Mol Gen Genet 178: 285–291

Klaer R, Kühn S, Tillman E, Fritz H-J, Starlinger P (1981) The sequence of IS4. Mol Gen Genet 181: 169–175

Kohara Y, Akiyama K, Isono K (1987) The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell 50: 495–508

Komoda Y, Enomoto M, Tominaga A (1991) Large inversion in *Escherichia coli* K-12 1485IN between inversely oriented IS3 elements near lac and cdd. Genetics 129: 639–645

Kulkosky J, Jones KS, Katz RA, Mack JPG, Skalka AM (1992) Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol Cell Biol 12: 2331–2338

Kunze ZM, Wall S, Appelberg R, Silva MT, Portaels F, McFadden JJ (1991) IS901, a new member of a widespread class of atypical insertion sequences is associated with pathogenicity in *Mycobacterium avium*. Mol Microbiol 5: 2265–2272

Lawrence JG, Ochman H, Hartl DL (1992) The evolution of insertion sequences within enteric bacteria. Genetics 131: 9–20

Lenich AG, Glasgow AC (1994) Amino acid sequence homology between Piv, an essential protein in site-specific DNA inversion in *Moraxella lacunata,* and transposases of an unusual family of insertion elements. J Bacteriol 176: 4160–4164

Leskiw BK, Mevarech M, Barritt LS, Jensen SE, Henderson DJ, Hopwood DA, Bruton CJ, Chater KF (1990) Discovery of an insertion sequence, IS116, from *Streptomyces clavuligerus* and its relatedness to other transposable elements from actinomycetes. J Gen Microbiol 136: 1251–1258

Luthi K, Moser M, Ryser J, Weber H (1990) Evidence for a role of translational frameshifting in the expression of transposition activity of the bacterial insertion element IS1. Gene 88: 15–20

Machida C, Machida Y (1989) Regulation of IS1 transposition by the insA gene product. J Mol Biol 208: 567–574

Machida C, Machida Y, Ohtsubo E (1984) Both inverted repeat sequences located at the ends of IS1 provide promoter functions. J Mol Biol 177: 247–267

Machida Y, Machida C, Ohtsubo H, Ohtsubo E (1982) Factors determining frequency of plasmid cointegration mediated by insertion sequence IS1. Proc Natl Acad Sci USA 79: 277–281

Machida Y, Machida C, Ohtsubo E (1984) Insertion element IS1 encodes two structural genes required for its transposition. J Mol Biol 177: 229–245

Maekawa T, Ohtsubo E (1994) Identification of the region that determines the specificity of binding of the transposases encoded by Tn3 and γδ to the terminal inverted repeat sequences. Jpn J Genet 69: 269–285

Mahillon J, Seurinck J, van Rompuy L, Delcour J, Zabeau M (1985) Nucleotide sequences and structural organization of an insertion sequence element (IS231) from *Bacillus thuringiensis* strain berliner 1715. EMBO J 4: 3985–3899

Mahillon J, Seurinck J, Delcour J, Zabeau M (1987) Cloning and nucleotide sequence of different iso-IS231 elements and their structural association with the Tn4430 transposon in *Bacillus thuringiensis*. Gene 51: 187–196

Malamy MH (1966) Frameshift mutations in the lactose operon of *E. coli*. Cold Spring Harb Symp Quant Biol 31: 189–201

Malamy MH (1970) Some properties of insertion mutations in the lac operon. In: Beckwith JR, Zipser D (eds) The loctose operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 359–373

Malamy MH, Fiandt M, Szybalski W (1972) Electron microscopy of polar insertions in the lac operon of *Escherichia coli*. Mol Gen Genet 119: 207–222

Malamy MH, Rahaim PT, Hoffman CS, Baghdoyan D, O'Connor MB, Miller JF (1985) A frameshift mutation at the junction of an IS1 insertion within lacZ restores β-galactosidase activity via formation of an active lacZ-IS1 fusion protein. J Mol Biol 181: 551–555

Martin C, Timm J, Rauzier J, Gomez-Lus R, Davies J, Gicquel B (1990) Transposition of an antibiotics resistance element in mycobacteria. Nature 345: 739–743

Matsutani S, Ohtsubo E (1993) Distribution of the *Shigella sonnei* insertion elements in Enterobacteriaceae. Gene 127: 111–115

Matsutani S, Ohtsubo H, Maeda Y, Ohtsubo E (1987) Isolation and characterization of IS elements repeated in the bacterial chromosome. J Mol Biol 196: 445–455

McClintock B (1956) Controlling elements and the gene. Cold Spring Harb Symp Quant Biol 21: 197–216

McClintock B (1965) The control of gene action in maize. Brookhaven Simp Biol 18: 162–184

Mendiola MV, de la Cruz F (1989) Specificity of insertion of IS91, an insertion sequence present in alpha-hemolysin plasmids of *Escherichia coli*. Mol Microbiol 3: 979–984

Mendiola MV, de la Cruz F (1992) IS91 transposase is related to the rolling-circle-type replication protein of the pUB110 family of plasmids. Nucleic Acids Res 20: 3521

Mendiola MV, Jubete Y, de la Cruz F (1992) DNA sequence of IS91 and identification of the transposase gene. J Bacteriol 174: 1345–1351

Mills JA, Venkatesan MM, Baron LS, Buysse JM (1992) Spontaneous insertion of an IS1-like element into the *virF* gene is responsible for avirulence in opaque colonial variants of *Shigella flexneri* 2a. Infect Immun 60: 175–182

Moore R, Dixon M, Smith R, Peters G, Dickson C (1987) Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of gag and pol. J Virol 61: 480–490

Moss MT, Malik ZP, Tizard MLV, Green EP, Sanderson JD, Hermon-Taylor J (1992) IS902, an insertion element of the chronic-enteritis-causing *Mycobacterium avium* subsp. *silvaticum*. J Gen Microbiol 138: 139–145

Nakatsu C, Ng J, Singh R, Straus N, Wyndham C (1991) Chlorobenzoate catabolic transposon Tn5271 is a composite class I element with flanking class II insertion sequences. Proc Natl Acad Sci USA 88: 8312–8316

Nyman K, Nakamura K, Ohtsubo H, Ohtsubo E (1981) Distribution of the insertion sequence IS1 in gram-negative bacteria. Nature 289: 609–612

Ohtsubo E, Zenilman M, Ohtsubo H (1980) Plasmids containing insertion elements are potential transposons. Proc Natl Acad Sci USA 77: 750–754

Ohtsubo E, Zenilman M, Ohtsubo H, McCormick M, Machida C, Machida Y (1981) Mechanism of insertion and cointegration mediated by IS1 and Tn3. Cold Spring Harb Symp Quant Biol 45: 283–295

Ohtsubo E, Ohtsubo H, Doroszkiewicz W, Nyman K, Allen D, Davison D (1984) An evolutionary analysis of iso-IS1 elements from *Escherichia coli* and *Shigella* strains. J Gen Appl Microbiol 30: 359–376

Ohtsubo H, Ohtsubo E (1978) Nucleotide sequence of an insertion element, IS1. Proc Natl Acad Sci USA 75: 615–619

Ohtsubo H, Nyman K, Doroszkiewicz W, Ohtsubo E (1981) Multiple copies of iso-insertion sequences of IS1 in *Shigella dysenteriae* chromosome. Nature 292: 640–643

Ohtsubo H, Zenilman M, Ohtsubo E (1980) Insertion element IS102 resides in plasmid pSC101. J Bacteriol 144: 131–140

Olasz F, Stalder R, Arber W (1993) Formation of the tandem repeat (IS30)2 and its role in IS30-mediated transpositional DNA rearrangements. Mol Gen Genet 239: 177–187

Ou JT, Huang CJ, Houng HS, Baron LS (1992) Role of IS1 in the conversion of virulence (Vi) antigen expression in Enterobacteriaceae. Mol Gen Genet 234: 228–232

Pabo C, Sauer R (1984) Protein-DNA recognition. Annu Rev Biochem 53: 293–321

Polard P, Prère MF, Chandler M, Fayet O (1991) Programmed translational frameshifting and initiation at an AUU codon in gene expression of bacterial insertion sequence IS911. J Mol Biol 222: 465–477

Polard P, Prère MF, Fayet O, Chandler M (1992) Transposase-induced excision and circularization of the bacterial insertion sequence IS911. EMBO J 11: 5079–5090
Prère MF, Chandler M, Fayet O (1990) Transposition in *Shigella dysenteriae*: isolation and analysis of IS911, a new member of the IS3 group of insertion sequences. J Bacteriol 172: 4090–4099
Rådström P, Sköld O, Swedberg G, Flensburg J, Roy PH, Sundström (1994) Transposon Tn5090 of plasmid R751, which carries an integron, is related to Tn7, Mu and retroelements. J Bacteriol 176: 3257–3268
Ramirez SJ, Alvarez G, Cisneros E, Gomez EM (1992) Distribution of insertion sequence IS1 in multiple-antibiotic resistant clinical Enterobacteriaceae strains. FEMS Microbiol Lett 72: 189–193
Reimmann C, More R, Little S, Savioz A, Willetts NS, Haas D (1989) Genetic structure, function and regulation of the transposable element IS21. Mol Gen Genet 215: 416–424
Rezsöhazy R, Hallet B, Delcour J (1992) IS231D, E, and F, three new insertion sequences in *Bacillus thuringiensis:* extension of the IS231 family. Mol Microbiol 6: 1959–1967
Rezsöhazy R, Hallet B, Delcour J, Mahillon J (1993a) The IS4 family of insertion sequences: evidence for a conserved transposase motif. Mol Microbiol 9: 1283–1295
Rezsöhazy R, Hallet B, Mahillon J, Delcour J (1993b) IS231V and W from *Bacillus thuringiensis* subsp. *israelensis,* two distant members of the IS231 family of insertion sequences. Plasmid 30: 141–149
Rice NR, Stephens RM, Burny A, Gilden RV (1985) The gag and pol genes of bovine leukemia virus: nucleotide sequence and analysis. Virology 142: 357–377
Rodriguez H, Snow ET, Bhat U, Loechler (1992) An *Escherichia coli* plasmid-based, mutational system in which supF mutants are selectable: insertion elements dominate the spontaneous spectra. Mutat Res 270: 219–231
Romantschuk M, Richter GY, Mukhopadhyay P, Mills (1991) IS801, an insertion sequence element isolated from *Pseudomonas syringae phaseolicola.* Mol Microbiol 5: 617–622
Rose AM, Snutch TP (1984) Isolation of the closed circular form of the transposable element Tc1 in *Caenorhabditis elegans.* Nature 311: 485–486
Ruan KS, Emmons SW (1984) Extrachromosomal copies of transposon Tc1 in the nematode *Caenorhabditis elegans.* Proc Natl Acad Sci USA 81: 4018–4022
Schwartz E, Kröger M, Rak B (1988) IS150: distribution, nucleotide sequence and phylogenetic relationships of a new *E. coli* insertion element. Nucleic Acids Res 16: 6789–6802
Scott JR, Kirchman PA, Caparon MG (1988) An intermediate in transposition of the conjugative transposon Tn916. Proc Natl Acad Sci USA 85: 4809–4813
Sekine Y, Ohtsubo E (1989) Frameshifting is required for expression of IS1 transposase. Proc Natl Acad Sci USA 86: 4609–4613
Sekine Y, Ohtsubo E (1991) Translational frameshifting in IS elements and other genetic systems. In: Kimura M, Takahata N (eds) New aspects of the genetics of molecular evolution. Japan Sci Soc Press, Tokyo/Springer, Berlin Heidelberg New York, pp 243–261
Sekine Y, Ohtsubo E (1992) DNA sequences required for translational frameshifting in production of the transposase encoded by IS1. Mol Gen Genet 235: 325–332
Sekine Y, Nagasawa H, Ohtsubo E (1992) Identification of the site of translational frameshifting required for production of the transposase encoded by insertion sequence IS1. Mol Gen Genet 235: 317–324
Sekine Y, Eisaki N, Ohtsubo E (1994) Translational control in production of transposase and in transposition of insertion sequence IS3. J Mol Biol 235: 1406–1420
Sekine Y, Eisaki N, Ohtsubo E (1995) Identification and characterization of the linear IS3 molecules generated by staggered breaks. J Biol Chem (in press)
Sekino N, Sekine Y, Ohtsubo E (1995) IS1-encoded proteins, InsA and the InsA-B'-InsB transframe protein (transposase): functions deduced from their DNA-binding ability. Adv Biophys 31: 209–222
Shapiro JA (1969) Mutations caused by insertion of genetic material into the galactose operon of *Escherichia coli.* J Mol Biol 40: 93–105
Sherman PA, Fyfe JA (1990) Human immunodeficiency virus integration protein expressed in *Escherichia coli* possesses selectine DNA cleaving activity. Proc Natl Acad Sci USA 87: 5119–5123
Sherratt D (1989) Tn3 and related transposable elements: site-specific recombination and transposition. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington DC, pp163–184
Shimotohno K, Takahashi Y, Shimizu N, Gojobori T, Golde DW, Chen IS, Miwa YM, and Sugimura T (1985) Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: an open reading frame for the protease gene. Proc Natl Acad Sci USA 82: 3101–3105
Skaliter R, Eichenbaum Z, Shwartz H, Ascarelli GR, Livneh Z (1992) Spontaneous transposition in bacteriophage lambda *cro* gene residing on a plasmid. Mutat Res 267: 139–151

Sommer H, Cullum J, Saedler H (1979) Integration of IS3 into IS2 generates a short sequence duplication. Mol Gen Genet 177: 85–89

Spielmann-Ryser J, Moser M, Kast P, Weber H (1991) Factors determining the frequency of plasmid cointegrate formation mediated by insertion sequence IS3 from *Escherichia coli*. Mol Gen Genet 226: 441–448

Stark WM, Boocock MR, Sherratt DJ (1992) Catalysis by site-specific recombinases. Trends Genet 8: 432–439

Starlinger P, Saedler H (1976) IS-elements in microorganisms. In: Compans RW, Cooper M, Koprowski H et al. (eds) Current topics in microbiology and immunology, vol 75. Springer, Berlin Heidelberg New York, pp 111–152

Tenzen T, Ohtsubo E (1991) Preferential transposition of an IS630-associated composite transposon to TA in the 5'-CTAG-3' sequence. J Bacteriol 173: 6207–6212

Tenzen T, Matsutani S, Ohtsubo E (1990) Site-specific transposition of insertion sequence IS630. J Bacteriol 172: 3830–3836

Tenzen T, Matsuda Y, Ohtsubo H, Ohtsubo E (1994) Transposition of Tnr1 in rice genomes to 5'-PuTAPy-3' duplicating the TA sequence. Mol Gen Genet 245: 441–448

Terry R, Soltis DA, Katzman M, Cobrinik D, Leis J, Skalka AM (1988) Properties of avain sarcoma leukosis virus pp32–related pol-endonuclease produced in *Escherichia coli*. J Virol 62: 2358–2365

Timmerman KP, Tu CD (1985) Complete sequence of IS3. Nucleic Acids Res 13: 2127–2139

Toba MM, Hashimoto GT (1992) Characterization of the spontaneous elimination of streptomycin sensitivity (SmS) on high-copy-number plasmids: SmS-enforcement cloning vectors with a synthetic *rpsL* gene. Gene 121: 25–33

Trinks K, Habermann P, Beyreuther K, Starlinger P, Ehring R (1981) An IS4-encoded protein is synthesized in minicells. Mol Gen Genet 182: 183–188

Tsuchihashi Z, Brown PO (1992) Sequence requirements for efficient translational frameshifting in the *Escherichia coli dnaX* gene and the role of an unstable interaction between $tRNA^{Lys}$ and an AAG lysine codon. Genes Dev 6: 511–519

Tsuchihashi Z, Kornberg A (1990) Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme. Proc Natl Acad Sci USA 87: 2516–2520

Tudor M, Lobocka M, Goodell M, Pettitt J, O'Hare K (1992) The pogo transposable element family of *Drosophila melanogaster*. Mol Gen Genet 231: 126–134

Umeda M, Ohtsubo E (1989) Mapping of insertion elements IS1, IS2 and IS3 on the *E. coli* K-12 chromosome: role of insertion elements in formation of Hfrs and F-prime factors and in rearrangement of bacterial chromosomes. J Mol Biol 208: 601–614

Umeda M, Ohtsubo E (1990a) Mapping of insertion element IS5 in the *Escherichia coli* K-12 chromosome: chromosomal rearrangement mediated by IS5. J Mol Biol 213: 229–237

Umeda M, Ohtsubo E (1990b) Mapping of insertion element IS30 on the *Escherichia coli* K-12 chromosome. Mol Gen Genet 222: 317–322

Umeda M, Ohtsubo E (1991) Four types of IS1 with difference in nucleotide sequences residue in the *Escherichia coli* K-12 chromosome. Gene 98: 1–5

Umeda M, Ohtsubo H, Ohtsubo E (1991) Diversification of the rice *wx* gene by insertion of mobile DNA elements into introns. Jpn J Genet 66: 569–586

van der Meer JR, Zehnder AJB, de Vos WM (1991) Identification of a novel composite transposable element, Tn5280, carrying chlorobenzene dioxygenase genes of *Pseudomonas* sp. strain P51. J Bacteriol 173: 7077–7083

van Gent DC, Oude Groeneger AAM, Plasterk RHA (1992) Mutational analysis of the integrase protein of human immunodeficiency virus type 2. Proc Natl Acad Sci USA 89: 9598–9602

van Hove B, Staudenmaier H, Braun V (1990) Novel two-component transmembrane transcriptional control: regulation of iron dicitrate transport in *Escherichia coli* K-12. J Bacteriol 172: 6749–6758

Varmus H, Brown P (1989) Retroviruses. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington DC, pp 53–108

Vögele K, Schwartz E, Welz C, Schiltz E, Rak B (1991) High-level ribosomal frameshifting directs the synthesis of IS150 gene products. Nucleic Acids Res 19: 4377–4385

Weiss RB, Dunn DM, Atkins JFM, Gesteland RF (1987) Slippery runs, shifty stops, backward steps, and forward hops: –2, –1, +1, +2, +5, and +6 ribosomal frameshifting. Cold Spring Harb Symp Quant Biol 52: 687–693

Yoshioka Y, Ohtsubo H, Ohtsubo E (1987) Repressor gene *finO* in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS3 into *finO*. J Bacteriol 169: 619–623

Yoshioka Y, Fujita Y, Ohtsubo E (1990) Nucleotide sequence of the promoter distal region of the tra operon including tral (DNA helicase I) and *traD* genes of plasmid R100. J Mol Biol 214: 39–53

Zerbib D, Jakowec M, Prentki P, Galas DJ, Chandler M (1987) Expression of proteins essential for IS1 transposition: specific binding of InsA to the ends of IS1. EMBO J 6: 3163–3169

Zerbib D, Polard P, Escoubas JM, Galas D, Chandler M (1990a) The regulatory role of the IS1-encoded InsA protein in transposition. Mol Microbiol 4: 471–477

Zerbib D, Prentki P, Gamas P, Freund E, Galas DJ, Chandler M (1990b) Functional organization of the ends of IS1: specific binding for an IS1-encoded protein. Mol Microbiol 4: 1477–1486

Zuber U, Schumann W (1993) The eighth copy of IS1 in *Escherichia coli* W3110 maps at 49.6 min. J Bacteriol 175: 1552

Zuerner RL (1994) Nucleotide sequence analysis of IS1533 from *Leptospira borgpetersenii*: identification and expression of two IS-encoded proteins. Plasmid 31: 1–11