Functional comparison of the transposition core machineries of phage Mu and *Haemophilus influenzae* Mu-like prophage Hin–Mu reveals interchangeable components

Anna-Helena Saariaho\(^a\), Arja Lamberg\(^a\), Seija Elo\(^b\), Harri Savilahti\(^a,\text{*}^\)

\(^a\)Institute of Biotechnology, Program in Cellular Biotechnology, Viikki Biocenter, University of Helsinki, PO Box 56, Viikinkaari 9, 00014 Helsinki, Finland

\(^b\)Division of General Microbiology, Department of Biosciences, Viikki Biocenter, University of Helsinki, 00014 Helsinki, Finland

Received 27 October 2003; returned to author for revision 23 November 2003; accepted 16 September 2004

Available online 6 November 2004

**Abstract**

Bacteriophage Mu uses DNA transposition for propagation and is a model for transposition studies in general. Recent identification of Mu-like prophages within bacterial genomes offers new material for evolutionary and comparative functional studies. One such prophage, Hin–Mu of *Haemophilus influenzae* Rd, was studied for its transpositional properties. The components of its transposition core machinery, the encoded transposase (MuA\(_{\text{hin}}\)) and the transposase binding sites, were evaluated for functional properties by sequence comparisons and DNase I footprinting. Transpositional activity of Hin–Mu was examined by in vitro assays directly assessing the assembly and catalytic function of the transposition core machinery. The Hin–Mu components readily assembled catalytically competent protein–DNA complexes, transpososomes. Thus, Hin–Mu encodes a functional transposase and contains critical transposase binding sites. Despite marked sequence differences, components of the Hin–Mu and Mu transposition core machineries are partially interchangeable, reflecting both conservation and flexibility in the functionally important regions within the transpososome structure.

© 2004 Elsevier Inc. All rights reserved.

**Keywords**: Bacteriophage Mu; DNase I footprinting; In vitro DNA transposition; MuA transposase; Prophage

**Introduction**

DNA transposition plays an important role in the evolution of both eukaryotic and prokaryotic genomes and involves a series of DNA restructuring reactions by which certain mobile genetic elements move within and between genomes (Craig, 1995). In prokaryotes, this class of mobile elements includes insertion sequences (IS elements), composite transposons, and certain viruses that propagate with the aid of DNA transposition, that is, transposing bacteriophages such as Mu (Craig et al., 2002). Only a few similar phages have been characterized (DuBow, 1987) including the coliphage D108 (Hull et al., 1978) and *Pseudomonas aeruginosa* phage D3112 (Wang et al., 2004). However, completion of several bacterial genome projects has enabled identification of Mu-like phages within genomes where they represent integrated prophages or their remnants (Morgan et al., 2001). Recently, some of these prophages, such as that of *Haemophilus influenzae* (Fleischmann et al., 1995), have been scrutinized for their genomic organization (Morgan et al., 2001), but none of the identified prophages have been characterized for their transpositional properties. Such characterizations should reveal new activities and provide novel means for functional and evolutionary comparisons of this group of elements.

In general, DNA transposition proceeds within the context of a higher-order nucleoprotein complex, a transpososome (Chaconas et al., 1996; Craig et al., 2002; Davies et al., 2000). Within the core of this molecular machine, a multimer of a transposon-encoded transposase synapses two transposon ends via specific protein–protein and protein–DNA interactions.
DNA interactions. While transpososomes may also contain other proteins, the minimal catalytic core comprises the transposon ends and a few transposase protomers only. DNA transposition reaction proceeds through several distinct but chemically similar transposase-catalyzed steps (Craig, 1995), including initial cleavages at the transposon–host DNA boundaries (donor cleavage) and covalent integration of the transposon DNA into the target DNA (strand transfer).

The 36,717-bp genome of the bacteriophage Mu (Morgan et al., 2001) is one of the largest, most efficient, and complex transposons known (Chaconas and Harshey, 2002). Despite its complexity, phage Mu has served as a model system for transposition studies (Mizuuchi, 1992) primarily due to the high efficiency of in vivo transposition (Symonds et al., 1987) and early development of an efficient in vitro transposition system (Mizuuchi, 1983). Mu transposition involves a number of phage-specific DNA factors, several phage- and host-encoded proteins, as well as certain DNA topology (reviewed by Chaconas and Harshey, 2002). While certain accessory DNA sites are also involved in Mu transposition, the most critical DNA factors include the phage genome ends carrying three transposase binding sites each: L1, L2, L3 in the left (L-end) and R1, R2, R3 in the right (R-end). The most important protein is the phage-encoded 75-kDa MuA transposase (Fig. 1A), a multidomain protein that in its core domain contains a phylogenetically conserved and catalytically important triad of acidic amino acids known as the DDE motif. MuA binds as an inert monomer to each of its six binding sites within the genome ends and initiates an elaborate transpososome assembly pathway that leads to synopsis of the genome ends and concomitant tetramerization of the transposase. However, within the fully assembled transpososome, MuA binds tightly to only three of the binding sites: L1, R1, and R2. Efficient transposition also requires phage-encoded MuB, an ATP-dependent DNA-binding protein that functions in transpososome activation and transposition target selection. In addition, host-encoded proteins are involved in Mu transposition, and they include DNA-bending proteins (HU and IHF) as well as protein remodeling and DNA replication factors. Five base pairs of target DNA are duplicated as a consequence of transposition, and the duplication can be detected as a pair of directly repeated sequences that flank the transposon DNA.

The original Mu in vitro transposition reaction involves a model superhelical plasmid substrate that contains the critical DNA sequences, MuA, MuB, and at least one of the DNA bending proteins (Craigie et al., 1985; Mizuuchi, 1983). However, under altered conditions, the reaction can be performed with MuA and a short Mu end-specific DNA segment as the only macromolecular components (Savilahti et al., 1995). This minimal reaction faithfully reproduces transpososome assembly, donor cleavage, and strand transfer steps; and it has been used effectively in a number of detailed studies of transpososome function and organization (reviewed by Chaconas and Harshey, 2002).

Here, we identified and characterized the molecular components of the *H. influenzae* Mu-like prophage Hin–Mu transposition core machinery, including the transposase binding sites at the Hin–Mu genome ends and the product of the Hin–Mu A gene, a MuA-like transposase termed *MuA*<sub>Hin</sub>. Using in vitro reactions, similar to those developed for phage Mu, we demonstrate that a short Hin–Mu right end segment and *MuA*<sub>Hin</sub> can assemble a functional Hin–Mu transpososome core. Comparative studies revealed not only important differences between the Hin–Mu and Mu transpososome core components but also a degree of conservation and interchangeability.

**Results**

**Hin–Mu prophage and its ends**

The Mu-like prophage (Hin–Mu) within the genome of *H. influenzae* Rd (Fleischmann et al., 1995) shares significant homology with the genome of phage Mu and is colinear with respect to genomic organization (Morgan et al., 2001). To distinguish the exact Hin–Mu prophage termini, we compared the DNA sequence of *H. influenzae* Rd to that of a strain T7386 that did not contain a prophage at the corresponding locus. This comparison showed that Hin–Mu prophage spans the *H. influenzae* Rd genome (GenBank acc. no. NC_000907) between coordinates 1,559,722–1,594,398 and is 34,676 bp. The prophage DNA ended with the sequence CA-3′ at both termini (similar to Mu) and contained identifiable transposase binding sites that aligned well with those of Mu and its closest relative, phage D108 (see below). The prophage was flanked on both sides by a 5-bp direct repeat (ACGCA), present only as a single copy in the T7386 genome. These data indicate that Hin–Mu represents a full-length integrated copy of a Mu-like phage genome and suggest that its genomic integration was a result of DNA transposition.

**MuA<sub>Hin</sub> transposase**

Phage Mu-encoded MuA transposase is a 663-amino acid (aa) product of gene A and can be divided into several structurally and functionally distinct domains (Fig. 1A). A similar gene in the Hin–Mu prophage encodes a putative 687-aa transposase homologous to MuA, referred to here as *MuA*<sub>Hin</sub>. The amino acid sequences of *MuA* and *MuA*<sub>Hin</sub> were aligned and compared with regard to structural and functional characteristics on the basis of information available for MuA. The two proteins share significant amino acid similarity in each domain, and their domain organization is colinear (Fig. 1B, Table 1). The length difference is due to relatively short (up to 10 aa) insertions/
deletions, which are primarily located within the domain I. The similarity is highest (63%) within the domain II, and the MuA<sub>Hin</sub> motif (Asp 279, Asp 344, Glu 400) aligns well with the DDE motif of MuA (Asp 269, Asp 336, Glu 392). The domains I and III also exhibit significant but relatively lower sequence similarity.
Putative Hin–Mu binding sites

Mu and its closest relative, phage D108, carry three transposase binding sites, L1–L3 and R1–R3, with identical organization at the L-end and R-end of their genomes, respectively. To determine potential transposase binding sites at Hin–Mu ends, we aligned and compared the sequences of Mu, D108, and Hin–Mu (Fig. 2). At the R-end, potential Hin–Mu R1 and R2 binding sites are very similar to those of Mu and D108, both in sequence
and location. However, two alternatives exist for a potential Hin–Mu R3 site, 9 bp apart, and they were designated R3 and R3*. At the L-end, potential Hin–Mu L1 and L2 binding sites are also identifiable as well as two potential L3 site alternatives (8 bp apart, designated L3 and L3*).

Binding site similarities were further studied by aligning all Mu, D108, and putative Hin–Mu transposase binding sites and comparing them to a 22-bp consensus sequence generated from all six Mu sites: R1–R3 and L1–L3 (Fig. 3). Each binding site was additionally scored with regard to nucleotide conservation (Fig. 3). While the Hin–Mu binding site scores were generally somewhat lower than those of Mu and D108, sequence conservation was evident. Of the potential Hin–Mu binding site alternatives, the R3* and L3* sites matched better to the consensus sequence than the R3 and L3 sites.

DNase I footprinting analysis

Interaction of transposase molecules with the specific binding sites at the transposon ends constitutes an important early step in DNA transposition. In general, this initial binding can be studied by using various footprinting techniques under reaction conditions which allow transposase binding but do not promote transpososome assembly (restrictive conditions). We chose to use DNase I footprinting for such MuAHin binding studies and, as a control, also included MuA and Mu ends in our studies. Somewhat surprisingly, while clear MuA footprints could be readily obtained on both Mu ends (Figs. 4A, B), we were not able to obtain MuAHin footprints on either of the Hin–Mu ends (not shown). These data indicate that either MuAHin does not bind to its putative binding sites or the binding affinities to these sites are low and fall below the detection limit.

As both the encoded transposases and the transposon ends of Mu and Hin–Mu share significant similarity, we examined if Mu ends could be footprinted with MuAHin and, vice versa, whether MuA would footprint Hin–Mu ends, possibly revealing conservation between the two systems. In these experiments, MuAHin yielded no detectable footprints on either of the Mu ends (not shown). Similarly, MuA did not footprint the Hin–Mu L-end (Fig. 4B). However, clear MuA protection patterns were detectable on the Hin–Mu R-end, and the patterns were in good general agreement with the binding site predictions (Fig. 4A).

Transpososome assembly and catalytic activities

MuAHin did not generate detectable footprints in the above studies. As the relationship between transposase binding site-affinities and transpososome assembly is not known, we examined whether MuAHin was nevertheless able to promote assembly of catalytically competent transpososomes. This was studied by using a minimal component in vitro reaction either with short radioactively labeled DNA segments as transposon-specific substrates alone or with addition of plasmid DNA as a target (Fig. 5A). The reaction conditions employed were identical to those generally used for analogous Mu reactions (Savilahti et al., 1995).

| Position | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | Score |
|----------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|
| R1 Mu    | C | G | G | G | C | A | C | A | A | A | A | A | A | C | G | G | G | G | A | A | A | G | 14.5 |
| Hin-Mu   | C | G | A | G | C | A | T | A | A | A | A | C | T | C | G | G | A | A | A | A | 14.0 |
| R2 Mu    | C | G | T | T | T | C | A | C | G | A | T | A | A | A | T | C | G | G | G | A | A | 17.0 |
| D108     | C | G | T | T | T | C | A | A | A | A | A | A | A | A | A | A | A | A | A | A | 15.5 |
| Hin-Mu   | A | A | A | A | C | C | A | T | A | A | A | A | A | C | T | C | A | G | A | A | A | 12.5 |
| R3 Mu    | T | G | T | T | T | C | C | T | G | A | A | G | C | G | C | G | A | A | A | G | 17.5 |
| D108     | T | G | T | C | T | C | A | T | G | A | A | G | C | G | C | G | A | A | A | G | 16.0 |
| Hin-Mu   | T | A | A | A | A | A | A | C | G | A | A | A | A | A | A | A | A | A | A | A | 10.5 |
| R3* Hin-Mu | T | G | A | G | T | C | A | T | A | A | A | A | A | C | C | G | A | A | A | A | 14.5 |
| L1 Mu    | T | G | A | T | C | A | C | T | G | A | G | G | T | A | G | A | A | A | A | 16.5 |
| Hin-Mu   | A | A | A | A | A | A | A | C | C | A | A | A | A | A | A | A | A | A | A | 11.5 |
| L2 Mu    | G | T | T | A | A | T | C | A | A | T | G | A | A | A | C | G | C | G | A | A | A | 12.5 |
| D108     | G | T | T | A | T | A | C | T | C | T | G | A | A | A | C | G | C | G | A | A | 13.5 |
| Hin-Mu   | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | G | G | A | A | A | 11.0 |
| L3 Mu    | T | G | T | T | T | C | T | A | A | A | A | A | A | A | A | A | A | A | A | A | 17.0 |
| D108     | T | G | A | T | T | C | A | T | A | A | A | G | C | A | C | G | A | A | A | 15.5 |
| Hin-Mu   | G | T | G | T | G | A | T | T | A | G | T | C | A | A | A | A | T | G | A | 8.5 |
| L3* Hin-Mu | T | T | A | G | T | C | A | A | A | T | G | A | A | C | C | G | A | A | A | 11.5 |

Fig. 3. Alignment of the 22-bp transposase binding site sequences of Mu, D108, and Hin–Mu genomes. The six transposase binding sites of Mu and D108, and the eight putative Hin–Mu transposase binding sites are compared to a consensus derived from all six Mu binding sites (bold letters at the top). Nucleotides that match the consensus sequence are shaded with gray. Scores for binding sites were calculated as described in Materials and methods and are shown on the rightmost panel. Squared and circled nucleotides denote particularly critical and less critical nucleotides, respectively, as defined by in vivo analysis (Groenen and van de Putte, 1986).
First, the formation of stable protein–DNA complexes was studied by incubating the transposases with several types of transposon end-specific DNA fragments (Fig. 5B), and the reaction products were analyzed by native agarose gel electrophoresis (Fig. 5C). Mu-specific fragments in either uncut or precut form promoted transpososome assembly by MuA but not by MuA\textsubscript{Hin} (the two leftmost panels). However, Hin–Mu-specific fragments promoted assembly of potential transpososomes not only by MuA\textsubscript{Hin} but also by MuA (the two rightmost panels). The most efficient assembly was achieved with precut fragments in that the assembly did not require divalent metal ions and proceeded readily on ice.

Next, the catalytic steps were studied by reactions similar to those above but with the addition of a plasmid target. Following disassembly of protein–DNA complexes, the reaction products were analyzed by agarose gel electrophoresis (Fig. 5D). Mu end-fragments generated products with MuA but not with MuA\textsubscript{Hin} (lanes 1–8). However, Hin–Mu end-fragments generated products not only with MuA\textsubscript{Hin} but also with MuA (lanes 9–20). In combination with MuA\textsubscript{Hin} uncut and frayed Hin–Mu fragments generated only a limited number of single-ended integration products (lanes 11 and 19). However, the precut Hin–Mu fragment generated both single- and double-ended integration products in substantial quantities (lanes 14 and 15). All three types of Hin–Mu fragments also generated both kinds of reaction products in the presence of MuA (lanes 12, 16, and 20).

![Fig. 4. DNase I footprinting of MuA on Mu and Hin–Mu genome ends. The footprinted DNA fragments were labeled at the 5'-terminus ending on phage sequence. Prior to DNase I digestion, the transposon R-end- (A) or L-end-specific (B) DNA fragments were incubated in the absence of MuA (lane 1) or in the presence of 170 \( \mu \)g (lane 2), 34 \( \mu \)g (lane 3), 17 \( \mu \)g (lane 4) or 1.7 \( \mu \)g (lane 5) of MuA per milliliter. The arrows indicate the junction between the transposon and flanking DNA. Numbers to the right of each autoradiogram are coordinates from the transposon end and solid lines to the left indicate MuA-protected areas. DNase I-sensitive sites within the protected areas are indicated by vertical arrowheads below the sequence.](image-url)
Identity of the transposition reaction products

To verify the identity of the MuA_Hin transposition reaction products (Fig. 5D), several transposon/target DNA junctions were initially cloned and subsequently sequenced. The strategy employed mainly scores double-ended integration products (DEPs), thus revealing duplicated target site sequence flanking the transposon ends. It also scores single-ended integration products (SEPs) in cases where two transposon ends integrate independently into opposite strands of the target within close proximity of each other. In the latter case, the sequence determination will not reveal a target site duplication but a short deletion within the target DNA. Of the five clones sequenced, MuA_Hin-catalyzed reaction products on Hin–Mu ends produced three clones with an accurate target site duplication and two clones with a short deletion representing DEPs and SEPs, respectively. Similar results were obtained with MuA on Hin–Mu ends, as four clones represented DEPs and one clone SEPs. In each case, the 3'-end of the transposon was accurately joined to the pUC19 target plasmid DNA (data not shown).

Discussion

We identified and isolated core components of the H. influenzae prophage Hin–Mu transposition machinery and assessed functionality of the machinery by using in vitro reactions with purified MuA_Hin transposase and
an R-end segment of the prophage as the critical macromolecular components. In general, the advantage of these types of assays is that no phage propagation is required. Thus, phage-derived DNA within bacterial genomes, including functional prophages as well as various types of defective prophages, can be scrutinized effectively for transpositional recombination activity. In principle, critical activities can be revealed even in cases when the full transpositional activity of a given element would require cofactors and in cases when accumulated mutations have altered original activities. This is in striking contrast to generally used in vivo transposition assays that are able to detect full activities only.

---

**Fig. 5. Complex formation and strand transfer.** (A) In vitro transposition reaction schematics. Donor DNA fragments and a transposase assemble a transposition complex. Under reaction conditions with Mg\(^{2+}\), the complex then executes successive donor DNA cleavage and strand transfer reactions. With precleaved donor DNA fragments, the cleavage step is bypassed. Transfer into circular target DNA generates two major products (shown earlier for Mu reaction, Goldhaber-Gordon et al., 2002a; Lee and Harshley, 2001). A double-ended integration product (DEP) is generated when both donor DNA fragments are properly transferred to the target DNA, and concomitantly the target DNA becomes linearized. A single-ended integration product (SEP) is generated when only one of the donor DNA fragments is transferred into the target DNA, and the supercoiled circular target DNA becomes relaxed. Radiolabels at the 5'-end of the donor DNA fragments (asterisks) reveal the products following agarose gel electrophoresis and autoradiography. (B) DNA fragments used as donor DNA substrates for in vitro transposition reactions. They contain the first 50 bp of Mu or Hin–Mu phage genome right end sequence, including two transposase binding sites, R1 and R2 (rectangles), and several flanking nucleotides as depicted. Fragments were labeled at the 5'-end of the bottom strand (asterisks). The vertical arrows indicate the transposon end cleavage point. (C) Complex formation of Mu- and Hin–Mu-specific donors with MuA or MuA\(_{Hin}\). Transposition reactions contained a standard (1×) or 4-fold (4×) concentration of both donor DNA and transposase, and they were incubated either with Mg\(^{2+}\) at 30°C or on ice without Mg\(^{2+}\). Complex formation was analyzed by native agarose gel electrophoresis and autoradiography. Mu transpososomes (Savilahti et al., 1995) and a major species of similarly migrating stable MuA\(_{Hin}\)-specific complexes are indicated by "C". The identity of the latter complexes is not known, but their migration and correlation to strand transfer activity (panel D) suggest them being active transpososomes. The faster migrating MuA-specific complexes (Savilahti et al., 1995) and most probably also faster-migrating MuA\(_{Hin}\)-specific complexes (as they form efficiently upon incubation on ice, this study) do not represent transpososomes (indicated by "*"). The slower migrating species seen with MuA\(_{Hin}\) are indicated by "**", and they may represent alternative oligomeric forms of complexes (active or inactive). Unreacted donor fragments are indicated by "bDQ". (D) Strand transfer products of transposition reactions using Mu or Hin–Mu donors with MuA or MuA\(_{Hin}\). Single- and double-ended integration products are indicated by “SEP” and “DEP”, respectively, and unreacted donor molecules by “D”. The slower migrating minor products (indicated by "***") represent strand transfer products that contain, in addition to one Mu-end fragment, two covalently connected target DNA (pUC19) molecules. They are generated when transpososomes form with a pair of one Mu end fragment and a "pseudo-Mu-end" present within the target DNA (Goldhaber-Gordon et al., 2002b).
MuA<sub>him</sub>, a transposase structurally similar to MuA

MuA<sub>him</sub> and MuA transposases are similar, and the similarity is significant (>31%) within all domains and subdomains (Table 1). The DNA binding domain I is 41% similar despite several insertions/deletions. As transposases and their binding sites in DNA are expected to evolve as a pair, it is likely that the observed conservation in this domain reflects the conservation of the transposase binding sites. The highest degree of amino acid similarity between the transposases (63%) lies within domain II, in particular within and around the catalytically important DDE motif. The motif is highly conserved with respect to both residues and spacing underscoring the importance of its structural architecture. Domain III, responsible for nonspecific DNA binding as well as interaction with MuB and protein remodeling factor ClpX, is the least similar (32%). The contacts between this domain and its binding partners are expected to be less specific or less conserved, and thus, a high degree of amino acid conservation is not essential. The overall sequence conservation between the two proteins suggests that MuA<sub>him</sub> likely exhibits very similar structural features as MuA. Consequently, several of the distinct functions typical of transposases of transposing bacteriophages, as typified by the functions of MuA, were likely present in MuA<sub>him</sub> when <i>H. influenzae</i> Rd was lysogenized by Hin–Mu. It is possible that at least some of these functions have been retained in the present-day MuA<sub>him</sub>.

Transposase binding sites

While sequence comparisons identified potential transposase binding sites within Hin–Mu prophage ends, we were not able to detect transposase protection in DNase I footprinting studies with MuA<sub>him</sub>. Accumulation of deleterious mutations in the MuA<sub>him</sub> transposase and/or within the binding sites during Hin–Mu prophage evolution is consistent with these observations. However, MuA generated footprints on the Hin–Mu R-end, indicating conservation between the systems and/or flexibility in the MuA binding characteristics. MuA protection patterns on the Hin–Mu R-end include a characteristic footprint with one major DNase I-sensitive site in each of the predicted transposase binding sites. As these sensitivity are located in identical positions within the binding sites in both Hin–
Mu and Mu ends (Fig. 6), it is likely that MuA interacts with all of these sites in a qualitatively similar manner. These sensitivity sites detected likely reflect local DNA distortion by MuA. DNA bending is probably involved in the process, as it is known that MuA can effectively bend its binding sites (Kuo et al., 1991; Zou et al., 1991). Two of the putative transposase binding sites in the Hin–Mu ends were identified ambiguously. While the L3* and R3* sites scored better than L3 and R3 in the nucleotide conservation analysis, it is not clear which of the sites represent the major binding sites. However, an appropriately positioned DNase I-sensitive site within the footprinted area (Figs. 4A and 6) suggest that R3* probably is the major binding site in the R-end. Our MuA footprinting data confirm the existing data on MuA binding to Mu genome ends (Craigie et al., 1984; Zou et al., 1991) and reveal some new information. We detected a DNase I-sensitive site residing within Mu R3 that has not been reported previously (Fig. 4A). In addition, a similar but barely detectable sensitivity site, also not reported earlier, is apparent in Mu L2 (Fig. 4B). These two sensitivity sites match the positions of other sensitivity sites that have been mapped (Fig. 6), again most likely reflecting qualitatively similar MuA binding characteristics.

Some common features of Mu-like phage transposase binding site sequences could be identified in the nucleotide conservation analysis. In general, the first half (nt 1–11) of the binding sites appeared less conserved than the second half (nt 12–22), suggesting that the transposase contacts within the first half may be less specific or less important for function. Earlier mutation studies (Groenen and van de Putte, 1986) delineated both essential and non-essential nucleotides within MuA binding sites (highlighted in Fig. 3). Interestingly, within the Hin–Mu right end binding sites all the essential nucleotides are conserved, and these binding sites generated clear MuA footprints. In contrast, in each Hin–Mu left end binding site at least one essential nucleotide is apparently mutated, probably explaining the lack of detectable footprints in this end.

Transpososome assembly and transposase activity

Active Hin–Mu transpososomes could be formed, as evidenced by the activity assays, even though the initial DNA binding activity of MuA_Hin on Hin–Mu ends was below the limit of detection by DNase footprinting. This result underscores the importance of reaction conditions and shows that the lack of footprints in binding studies does not necessarily indicate a non-functional transposase.

In general, the amount of detected Mu transpososomes and similarly migrating Hin–Mu complexes correlated well with the corresponding strand transfer activities, indicating that the observed Hin–Mu complexes most likely represent genuine Hin–Mu transpososomes. A substantial level of MuA_Hin activity was obtained with the precut Hin–Mu substrate, indicating that MuA_Hin is capable of assembling a transpososome and has retained the strand transfer activity of its catalytic core. However, the MuA_Hin activity with uncut Hin–Mu substrate was almost undetectable. As the corresponding MuA-catalyzed reactions on both precut and uncut Mu substrates yielded similar levels of activity, the above results suggest that MuA_Hin is compromised in one or several steps preceding strand transfer. In Mu system, uncut substrates with several non-complementary nucleotides in the transposon DNA flank (i.e., as in the frayed substrate) are among the most proficient in transpososome assembly and subsequent catalytic steps (Savilahti et al., 1995). In the Hin–Mu system, utilization of the frayed substrate was not able to rescue the above-mentioned defect, suggesting that probably the donor cleavage activity is affected. However, some other, not mutually exclusive possibilities exist, and they include assembly defects and/or decreased stability properties of transpososomes.

MuA was able to promote efficient assembly of active transpososomes on Hin–Mu ends both with uncut and precut substrates. These data are in accordance with the footprinting results and show interchangeability between the components of the two systems. In addition, the data indicate structural conservation in the reaction mechanisms of the two systems as well as evolutionary close relatedness.

Is Hin–Mu a live virus?

Hin–Mu is similar to phage Mu with regard to DNA sequence and genomic organization (Morgan et al., 2001). While the sequence conservation is strikingly non-uniform along the genome with a total lack of identifiable similarity
in certain regions, similar regions predominate and are amenable to reasonable comparisons. For example, the degree of similarity (at the amino acid level) ranges from 22% to 60% among identifiable protein homologues (Morgan et al., 2001). The genomic conservation and the degree of similarity (at the amino acid level) ranges from amenable to reasonable comparisons. For example, the in certain regions, similar regions predominate and are amenable to reasonable comparisons. For example, the degree of similarity (at the amino acid level) ranges from 22% to 60% among identifiable protein homologues (Morgan et al., 2001). The genomic conservation and the retained catalytic activity of Hin–Mu transposition machinery strongly suggest an evolutionarily recent integration. But can Hin–Mu still function as a virus? Our recent results indicate that such is probably not the case as virus plaques could not be generated by plating all the other H. influenzae strains in this study with supernatants of the H. influenzae Rd strain even after treatment with chloroform (our unpublished results). Accordingly, we were not able to directly investigate potentially interesting phage immunity characteristics of Hin–Mu in comparison to those of its relatives, as this would have required isolation of infective viruses.

Other Mu-like prophages

Apparently full-length Mu-like prophages have recently been identified at least in the genomes of H. influenzae, Neisseria meningitidis, Deinococcus radiodurans, and Escherichia coli (Morgan et al., 2001 and references therein). In addition, partial Mu-like sequences or defective prophages exist, for example, in Vibrio, Shigella, Campylobacter, Haemophilus, Pasteurella, and Yersinia (Allison et al., 2002; Morgan et al., 2001 and references therein). It is likely that the apparent host range of Mu-like phages will increase as genome sequencing projects proceed, and the prevalence and distribution of these transposons in different species of bacteria will be clarified over time. Accordingly, functional comparisons between related systems will become possible and may generate a wealth of new information regarding the genetic mechanisms and evolution of these elements.

Materials and methods

Bacterial strains

E. coli strains DH5α (standard cloning host), DH10B, and BL21(DE3) were from Invitrogen. Nine H. influenzae clinical strains (T2369, T2370, T2377, T2408, T2409, T5896, T7386, T8217, T11457) and two culture collection strains (ATCC 33391, ATCC 49247) were obtained from Dr. M. Vaara (National Public Health Institute, Helsinki, Finland), and they were cultured as described (Virkola et al., 1996).

DNA and oligonucleotides

Plasmid pUC19 was from New England BioLabs, pBluescript SK+ from Stratagene, and pET3c from Novagen. Plasmid pMK586 (Mizuuchi et al., 1991) was obtained from K. Mizuuchi. H. influenzae Rd (ATCC 519070) chromosomal DNA was from ATCC. Chromosomal DNA of other H. influenzae strains was isolated using standard techniques (Sambrook et al., 1989). Primer oligonucleotides were from commercial sources (sequences available on request). Strands of the Mu end-specific DNA segments (Fig. 5B) were synthesized by the Keck Oligonucleotide Synthesis Facility at Yale University and purified by urea-polyacrylamide gel electrophoresis (Sambrook et al., 1989) prior to use.

Reagents, enzymes, and DNA techniques

Streptavidin-coated beads were from Roche, [γ-33P]ATP (1000–3000 Ci/mmol) from Amersham, and agaroses from Cambrex. Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase (PNK), DNA polymerase I Large fragment (Klenow), and Vent DNA polymerase were from New England BioLabs. DNA polymerase preparations Dynazyme II and Dynazyme EXT as well as deoxynucleotides were from Finnzymes. RQ1 RNase-Free DNase I was from Promega. All the commercial enzymes were used under the reaction conditions recommended by the suppliers. Standard DNA techniques were performed as described (Sambrook et al., 1989). DNA sequencing was performed at the sequencing service unit of the Institute of Biotechnology, University of Helsinki. Plasmids were isolated, and PCR fragments were purified using appropriate kits from Qiagen. When required, DNA single strands were radiolabeled at the 5’-end with [γ-33P]ATP using PNK. To generate radiolabeled transposon end-specific DNA fragments (Fig. 5A), the 33P-labeled strands were purified and annealed with unlabeled strands as described (Savilahti et al., 1995).

Purification of transposase proteins

MuA transposase was overexpressed and purified as described (Haapa et al., 1999a). The gene encoding MuAHin was amplified by PCR using Vent DNA polymerase with the primer pair HSP36/HSP37 and chromosomal DNA of H. influenzae Rd as a template. The resulting PCR fragment was cleaved with NdeI and BamHI and cloned into plasmid pET3c cleaved with the same two enzymes. The DNA sequence of the construct (pALH1) was verified by sequencing, and the plasmid was transformed into the E. coli strain BL21(DE3) for MuAHin expression. The expression and purification of MuAHin primarily followed that of MuA except that the hydroxyapatite column was from BioRad (BioGel HT Gel) and 200 mM KCl was used in the loading buffer of chromatographic steps.

Southern analysis

PstI-digested genomic DNA of the 12 H. influenzae strains was electrophoresed on a 0.8% agarose gel and
transferred to a Hybond-N+ membrane (Amersham) using the alkaline method (Sambrook et al., 1989). Southern hybridization (Sambrook et al., 1989) was carried out using the Hin–Mu-specific NdeI–BamHI restriction fragment of pALH1 as a γ-32P-labeled probe. In addition to Hin–Mu prophage-containing H. influenzae Rd DNA, the probe hybridized to DNA of the strains T7386 and T2377 (data not shown), suggesting that also these two strains may contain Mu-like DNA (length and genomic location unknown).

Identification of Hin–Mu prophage termini

A PCR fragment was generated by amplification using Dynazyme EXT and the primer pair HSP180/HSP177 with genomic DNA (100 ng) from H. influenzae strain T7386 as a template (prophage absent in the amplified locus). The sequence of the resulting 1.2-kb PCR product was determined, and the sequence was aligned with the corresponding sequence from H. influenzae Rd strain (prophage present) to localize the prophage termini.

Cloning of Hin–Mu prophage ends

(i) For the L-end connected to its chromosomal flank, a PCR fragment was amplified using Vent DNA polymerase with the primer pair HSP181/HSP182 and H. influenzae Rd genomic DNA (100 ng) as a template (prophage absent in the amplified locus). The PCR product was cleaved with BamHI and HindIII and cloned into plasmid pBluescript SK+ as a template for PCR amplification by Dynazyme Ext DNA polymerase with a single Hin–Mu end-specific primer HSP496. A 2.8-kb amplification product that corresponded to streptavidin-coated beads in MuA binding buffer (the buffer composition described in Craigie et al., 1984). Binding reactions (20 μl) were incubated for 30 min on ice. Subsequently, DNase I (0.2 U) was added, and the reactions were incubated at room temperature for 5 min. Following termination and washing steps, the beads were suspended in 8 μl loading dye (95% formamide, 10 mM EDTA, 0.03% bromophenol blue, 0.03% xylene cyanol). The samples were incubated at 100 °C for 5 min and analyzed (along with appropriate sequencing markers) by denaturing 7 M urea-10% polyacrylamide gel electrophoresis in 1 × TBE buffer (Sambrook et al., 1989). The gel was dried at 80 °C onto a paper support and visualized by autoradiography using a Fuji BAS 2000 phosphorimager.

In vitro transposition reactions

Standard transposition reactions (25 μl) contained 1 pmol radioactively labeled transposon end DNA fragment (donor DNA), 250 ng pUC19 (target DNA), 200 ng transposase, 25 mM Tris–HCl pH 8.0, 100 μg/ml BSA, 15% (w/v) glycerol, 0.05% (w/v) Triton X-100, 15% (w/v) DMSO, 126 mM NaCl, and 10 mM MgCl2. For some reactions, in order to generate a higher number of complexes, the concentrations of both donor DNA and transposase were increased stoichiometrically. For example, a 4-fold concentrated reaction utilized 4 pmol of donor DNA and 800 ng of transposase. Reactions were carried out at 30 °C for 1 h and terminated by freezing in liquid nitrogen. Following addition of 0.2 vol loading dye (0.1% bromophenol blue, 2.5% SDS, 50 mM EDTA, 25% Ficoll 400), reaction products were analyzed by electrophoresis on a native 3.5% SeaKem HGT agarose gel for 1.5 h at 5.3 V/cm in 1 × TAE buffer (Sambrook et al., 1989). To detect stable protein–DNA complexes, transposition reactions were performed as above except the target DNA was omitted, and some reactions were performed without Mg2+ and incubated on ice instead of 30 °C (indicated in Fig. 5C). Following addition of Ficoll 400 (0.2 vol, 25% Pharmacia), samples were analyzed at room temperature by electrophoresis on a native 3.5% NuSieve 3:1 agarose gel containing heparin (87.5 μg/ml) and BSA (87.5 μg/ml) for 2 h at 5.3 V/cm in 1 × TBE buffer with buffer circulation. Gels were dried onto DEAE paper (DE81, Whatman) and visualized as above.

Analysis of donor-target DNA junctions of transposition reaction products

An 8-fold concentrated transposition reaction was diluted 1:10 with water, and 1 μl of the diluted reaction was used as a template for PCR amplification by Dynazyme Ext DNA polymerase with a single Hin–Mu end-specific primer HSP496. A 2.8-kb amplification product that corresponded to transposon end segments attached to pUC19 was isolated.
via preparative agarose gel electrophoresis, treated with Klenow in the presence of dNTPs to generate blunt ends (Sambrook et al., 1989), and ligated by T4 DNA ligase into a Klenow-treated 1.1 kb chloramphenicol acetyl transferase gene-containing BamHI fragment (cat) of artificial transposon cat-Mu (Haapa et al., 1999b). Ligation products were electrotitransformed (Lamborg et al., 2002) into E. coli DH10B cells and selected on LB plates containing ampicillin (100 μg/ml) and chloramphenicol (10 μg/ml). Plasmid DNA was isolated from clones resistant to both antibiotics, and the sequences of the transposon DNA–target DNA junctions were determined by using cat-specific primers HSP350 and HSP349.

Scoring the transposase binding site sequences

A consensus sequence for the 22-bp transposase binding site was first compiled using the six MuA binding sites of the right and left end of the Mu genome (Fig. 3). Each putative Hin–Mu transposase binding site was then scored with the following rules: (i) a match to a consensus comprising a single nucleotide (1.0 points), (ii) a match to a consensus comprising two alternate nucleotides (0.5 points). The total score for each site was calculated as a sum of scores from 22 individual nucleotide positions.

Acknowledgments

Sari Tynkkynen, Pirjo Rahkola, and Auli Saarinen are acknowledged for excellent technical assistance. Martti Vaara kindly provided the H. influenzae strains. We thank Anja Paatero and Maria Pajunen for critical reading of the manuscript. This research was funded by The Viiikki Graduate School in Biosciences grant to A.-H. S. and the Academy of Finland grants to A. L. and H. S.

References


