In Vitro Shuttle Mutagenesis
Using Engineered Mariner Transposons

Kelly A. Robinson, Sophie Goyard, and Stephen M. Beverley

Summary

Advances in our understanding of the protozoan parasite *Leishmania* have been facilitated by the development of molecular and genetic tools. One powerful approach for gene identification and analysis is transposon mutagenesis. This can be performed directly in vivo, but often it is more convenient to generate transpositions in vitro for subsequent analysis in vivo, in a process termed "shuttle mutagenesis." The *Drosophila* element *mariner* is well suited for application by either route. Minimal *mariner* elements containing *cis*-acting elements required for transposition have been generated, which can be further modified to suit the needs of the experimenter. Additional genetic markers and/or reporters can be introduced, which are useful for procedures such as insertional mutagenesis, shotgun sequencing, or the generation of protein and transcriptional fusions for subsequent analysis. Active transposase can readily be generated following expression in *Escherichia coli*, and efficiencies of $10^{-7}$/target can be obtained, allowing the generation of large transposon insertion libraries suitable for subsequent screening in vivo. This chapter explains the steps necessary to purify active *Mos1* transposase and conduct an in vitro transposition reaction. We also discuss some of the considerations relevant to the design and application of functional *mariner* elements (donor plasmids) relevant to studies in *Leishmania* and other organisms.

Key Words: In vitro transposition; *Mos1* transposon; protozoan parasite; shuttle mutagenesis; *Tc1/mariner* transposon family.

1. Introduction

Protozoan parasites such as *Leishmania* are responsible for numerous illnesses that cause significant mortality and morbidity throughout the world (1). Genetic and genomic tools now available for the study of the parasite promise to increase greatly our understanding of how this parasite survives and causes disease and ultimately lead to improved methods for overcoming this disease by immunization or chemotherapy (2–5).
Two common tasks in parasite genetics are, first, the identification of genes mediating interesting functions, and second, dissection of the role, regulation, and localization of encoded proteins. A powerful tool suitable for this task in many organisms is transposon-based mutagenesis (6). This can be performed directly in vivo, whereby both the transposon and active transposase are introduced or expressed in the parasite, or in vitro, whereby transposition is performed in vitro and the products introduced into the parasite for subsequent analysis (Fig. 1). In vivo strategies are especially powerful when incorporated into forward genetic approaches, as mutants generated are simultaneously tagged by the transposon, which can then be used to recover the affected gene. Unfortunately, Leishmania is an asexual diploid in the laboratory, and for most loci, recovery of loss-of-function mutations requires at least two genetic events (7). However, this approach is widely used in haploid organisms, or ones in which homozygosity can be readily attained in some manner (for example, by sexual crossing).

In vivo transposition systems can be challenging to set up because of the constraints inherent in engineering transposase expression and controlling transposition. For many purposes, in vitro transposition is more convenient and as powerful. In a process termed “shuttle mutagenesis,” transflectable molecular constructs (for example, Leishmania DNA cloned in the shuttle Escherichia coli-Leishmania vector cLHYG [8]) are subjected to transposition in vitro, and then the population of independent insertions is scored for phenotypes following transfection back into Leishmania (Fig. 1). The transposon insertion library also can be used for rapid and systematic DNA sequencing if necessary, using primers situated within the transposon. We have found this approach especially useful in mapping the active gene within cosmids recovered in various functional genetic screens in Leishmania.

Beyond their role as insertional mutagens, transposons can be engineered to contain reporters such as the green fluorescent protein (GFP), β-galactosidase, β-glucosidase, or β-lactamase, or selectable markers such as NEO, HYG, or PHLEO, which mediate resistance to G418/geneticin, hygromycin B, or phleomycin/zeocin, respectively. Following transposition, activation of the reporter or marker can then be used to identify and/or select for transcribed or translated regions of the genome, a procedure commonly referred to as gene/protein “trapping.” By studying expression and/or localization of the reporter proteins, one can then conveniently (albeit indirectly) monitor gene expression and protein localization (9,10).

Several transposon systems have been engineered to the point that they are readily incorporated into shuttle transposon mutagenesis strategies; these include Tn7 (11), Tn5 (12), Ty1 (13), Mu (14), and several Tc1/mariner family elements (15–17). Relevant factors include the availability, cost, and/or ease of purifica-
Fig. 1. Transposon mutagenesis strategies. (A) Shuttle mutagenesis begins with in vitro transposition into target DNA (plasmid or cosmid) to create an insertion library. The transposon library is then transfected into *Leishmania*, and recovered by selection on the drug-resistance marker, found on the transposon, in various ways. (B) In vivo mutagenesis requires the establishment of an active transposition system in the parasite itself. This could be accomplished in several ways; stable expression of transposase and stable introduction of the transposon, stable expression of transposase and stable introduction of the transposon, stable expression of transposase and stable introduction of the transposon, and stable introduction of donor transposons. Alternatively, one could form a transposase–transposon complex in vivo, and then introduce this for subsequent transposition in vitro. This has not yet been demonstrated in the *mariner* system, but it works well with transposon Tn5 in various eukaryotes (35,36). (Reprinted with permission from ref. 37.)
The Drosophila element *Mosl* is a member of the mariner/Tc1 family, which occurs in most kingdoms of living organisms (18, 19). Typically, mariner/Tc1 elements are small, encoding only the transposase and cis-acting elements required for transposition, such as the flanking IRs (20, 21). Transposition occurs through a cut-and-paste mechanism (21), in which recognition of the IRs by the transposase results in excision of the donor element. This is subsequently inserted into a TA dinucleotide of the target molecule, and accompanied by duplication of the TA flanking the insertion site (22). In vitro experiments have shown that the transposition reaction requires only transposase and transposon cis-elements, without the need for cellular factors (15, 17, 21, 23, 24). For Drosophila *Mosl*, the cis-acting elements required for transposition include the 28 bp 5' and 3' IRs, along with some internal nucleotides (no greater than 38 and 5 additional internal nucleotides on the 5'- and 3'-sides, respectively;17, 23). Although here we focus on in vitro applications, it is notable that the mariner system has been shown to function in vivo in a variety of different organisms including *Leishmania*, insects, and vertebrates (25–29). Most of the mariner transposon derivatives described below can be used in vivo as well, in any species.

Our understanding of the mechanism of *Mosl* transposition has lead to the development of a minimal, “empty” transposon donor, pELHY6Δ-0 (Fig. 2), into which various transposon “cargos” have been inserted previously (23).

---

**Fig. 2. (Opposite page) Mosl vectors and the in vitro transposase reaction. (A)** The “empty” donor plasmid pELHY6Δ-0 contains the minimal cis-element (open arrow heads) with the 5’- and 3’-IRs and some internal nucleotides (see Fig. 3A). The vector contains the *E. coli* OriR6K origin (striped box) for propagation in a λpir+ strain and a HYG-selectable marker (internal black line); this particular marker contains a *Leishmania* splice acceptor site (AG) for expression in *Leishmania* and an *E. coli* promoter (black arrow). For specific uses, one can insert various “cargo” within the IRs, at the unique *MslI*, XbaI, or *SbfI* sites (see Fig. 3A). In this figure, the donor element is pELHY6Δ-/GEP3/, created by insertion of the /GEP3/ (see Table 1). For simplicity, only the *E. coli* PHLEO-resistance marker is shown in /GEP3/. **(B)** The basic in vitro transposase reaction contains a donor element, target DNA, and transposase (shaded area). In this example, the transposon target is a cosmid DNA (bottom left in Panel B), which contains an *E. coli* ampicillin-resistance marker (Ap) and an OriC origin of replication (open box). After in vitro transposition of the donor plasmid into the TA dinucleotide indicated on the cosmid, the DNA is transformed into a λpir- strain (such as DH10B). Donor plasmids bearing the OriR6K origin of replication cannot replicate in such strains and are selectively lost. Bacterial transformants are plated on LB medium containing Ap/PHLEO to select for transposition. Target DNAs containing transpositions can then be transfected into *Leishmania*, and fusion proteins identified by selecting or screening for PHLEO or GFP expression, respectively.
Insertion of /GFP3/ ‘cargo’ between Mosl IRs

empty ‘mini-Mosl’ vector with E. coli marker (Leishmania marker shown is optional)

pELHY6A-0 (propagated in pir+ E. coli)
New cargos may be designed and rapidly introduced into this vector as desired by the experimenter. Various transposons have been created previously, and potential applications such as insertional mutagenesis and gene trapping are summarized in Table 1. A typical in vitro Mosl transposition reaction consists of the donor plasmid (e.g., pGEP31, Fig. 2, Table 1), target DNA (e.g., a cLHYG-based cosmid; Fig. 2), and purified transposase. The properties and requirements for these elements are discussed below.

The donor plasmid pELHY6A-0 contains outside of the minimal transposon a bacterial/Leishmania-selectable marker (HYG here) and origin of replication (OriR6K) (Fig. 2). For propagation in E. coli, the OriR6K origin requires the pir gene product, which is provided by the use of appropriate E. coli host strains when growing this plasmid (often harboring a λpir lysogen). This allows one to select against the donor plasmid following transformation of the in vitro transposition reaction mix into pir- E. coli (this comprises virtually all common E. coli recipients). In this particular donor plasmid, the HYG marker is bifunctional, designed to function in E. coli because of the inclusion of a bacterial promoter, as well as Leishmania because of the presence of a parasite trans-splice acceptor site. Thus, this plasmid potentially can be used for in vitro transposition (E. coli) as well as in vivo transposition (Leishmania). Note that the signals for replication, transcription, and/or mRNA processing differ considerably between E. coli and Leishmania. Briefly, E. coli markers require promoters and plasmids require origins of replication for episomal maintenance; in contrast, in Leishmania all that is required is a trans-splice acceptor site upstream of the marker open reading frame (ORF), as transcriptional and replication origin requirements are quite relaxed (30). These differences must be taken into account when designing new transposons, and it can also be used to the researcher's advantage in various ways.

As a target plasmid, most common laboratory plasmid, cosmid, and bacterial artificial chromosome (BAC) vectors can be used; the specific requirements are that the target DNA should not contain an OriR6K for replication. The target marker should not be the same as ones borne within the transposon or in its donor background (note that in the cLHYG example shown, the HYG gene lacks an E. coli promoter and thus does not confer resistance in bacteria). Transposition efficiencies are highest if the target DNA is supercoiled and its "quality" is high (17).

Mosl transposase is purified and stored as described below. Typically, transposons require Mg²⁺ for activity; however, Mn²⁺ can be used, as this relaxes the requirement for insertion into TA dinucleotides (17). Whereas some mariner family transposases show a phenomenon called overproduction inhibition (15), Mosl shows simple saturation kinetics, and increasing transposase yields increasing transpositions until a plateau is reached (17).
Table 1
Examples of *MosI* Transposons and Their Properties and Applications

<table>
<thead>
<tr>
<th>Transposon</th>
<th>Bacterial marker</th>
<th>Eukaryotic reporter</th>
<th>Transpositional</th>
<th>In vitro</th>
<th>In vivo</th>
<th>TIMLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>pELHY65-GFP*K</td>
<td>---</td>
<td>---</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>pELHY65K-PG</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>PHLEO GUS</td>
<td>PHLEO GUS</td>
<td>+</td>
</tr>
<tr>
<td>pELHY65-GFP3*</td>
<td>---</td>
<td>PHLEO</td>
<td>PHLEO GFP</td>
<td>GPP</td>
<td>PHLEO</td>
<td>+</td>
</tr>
<tr>
<td>pELHY65-GFP2/</td>
<td>PHLEO</td>
<td>PHLEO</td>
<td>GPP</td>
<td>PHLEO</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>pELHY65-neo+ELSAT</td>
<td>SAT</td>
<td>NEO</td>
<td>SAT</td>
<td>NEO SAT</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>pELHY65-2x5</td>
<td>Km</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>+</td>
</tr>
</tbody>
</table>

All transposons listed can be used for insertional mutagenesis and sequencing. Symbols are as described: AG, *Leishmania* splice acceptor; SAT, nourseothricin resistance marker; Km, kanamycin resistance marker; GFP, modified green fluorescent protein; PHLEO, phleomycin/zeocin resistance marker; "" before or after a gene name indicates a gene lacking a start or stop codon respectively; "*" indicates an in-frame stop codon; black arrow, *E. coli* promoter; open triangle, 5'-IR; gray triangle, 3'-IR; oriC, oriC origin of replication; and BsrGI and SexAI, unique restriction sites used in TIMLI mutagenesis (described under Subheading 3.1.).
Following incubation, the in vitro transposition mix (which contains both donor and target plasmids as well as the desired transpositions) is transformed into pir- E. coli and plated on drugs that select for both the transposon (phleomycin in the example of /GEP3/ here) and the target (ampicillin here). Transposition efficiencies can be calculated by comparing platings on ampicillin alone vs ampicillin + phleomycin, and can approach 10⁻³/target.

The number of individual transpositions required depends on the particular application. Although mariner demands TA residues for insertion under standard conditions, these are sufficiently abundant even in the GC-rich Leishmania genome to provide plenty of potential target sites, and the requirement for TA can be relaxed if transposition is performed in the presence of Mn²⁺ (17). We have found that, for cosmid targets, several hundred independent insertions usually are sufficient for sequencing and inactivation of most potential target genes. For the recovery of specific gene fusions, larger libraries may be required, because one has the additional constraints of inserting into TAs in the appropriate strand and reading frame. For these purposes, 1000 independent insertions into a cosmid target should suffice. Note that with current in vitro transposition efficiencies and Leishmania transfection efficiencies (31), one may contemplate scoring libraries in excess of 10⁵ independent insertions.

This chapter describes how to express and purify active Mosl transposase and carry out in vitro transposition reactions using donor plasmids from the Mariner toolkit.

2. Materials

2.1. Vector

1. A suitable vector from the mariner toolkit (see Subheading 3.1, for details).

2.2. Expression of Mosl Transposase

1. E. coli strain expressing T7 polymerase (BLR [DE3]) from Novagen.
2. Vector expressing His₆-tagged Mosl transposase (pET19-Tpase, Beverley lab strain B4289; ref. 23).
3. 1 M IPTG (isopropyl-β-D-thiogalactopyranoside) stock solution.
4. Resuspension buffer: 20 mM Tris-HCl, pH 7.6, 2 mM MgCl₂, 25% sucrose, 0.6 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM benzamidine (BZA) and 1 mM dithiothreitol (DTT).
5. Liquid nitrogen or Sonicator.

2.3. Purification of Transposase

1. Lysis buffer: 20 mM Tris-HCl, pH 7.6, 4 mM EDTA, 200 mM NaCl, 1% deoxycholate, 1% nonylphenoxypolyethoxy ethanol (NP-40), 0.6 mM PMSF, 1 mM BZA, 1 mM DTT.
2. DNaseI.
3. 1 M MgCl₂.
4. Lysozyme.
5. Buffer A: 20 mM Tris-HCl, pH 8, 500 mM NaCl, 6 M guanidine-HCl, 1% NP-40, 70 mM imidazole.
6. Wash buffer: 50 mM NaH₂PO₄, 300 mM NaCl, 100 mM imidazole, adjusted to pH 8.0 using NaOH.
7. Elution buffer: 50 mM NaH₂PO₄, 300 mM NaCl, 500 mM imidazole, adjusted to pH 8.0 using NaOH.
8. Ni-NTA agarose (Qiagen, cat. no. 30210).
9. Purification column (Qiagen, cat. no. 34964).
10. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel.
11. Dialysis slide (Slide-A-Lyzer® dialysis cassette, 10,000-MW cutoff; Pierce cat. no. 66425).
12. Dialysis buffer A: 10% glycerol, 25 mM Tris-HCl, pH 7.6, 50 mM NaCl, 5 mM MgCl₂, 2 mM DTT.
13. Dialysis buffer B: 10% glycerol, 25 mM Tris-HCl, pH 7.6, 50 mM NaCl, 5 mM MgCl₂, 0.5 mM DTT.
14. 100% Glycerol.

2.4. In Vitro Transposition Assay
1. 10X transposition buffer: 250 mM HEPES, pH 7.9, 10 mM DTT, 50 mM MgCl₂, 1 M NaCl.
2. 100% glycerol.
3. 10-mg/mL purified bovine serum albumin (BSA) (New England Biolabs).
4. Donor plasmid (see Subheading 3.1. for details).
5. Target DNA consists of any plasmid, cosmid, or BAC that contains a different selectable marker than that used on the donor plasmid and that does not contain an OriR6K origin of replication.
6. Transposase (see Subheading 3.2.).
7. Sterile distilled H₂O (sdH₂O).
8. Stop buffer: 50 mM Tris-HCl, pH 7.6, 0.5-mg/mL proteinase K, 10 mM EDTA, 250-mg/mL yeast tRNA.
9. 3 M sodium acetate.
10. 100% ethanol.
11. 70% ethanol.
13. pir- E. coli electrocompetent cells, such as DH10B.
15. 10 mM Tris, pH 7.5.
16. Ampicillin, hygromycin, nourseothricin (Dr. Walter Werner; WeBioAge@aol.com) and Zeocin (Invitrogen).
17. Luria Bertani (LB) medium: 10 g Bacto-tryptone, 5 g yeast extract, and 10 g NaCl per liter (with appropriate drug)/LB agar plates: 20 g of agar to 1 L of LB medium.
3. Methods

Here, we describe a range of vectors available for transposition using mariner and outline the steps involved in preparing the transposase enzyme and performing an in vitro transposition reaction.

3.1. The mariner Toolkit

Table 1 describes some of the mariner derivatives that have been developed and used successfully in our laboratory. Various applications and transposons can be envisaged, and the ones below provide some perspective on the factors relevant to their design and utilization. The transposons are described briefly below; many can be used for the recovery and/or selection of gene fusions in Leishmania and other organisms, as they lack species-specific regulatory elements (the GEP transposon series, for example). All transposons can be used for primer-island sequencing, and insertional inactivation. Most transposons contain autonomous bacterial selectable markers and can be used in the in vitro system, except for pELHY6TK-PG (thereby restricting it to in vivo applications). Because Leishmania uses a polycistronic transcriptional mechanism to generate mRNAs and relies heavily on posttranscriptional regulatory mechanisms to control protein expression, we have given the most attention to transposons that facilitate the recovery of protein fusions.

Transposons /GEP3*, /GEP3/, and /GEP2/ contain a GFP-PHLEO fusion protein (Table 1); the linker peptide between the GFP and PHLEO additionally functions as an E. coli promoter, and in bacteria this cassette confers phleomycin resistance constitutively (as required for use in the in vitro system). Importantly, the GFP lacks an initiating ATG codon, and thus in eukaryotes GFP-PHLEO expression can only be obtained following insertion of the /GEP transposons into an ORF expressed by the target DNA (Fig. 3). Such insertions can be selected for by phleomycin resistance (only in eukaryotes), or screened for by GFP expression. Note that phleomycin resistance can be affected by compartmentalization of the fusion protein; if the PHLEO protein domain is restricted to a compartment such as the glycosome, which is segregated from the nucleus (the site of action of phleomycin), resistance will be abrogated (32).

/GEP3* differs from /GEP3/ and /GEP2/ in that it contains a stop codon following the GFP-PHLEO fusion protein. Thus, the protein fusions recovered bear the N- but not C-terminus of the trapped protein (Fig. 4A,C). In /GEP3/ and /GEP2/ the stop codon has been eliminated so that an intact reading frame is maintained across the entire transposon, enabling the recovering of protein fusions that bear both the N- and C-terminus of the trapped protein; this type of transposon is referred to as a “sandwich” transposon (Fig. 4B,C). Because mariner elements must insert into TAs, which can occur in any reading frame,
Fig. 3. Nucleotide sequences of *Mosl* cis-element and pELHY6Δ/-GEP3/, pELHY6Δ/-GEP2/ chimeric genes. (A) The minimal *Mosl* element used in our work contains essential cis-elements consisting of the 5’ and 3’-IRs (shaded gray block arrows) and the internal 38 and 5 internal nucleotides (nonshaded capital letters). The six potential *Mosl* reading frames, from the flanking DNAs across the IRs, are shown with arrows (labeled 1–6), whereas start and stop codons in each frame are shown by M or X, respectively. Unique restriction sites found within the empty transposon that are suitable for the addition of “cargo” are shown. (B) Putative chimeric genes created by insertion of the /GEP3/ and /GEP2/ transposons into target TA dinucleotides in the third or second reading frame, respectively. The “/” symbol represents points of potential fusion of reading frames, in this case with the GEP ORFs and those of the target. (Reprinted with permission from ref. 23.)
Fig. 4. Diagram of potential translational fusions obtained with transposons /GEP3*/ and /GEP3/. Both transposons yield translational fusions when inserted in-frame into target ORFs that express a GFP-PHLEO resistance fusion protein domain (Table 1). Note that phleomycin resistance can be affected by compartmentalization of the fusion protein; if the PHLEO protein domain is restricted to a compartment such as the glycosome, which is segregated from the nucleus (the site of action of phleomycin), resistance will be abrogated (32). (A) The pELHY6Δ-/GEP3* transposon encodes a bifunctional GFP-PHLEO protein with a stop codon after the PHLEO domain. Thus, fusion proteins contain only N-terminal sequence information from the target ORF. (B) The pELHY6Δ-/GEP3/ transposon contains an ORE across the entire transposon (both IRs and the bifunctional GFP-PHLEO protein). (C) Comparison of the use of “terminator” vs “sandwich protein fusions. In this example, the properties of fusion proteins generated by /GEP3* and /GEP3/ are compared following in-frame insertion into a typical membrane surface protein, which bears an N-terminal signal peptide and C-terminal membrane anchor. With /GEP3* the C-terminal segment is lost, resulting in secretion of the fusion protein from the cell, whereas with /GEP3/ retention of the C-terminal segment results in the formation of a surface membrane-anchored fusion protein.

/GEP3/ differs from /GEP2/ in which frame can be trapped (because of the sequence of the 5’ mariner IRs, it is not possible to make a “/GEP1/” transposon for protein trapping; Fig. 3). Although all GEP transposons can be used to study translational regulation, the ability of sandwich transposon to retain both
In Vitro Mariner Transposition

N- and C-terminal sequences, which often contain important protein targeting information (for example, membrane-anchoring domains; Fig. 4C), is important for some purposes.

Transposon /NEO*ELSAT (Table 1) creates a translational fusion to the CEO selectable marker. It should be noted that the ability of the NEO protein to tolerate N-terminal fusions varies considerably amongst different protein targets, in contrast to GFP, PHLEO, and β-galactosidase, which are more permissive. An additional element in this transposon is the selectable marker SAT (streptothricin resistance), which contains both a Leishmania splice acceptor site and an E. coli promoter, allowing for selection for the transposon with SAT and protein fusions with NEO in Leishmania.

Transposon GFP*K (Table 1) can be used to generate GFP fusions in a manner similar to the /GEP transposons, as its GFP also lacks an ATG start codon. Additionally, it bears a rare restriction endonuclease (I-Ppol), which is helpful in mapping transposon insertion sites in large targets or the genome in vivo.

The transposon carried in pELHY6TK-PG (Table 1) contains a PHLEO-GUS translational fusion containing its own start codon. This transposon lacks a constitutive bacterial selectable marker, and contains an E. coli oriC replication origin; thus it cannot be used in the in vitro system, and can only be used in vivo. There, the oriC replication origin facilitates the recovery of candidate transpositions back from Leishmania into E. coli if desired. This transposon donor is carried on pELHY6TK, which is a modified version of pELHY6A-0; it additionally bears a conditionally negative selectable marker (herpes virus thymidine kinase) that is active in Leishmania (33).

Transposon /-2x5 (Table 1) was designed for transposon-mediated linker-insertional mutagenesis (TIMLI; ref. 34). It contains an E. coli kanamycin-resistance marker, flanked by a "symmetric" mariner element in which the 5'-IR was duplicated. Importantly, this IR contains two sites that occur relatively infrequently in Leishmania, SexAI and BsrGI (and most importantly should not occur in targets where they are to be used). In TIMLI mutagenesis, one first generates a large library of transposition events into the target. Then, this transposition pool is collected en masse, DNA prepared, digested with SexAI or BsrGI, diluted and self-ligated, and transformed back into E. coli. This yields excision of the transposon, leaving behind only an insertion of 12 or 18 nucleotides (encoding 4 or 6 amino acids, respectively). Thus one can generate a library of short-peptide insertions for subsequent functional analysis, such as the mapping of protein domains and activities.

### 3.2. Expression of Mos1 Transposase

1. Transform BLR (DE3) with plasmid expressing Mos1 transposase; plate on ampicillin (100 μg/mL).
2. The following day, pick a single colony and resuspend into 5 mL of LB containing 100 µg/mL ampicillin and incubate at 37°C overnight.
3. The next day, use the overnight bacterial culture to inoculate 100 mL of fresh LB medium. Incubate this culture at 37°C to an OD600 of 0.6.
4. Induce expression of transposase by adding IPTG to a final concentration of 1 mM. After 5 h of induction, harvest cells by centrifugation at 1303g for 10 min at 4°C.
5. Resuspend bacterial pellet in 0.5 mL of resuspension buffer and flash-freeze by dipping the tube in liquid nitrogen. Store at -80°C (see Note 1).

3.3. Purification of Mos1 Transposase

1. Thaw cells at room temperature.
2. Add 1 mg/mL of lysozyme and incubate for 5 min at room temperature. Next, add 1 mL of lysis buffer and incubate at room temperature for 15 min and then for 20 min after adding 60 µg of DNAseI and MgCl2 to 10 mM.
3. All subsequent steps are conducted at 4°C.
4. Typically, most of the expressed transposase is insoluble and forms inclusion bodies. Pellet inclusion bodies at 14,000g in a microcentrifuge and wash three times with 1 mL of 100 mM Tris-HCl, pH 7.6. Resuspend inclusion bodies to a final volume of 4 mL in buffer A. Add 1 mL of 50 % Ni-NTA agarose and gently shake solution for 1 h.
5. Load mixture onto column (Qiagen) and collect flow-through (FT).
6. Wash column twice with 4 mL of wash buffer and collect wash fractions.
7. Elute four times with 0.5 mL of Elution buffer and collect 0.5-mL fractions for analysis later. The eluate contains the transposase protein.
8. Run 20 µL of eluted protein from each of the four tubes on an SDS-PAGE gel. On a 12% SDS-PAGE gel the purified His-tagged Mos1 transposase runs at approx 50 kDa (see Note 2). Also load an aliquot of the noninduced culture, the column FT, and the wash FT.
9. Pool the fractions containing the most transposase. Place this solution of transposase into a dialysis slide, ensuring not to overfill. Conduct dialysis in 1 L of dialysis buffer A for 6–8 h at 4°C. Replace dialysis buffer with 1 L of dialysis buffer B and incubate overnight at 4°C (see Note 3).
10. Centrifuge solution at 10,000g for 20 min at 4°C to transfer supernatant to a new microcentrifuge tube and discard precipitate.
11. Add glycerol to the transposase in solution to a final concentration of 50%. Store at -20°C.

3.4. In vitro Transposition

1. Set up a standard transposition reaction in a 0.6-mL microcentrifuge tube to a final volume of 20 µL.
   a. 2 µL of 10X transposition buffer.
   b. 2 µL of 100% glycerol (warm glycerol at 65°C to ease pipetting).
   c. 0.5 µL of BSA at 10 mg/mL.
**In Vitro Mariner Transposition**

d. 1 μL of donor plasmid (32 fmol).
e. 6 μL of target plasmid (10 fmol, see Note 4).
f. 5 μL of transposase (100 nM, see Note 5).
g. 3.5 μL of ddH₂O.

2. Incubate reaction at 30°C for 1 h to overnight.
3. Add 80 μL of stop buffer and incubate for 30 min at 37°C.
4. Add 100 μL of 25:24:1 phenylchloroform-isooamyl alcohol and vortex. Separate phases by centrifugation at 14,000g for 15 min. Remove approx 90 μL of the upper layer into a 1-mL microcentrifuge tube. Add 10 μL of 3 M sodium acetate and 250 μL of EtOH and incubate at -80°C for at least 1 h.
5. Precipitate DNA by centrifugation at 14,000g for 30 min at 4°C. Wash precipitated DNA with 1 mL of 70% EtOH and centrifuge at 14,000g for 15 min at 4°C. Resuspend the pellet in 10 μL of 10 mM Tris, pH 7.5.
6. Electroporate 2 μL of the purified transposition reaction into DH10B electrocompeotent cells (see Note 6). Add 1 mL of LB medium and incubate for 1 h at 37°C. Plate transfecntants onto selective LB plates (see Notes 7 and 8) and incubate at 37°C overnight.
7. Depending on the purposes, colonies may be picked individually or large pools made for DNA preparations and subsequent analysis en masse, for example, following transfection into *Leishmania* or other organisms (see Note 9; ref. 31).

**Notes**

1. Alternatively, one can sonicate cells using a microtip at 40–50% power for 20 bursts. The cells must be kept on ice during sonication. Afterward, one can proceed directly to Subheading 3.2., step 2, excluding the addition of lysozyme.
2. The predicted molecular weight of the His-tagged *Mosl* transposase is 43.6 kDa. The altered mobility of the transposase during electrophoresis may result from the presence of the basic histidine residues. Antibodies to the poly-histidine residues specifically recognize the 50-kDa band by Western blot hybridization.
3. The recovery of active, properly folded Tpase is very sensitive to the refolding conditions. Previous work has demonstrated that rapid dilution at low pH or dialysis of detergent-solubilized proteins results in no active enzyme (17). Omission of the column purification step also results in no enzyme activity, likely because of the presence of an unknown inhibitory factor. However, rapid dilution at pH 8.0 or refolding on a column using a linear urea gradient (8–0 M urea) has been shown to yield Tpase activity (17,24). Thus, although the refolding protocol described in this article has worked effectively in this laboratory, ultimately the optimal refolding conditions needs to be determined qualitatively by each investigator. We have also been able to purify active transposase from the soluble fraction. Cells are sonicated in lysis buffer and 1 mg/mL of lysozyme and centrifuged at 14,000g for 10 min. The supernatant is transferred to a 15-mL Falcon tube and the volume is brought up to 4 mL in buffer A. The protocol then continues as described above (Subheading 3.3., steps 4–11), eliminating the dialysis step (Subheading 3.3., steps 9 and 10). Our experience is that there is a lot of variability in the yield of soluble...
active protein and the efficiency of refolding; thus, it is essential that investigators carefully explore these parameters in their own laboratory to find what works best.

4. The quality of DNA is very important to transposition efficiencies. The preparation of donor and recipient DNA containing a high proportion of supercoiled DNA results in high transposition efficiencies. Qiagen midi preparations are generally suitable for this goal. However, when preparing cosmid DNA, one should take special care to avoid shearing the DNA. Transposition efficiency reaches a maximum at around 150 ng of donor plasmid.

5. The concentration of transposase is determined by the micro-BCA method (Pierce). Concentration may also be determined by UV absorbance ($\varepsilon_{280} = 76,989 \text{ M}^{-1} \text{ cm}^{-1}$); note that these measurements lead to differences in estimation of transposase concentration by a factor of 3 (17). Different batches of purified transposase can have different transposition efficiencies. This difference probably arises from batches of MosI transposase containing different amounts of correctly folded transposase. Therefore, each batch should be tested before conducting large-scale transposition reactions. The transposition efficiency reaches a plateau at a concentration of around 100 nM of transposase and remains at this level at higher concentrations (17).

6. The use of high-efficiency electrocompetent cells works best. Invitrogen GeneHogs® electrocompetent cells can yield $1 \times 10^{10}$ transformants per microgram of pUC vector.

7. Plate 10 µL of a 1/100 dilution of the transformed bacteria onto medium containing the appropriate antibiotic for the resistance marker found on the recipient plasmid (e.g., ampicillin [Ap] at 100 µg/mL, Fig. 2). The number of Ap$^R$ colonies multiplied by the dilution factor (in this case by 10,000) is the transformation efficiency. The remaining 990 µL of cells are plated onto medium containing antibiotics for the resistance markers found on the recipient plasmid and the transposon cassette (such as Ap and Phleo in the example shown in Fig. 2; 100 µg/mL and 50 µg/mL, respectively). Transposition efficiency is determined by dividing the number of Ap$^R$, Phleo$^R$ colonies by the transformation efficiency (obtained in previous step). Control transposition efficiencies should range from $10^{-4}$ to $10^{-3}$.

8. One can also estimate the transposition efficiency of a vector such as pELHY6TK-PG, which contains no bacterial-selectable marker. To accomplish this, a negative-selectable marker such as the product of the ccdB (control of cell death) gene placed in the bacterial plasmid, which additionally contains a positive-selectable marker (Kanamycin, Km$^R$). Examples of this are the pZERO system available from Invitrogen. Transformation of this plasmid into bacteria lacking the gyrase gene results in Km$^R$ colonies. However, cell death occurs when transfected into strains containing gyrase like TOP10. An in vitro transposition reaction is performed with equimolar amounts of this plasmid (Km$^R$) and a standard plasmid target (Chloramphenicol, Cm$^R$) in a strain lacking gyrase. The transposition reaction is subsequently transformed into TOP10 bacteria. One half of the transformation is plated onto Lb + chloramphenicol plates (30 µg/mL) and the other half is plated onto Lb + kanamycin plates (50 µg/mL). The number of colonies obtained
when plated on Kanamycin represents transposition events into the ccdB gene. The ratio of Km\(^{R}\)/Cm\(^{R}\) represents the transposition efficiency of this transposon.

9. If one is planning transfections into *Leishmania*, it is critical first to determine the sensitivity of your specific strain under the exact circumstances you plan to use. Drug sensitivities vary greatly among different strains and species, in different media, and interactions can occur if two drugs are used simultaneously. First determine the EC\(_{50}\) in liquid medium; then, carry out a “mock” transfection followed by inoculation or plating onto media containing drug concentrations ranging upward from three to four times that of the liquid media EC\(_{50}\). Because high drug concentrations inhibit the recovery of bona fide transfectants, the goal is to identify the minimal drug concentration that kills untransfected/drug-sensitive cells. In *Leishmania*, drug concentrations typically are 15–30 \(\mu\)g/mL for G418/geneticin, 25–40 \(\mu\)g/mL for phleomycin, 15–30 \(\mu\)g/mL for hygromycin B, and 100 \(\mu\)g/mL for nourseothricin, but exceptions are common.

Acknowledgments

The authors thank Lon-Fye Lye for assistance with the transposon protocols and Rosa Reguera for comments on the manuscript. This work is supported by a National Institutes of Health (NIH) Research Fellowship (AI50339) to K. A. Robinson and NIH grant AI29646 to S. M. Beverley.

References


