First biochemical steps on bacterial transposition pathways

Catherine Guynet[1]*, Emilien Nicolas[2], Bao Ton-Hoang[1], Jean-Yves Bouet[1] and Bernard Hallet[3]*

[1]Laboratoire de Microbiologie et Génétique Moléculaires, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université de Toulouse, UPS, F-31000 Toulouse, France.

[2]Fast-track Diagnostics, 29 rue Henri Koch, L-4354 Esch-sur-Alzette, Luxembourg.

[3]Institut des Sciences de la Vie, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.

*Corresponding authors: guynet@ibcg.biotoul.fr and bernard.hallet@uclouvain.be

Abstract

Transposons are found in a wide variety of forms throughout the prokaryotic world where they actively contribute to the adaptive strategies of bacterial communities and hence, to the continuous emergence of new multi-resistant pathogens. Contrasting with their biological and societal impact, only a few bacterial transposons have been the subject of detailed molecular studies. In this chapter, we propose a set of reliable biochemical methods as a primary route for studying new transposition mechanisms. These methods include (i) a straightforward approach termed "thermal shift induction" to produce the transposase in a soluble and properly folded configuration prior to its purification, (ii) an adaptation of classical electrophoretic mobility shift assays (EMSA) combined to fluorescently labeled DNA substrates to determine the DNA content of different complexes assembled by the transposase and (iii) a highly sensitive "in gel" DNA foot-printing assay to further characterize these complexes at the base pair resolution level. A combination of these approaches was recently applied to decipher the molecular organization of key intermediates in the Tn*3*-family transposition pathway, a mechanism that has long been refractory to biochemical studies.

Key Words

Transposase purification, transpososome, DNA-protein complex, thermal-shift induction, DNA binding, EMSA, fluorescent DNA labeling, nucleoprotein complexes stoichiometry, in gel DNA footprinting

Running title

Basic methods for transpososome biochemistry

1. Introduction

Transposons inhabit the genomes of all prokaryotes where they constitute the driving force of the "mobilome", allowing bacterial communities to face environmental challenges by exchanging and re-assorting phenotypically useful genetic traits. They are notably responsible for the dissemination of antimicrobial resistance genes and the emergence of an ever-growing number of multi-resistant pathogens, which is now regarded as one of the most serious threat for human health worldwide (e.g., see http://www.who.int/medicines/publications/WHO-

PPL-Short_Summary_25Feb-ET_NM_WHO.pdf). Therefore, understanding the molecular mechanisms that mediate and regulate the movement of transposons is crucial to fully assess their biological and societal impact, and possibly to find new strategies to control their activity.

Transposition is a multi-step process, during which a transposon-encoded specialized DNA recombinase, the "transposase", catalyzes DNA breakage and rejoining reactions required for transposon mobilization. Although transposases belong to structurally and biochemically distinct families or super-families of proteins, the transposition pathway generally goes through common successive steps (e.g. ***(1)***). The first step typically consists in specific binding of the transposase to the ends of the transposon. The next step involves the formation of a higher-order nucleoprotein complex termed "transpososome" in which the ends of the transposon are brought together usually through transposase oligomerization. Formation of this complex is critical to ensure that the chemical steps of transposition are catalyzed in an ordered and coordinated fashion and its architecture and dynamics are specific to each system. Transpososome formation and activity may be regulated by one or more host factors or by additional transposon-encoded accessory proteins. The precise nature of the DNA substrates (single or double stranded), the chemistry of the DNA cleavages and re-joining reactions catalyzed by the transposase (whether they occur by direct hydrolysis and trans-esterification or through the formation of a protein-DNA covalent intermediate), the structure of subsequent transposition intermediates and the degree of target specificity are defined by the biochemical family of the transposon and its mode of transposition.

In spite of this biochemical diversity and the large number of bacterial transposons currently listed in databases, only a handful of them have been characterized at the molecular level (Mu ***(2)***; IS*10* ***(3)***; Tn*7* ***(4)***; Tn*5* ***(5)***; IS*911* ***(6)***; Tn*552* ***(7)***; IS*1* ***(8)***; IS*608* ***(9)***; Tn*4430* ***(10)***). This is partly due to the complexity of the transposition reaction and the inherent difficulty to establish sensitive and quantitative biochemical assays recapitulating specific steps or the entire process in a controlled cell-free system. In this chapter, we describe three valuable methods for *in vitro* studies of transposition: (i) a thermal shift-induction method for transposase expression, (ii) a short–long Electrophoretic Mobility Shift Assay (EMSA) protocol coupled to differential fluorescent DNA labeling to characterize transpososome formation and composition, and (iii) an in-gel footprinting analysis of transposase-DNA interactions within the complexes. We have chosen these methods for their simplicity and the fact that they can conceptually be applied to decipher basic steps of any kind of transposition

mechanism.

A recurrent issue in *de novo* development of *in vitro* transposition assays is to purify sufficient amounts of active transposase to measure detectable levels of DNA binding, and to perform cleavage and strand transfer reactions under reconstituted conditions. This is most likely due to fact that transposase proteins are naturally expressed at a low level and exert their activity in the context of elaborate nucleoprotein complexes with strong conformational constrains and little or no turnover. In an extreme situation known as "*cis* preference", proper folding of the transposase is concomitant to its binding to the transposon ends, which regulates transposition by restricting the diffusion of the active transposase throughout the cell ***(11, 12)***. As a consequence, transposases are generally insoluble and have the tendency to aggregate when overexpressed, yielding poorly active and/or incompletely folded proteins, even when using standard denaturation/renaturation protocols (e.g., see ***(13)*** and references therein for discussion). To overcome these difficulties, we have set up a very simple protocol to optimize the overexpression of properly folded and soluble proteins in *Escherichia coli* cells (**Fig. 1**). This method, which we refer to as the "thermal shift induction" method, was shown to greatly facilitate purification of a number of "recalcitrant" proteins, including transposases (e.g. **(14-16)**). The method consists in growing cells carrying the transposase expression vector under heat-shock conditions (*i.e.*, at 42°C for *E. coli*) prior to induction. This allows to load the cells with heat-shock protein chaperones whose function is to interact with unfolded or misfolded polypeptides and to help them to adopt a proper conformation ***(17, 18)***. Expression of the transposase gene, which is under the control of an inducible promoter, is then induced at low growth temperature (*i.e.*, room T° or ~ 20°C) to benefit at the same time from low Brownian motion during protein production and from the activation of the protein quality control system by the heat-shock chaperones still present within the cells. Both conditions facilitate protein folding, thereby reducing aggregate formation and increasing the amount of transposase in the soluble fraction (**Fig. 1**). The protein can then be purified directly from cell extracts using appropriate protocols based, for example, on the use of small affinity tags or larger "carrier" tags, which may further increase protein solubility (e.g., for recent reviews: ***19-21).***

The two other methods described in this chapter are dedicated to determine how the transposase interacts with the transposon ends, and then how the resulting DNA-protein complexes develop to assemble the active transpososome. EMSA is one of the most widely used technique for experimental demonstration that the transposase binds specifically to the

transposon ends and to characterize its interaction with DNA (binding affinity, protein and DNA determinants of the interaction, stoichiometry and architecture of the transpososome) (e.g. ***(22-25)***). In classical EMSA, a labeled DNA fragment containing a specific sequence is incubated with the protein, and the protein-DNA complexes are subjected to polyacrylamide gel electrophoresis under native conditions ***(26, 27)***. In general, protein-DNA complexes migrate more slowly than the corresponding unbound DNA, generating upper-labeled bands within the gel. A limitation of this technique is that the electrophoretic mobility of a protein–DNA complex depends on many factors other than the size of the protein and DNA substrate, such as protein charge or DNA and protein/DNA conformational changes upon binding ***(28, 29)***. As a consequence, an observed mobility shift does not provide a straightforward determination of complexes composition. The binding of transposases to their ends often gives rise to more than one band shift corresponding to different complexes differing by the number of DNA substrates bound and/or the oligomerization state of the transposase in each complex. Here, we describe a "short–long" EMSA protocol coupled to differential fluorescent DNA labeling to facilitate determination of the transposase-DNA complexes composition (**Fig. 2**). In this assay, the transposase is incubated with a short and a long DNA fragment containing a transposon end, separately or together, in the same reaction. The fragments are separately labeled with either cyanidin5 (Cy5) or cyanidin3 (Cy3) fluorophores to assign them unambiguously to specific complexes. In the co-incubation reaction, retarded bands that are labeled with both fluorophores identify synaptic complexes in which the transposase is bound to two transposon ends (**Fig. 2**).

Details of the interactions between the transposase and DNA can be further investigated by in-gel 1,10-phenanthroline-copper (OP-Cu) footprinting analysis ***(30)***. This technique has been used to study the synaptic complexes inherent to different transposition mechanisms of the bacterial transposons IS*911*, IS*608* and Tn*4430* ***(10, 31, 32)***. The method that we propose here is an adaptation from Bouet et al., 1999 ***(33)***. It enables the determination of the DNA regions that are contacted by the transposase within each transposase-DNA complex previously identified by EMSA. OP-Cu has a chemical nuclease activity and interacts with DNA through the minor groove where it cleaves the phosphodiester backbone ***(34, 35)***. Resistance to cleavage, indicating that OP-Cu does not have access to the DNA, is caused by either protein steric hindrance or altered geometry of the DNA helix that narrows the minor groove. By contrast, enhanced OP-Cu cleavage results from altered DNA geometry that leads to a widened minor groove. Thus, the method not only indicates which DNA sequences are

protected by the transposase in the different complexes, it also provides additional information on the DNA distortions that are produced by the protein within or adjacent to its binding site.

In the assay, protein-DNA complexes are first separated by EMSA and the entire gel is treated with a solution containing OP-Cu to cleave the DNA. The complex bands and unbound DNA substrates are then excised from the gel, and after recovery of the treated DNA, OP-Cu cleavage products are resolved in a denaturing polyacrylamide gel. Comparison of the bound and unbound patterns with sequencing reaction products of the substrate can allow identification of protected sequences with single-nucleotide resolution. Compared to other footprinting reagents (such as DNase I, dimethyl sulfate or ferrous-EDTA), OP-Cu offers several advantages. Because of its small size and reactivity, it rapidly diffuses through the polyacrylamide gel matrix and cleaves DNA with relatively little sequence specificity, giving footprint patterns with a high precision. Also, OP-Cu cleaves DNA at physiological pH and temperature, and is not inhibited by glycerol, which is found in most protein storage buffers.

Although the ultimate fine molecular description of DNA-transposase complexes requires crystallographic, cryo-electron microscopy and/or NMR analyses, the techniques described in this chapter may provide important advances in the study of a particular transposition mechanism. They have been recently used to study the assembly of a transpososome of the Tn*3*-family, which has long been refractory to biochemical characterization due to the unusual large size and strong insolubility of the transposase ***(10)***.

2. Materials

Unless otherwise stated, reagents should be ‘molecular biology grade’ or better, and water should be Milli-Q® or any form of ultra-pure water with a sensitivity of 18.2 MΩ.cm. All 1.5 mL tubes used are DNA low binding grade. If not specified, buffers are stored at room temperature.

2.1 Thermal shift induction method for transposase purification

The standard procedure for protein purification includes the cloning of the gene of interest in an expression vector, its transformation into a host strain followed by induction of gene expression and then, protein purification. There are many molecular tools and protocols at hand for the production and purification of proteins in *E. coli*, including a large number of

expression vectors, host strains and purification methods (e.g. see ***(19, 21, 36).*** The method described here is based on the protocol used for Tn*4430* transposase ***(15)***. It is in principle compatible with any combination of host strain and expression vector and does not require any specific material. For batch cultivation, LB (Lysogeny broth) medium is commonly used, supplemented with appropriate selective antibiotics where necessary.

2.1.1 Expression plasmid

The catalog of available vectors for controlled gene expression is vast and diverse. Expression vectors are characterized mainly by the nature of the replicon (conferring the plasmid copy number), the type of promoter and the tags and/or fusion partners they contain (e.g., see ***19, 21*****).** The choice of the expression vector is linked to the choice of the host *E. coli* strain (**Subheading 2.1.2**) and the protocol that will be used for protein purification.

2.1.2 Bacterial strain

There is a variety of *E. coli* host strains for protein expression. The choice of a specific strain depends on the expression vector (see **Subheading 2.1.1**) and also on the nature of the transposase.

2.2 Short-long EMSA coupled to differential fluorescent DNA labeling

2.2.1 Fluorescently labeled DNA substrates preparation

1. Specific 5′ Cy5-labeled and 5′ Cy3-labeled oligonucleotides (see **Subheading 3.2.1.** and **Note 1**).

2. Hybridization buffer (2X): 20 mM Tris-HCl pH 8, 100 mM NaCl, and 2 mM Ethylenediaminetetraacetic acid (EDTA). This buffer is not required for PCR-amplified DNA substrates preparation.

3. Tris/EDTA (TE) buffer 1X: 10 mM Tris-HCl pH 8, 1mM EDTA.

4. Non-denaturing polyacrylamide gel: solutions and materials as described in **Subheading 2.2.3**.

2.2.2 Binding Reactions

1. Composition of DNA binding buffer depends on physicochemical properties of the

transposase protein. Our favorite 5X buffer is: 250 mM Tris-HCl pH 8, 250 mM NaCl, 25 mM EDTA pH 8, 5 mM DTT, and 50% (v/v) DMSO. Make fresh or aliquot and store at -20 °C (see **Note 2**).

2. Bovine serum albumin (BSA), 50 $mg.mL^{-1}$.

3. Poly-d[I-C], 1 mg/mL (see **Note 3**).

4. Protein dilution (or 'No') buffer: 50 mM Tris-HCl pH8, 500 mM NaCl, and 20% (v/v) glycerol. (see **Note 4**).

2.2.3 Non-Denaturing Polyacrylamide Gel

1. Dye-containing loading buffer 5X: 0.05% (w/v) orange G, 50% (w/v) sucrose. Make aliquots and store at – 20°C (see **Note 5**).

2. Electrophoresis buffer: Tris/borate/EDTA (TBE) buffer (10X): 890 mM Tris base, 890 mM boric acid, and 20 mM Ethylenediaminetetraacetic acid (EDTA) (see **Note 6**).

3. 40% (w/v) acrylamide:bis-acrylamide commercial solution, 19:1. Store at 4°C (see **Note 7**).

4. Ammonium persulfate (APS) (10%, w/v): dissolve 1g of APS in 10 mL water. Filter through a 22 μm membrane filter and store at 4°C.

5. Tetramethylethylenediamine (TEMED). Store at 4°C.

6. Special equipment: standard vertical apparatus for polyacrylamide gel electrophoresis, power supply and dedicated fluorescence scanner device.

2.3 In-gel footprinting analysis of DNA-transposases complexes

2.3.1 [^{32}P]-labeled DNA substrate preparation

1. Specific primers designed to PCR amplify the DNA region on which the footprint will be carried on.

2. T_4 polynucleotide kinase (T_4 PNK) and associated 10X buffer.

3. Radioisotope: ATP, [γ-^{32}P] (typical specific activity > 6.000 μCi/mmol) (see **Note 8**).

4. Items 3 to 5 in **Subheading 2.2.1**.

2.3.2 Transposase-DNA complexes separation by EMSA

All materials described in **Subheadings 2.2.2 and 2.2.3**.

2.3.3 In-gel footprinting with 1,10-phenanthroline-copper (OP-Cu)

1. Immersion buffer: 10 mM Tris-HCl pH 8.

2. Solution A: 0.45 mM $CuSO_4$, 2 mM 1,10-phenanthroline-copper (OP-Cu). Prepare just prior to use: OP solution (40 mM 1,10-phenanthroline): Dissolve 80 mg of 1,10-phenanthroline monohydrate in 10 mL of absolute ethanol; Cu^{2+} solution (9 mM $CuSO_4$): dissolve 72 mg of copper (II) sulfate anhydrous in 50 mL of water (see **Note 9**).

3. Solution B: 1:200 dilution of mercaptoproprionic acid in water. Add 100 µl of 3-Mercaptopropionic acid to 19,9 mL of water. Prepare just prior to use (see **Note 9**).

4. Solution C: dissolve 127 mg of neocuproine (2,9-dimethyl-1,10 phenanthroline) in 20 mL of ethanol. Prepare just prior to use (see **Note 9**).

5. Elution buffer for DNA recovery: 0.5 M ammonium acetate, 1 mM EDTA.

6. Phenol-chloroform-isoamyl alcohol mixture (49.5:49.5:1, v/v/v).

2.3.4 Preparation of the G + A sequencing ladder

1. Tris/EDTA (TE) buffer 1X: 10 mM Tris-HCl pH 8, 1mM EDTA.

2. 3 M sodium acetate (NaOAc).

3. 10 mg/ml tRNA.

4. DMS: ethanol solution: carefully dilute dimethyl sulphate (DMS) 10-fold in ethanol (see **Note 10**).

5. 1N NaOH.

6. Buffer A: 50 mM Sodium cacodylate pH 8.0, 150 mM KCl, 1 mM EDTA.
7. Buffer B: 20 mM KPO_4 pH 7.0, 1 mM EDTA.
8. Quench solution: 2 M NH_2OAc, 1 M β-mercaptoethanol, and 100 μg/ml of cold carrier DNA solution (e.g. sheared Salmon Sperm DNA solution with an average chain length ≤ 2000 bp). Store at 4°C.
9. Stop buffer: mix 1.8 mL of TE buffer, 0.2 mL of 3 M NaOAc and 4 □L of tRNA (10 mg/ml). Store at 4 °C.

2.3.5 Resolution of the footprint patterns in a denaturing polyacrylamide gel

1. Sequencing gel loading buffer: 98% formamide, 10 mM Na_2EDTA pH 8, 0.025% xylene cyanol, 0.025% Bromophenol Blue. Store at -20°C.

2. Solutions described in **items 2. to 5 in Subheading 2.2.3**.

3. Urea.

4. Whatman DE81 paper.

5. Special equipment: sequencing gel electrophoresis apparatus, power supply delivering high voltage (2500-3000V), gel dryer and all equipment for autoradiography.

3. Methods

3.1 Thermal shift induction method for transposase purification

1. On day 1, inoculate a 20 mL LB pre-culture, containing appropriate antibiotics with the bacterial strain harboring the transposase expression vector and incubate at 37°C under agitation overnight.

2. Pre-warm a flask containing 1 L of LB at 42°C.

3. On day 2, Dilute the pre-culture 50-fold in pre-warmed (42°C) fresh LB medium containing appropriate antibiotics and incubate at 42°C under agitation (200 rpm) until $OD_{600nm} \approx 0.6$ in order to express cellular heat-shock protein chaperones (Fig. 1) (See **Note 11**).

4. Rapidly chill the culture to 20°C by placing the flask on an ice bucket for ≈ 10 min with

vigorous shaking every 2 min (see **Note 12**).

5. Add the appropriate inductor agent to induce protein expression and incubate at 20°C for 4 hours (Fig. 1) (see **Note 13**).

6. Harvest cells by centrifugation (4000g) for 10 min at 4°C and wash the pellet once with 100 mL of the lysis buffer that will be used to re-suspend the pellet to start purification (see **Note 14**).

7. Re-centrifuge the suspension (4000 g) for 10 min at 4°C and store the pellet at -80°C until purification.

3.2 Short-long EMSA assay coupled to differential fluorescent DNA labeling

3.2.1 DNA substrates design

Double stranded DNA substrates may be generated either by PCR amplification (Fig. 2) (3.2.2) or by annealing specific and complementary synthetic oligonucleotides (3.2.3) (see **Note 15**). In both cases, the substrates will be designed to encompass the left or the right transposon end. In order to assign them unambiguously to specific complexes, the two fragments must have different lengths (see **Note 16**) and one substrate must be labeled with Cy5 fluorophore while the other with Cy3 fluorophores (see **Note 1**).

3.2.2 PCR-amplified DNA substrates preparation

1. PCR-amplify (standard procedures) the desired DNA fragment from a plasmid template containing the transposon ends using appropriate 5′ Cy5- labeled and 5′ Cy3-labeled oligonucleotides as primers (**Fig. 2**).

2. Prepare a 8% (w/v) non-denaturing (or native) polyacrylamide gel: for 50 mL, mix 10 mL of acrylamide:bis-acrylamide (19:1), 5 mL of TBE 10X buffer and water up to a volume of 50 mL. Add 300 µl of APS 10 %, 30 µl of TEMED and swirl the mixture gently. After adding APS and TEMED, pour the gel immediately (see **Notes 7** and **17**). Once the gel is polymerized, clean wells carefully with running buffer by using a syringe before loading the samples and pre-run the gel for 30 min at constant voltage (1V.cm^{-1}) at 4 °C (see **Note 18**).

3. Mix the PCR products with the appropriate amount of 5X loading buffer and load the samples on the gel after having rinsed the wells with a syringe and 1x TBE buffer (see **Note**

19).

4. Run the gel at room temperature with constant voltage ($1V.cm^{-1}$) until the marker dye has migrated the desired distance (with orange G-stained loading buffer, this is when the orange dye has reached the bottom of the gel, see **Note 5**).

5. Unmold the gel trying to keep it stuck on one of the glass plates, and locate the DNA of interest (see **Note 20**).

6. Use a clean sharp razor blade to cut out the band corresponding to the expected fragment and place the gel slice in a 1.5 mL micro centrifuge tube containing 350 µl of STE 1X buffer. Use a disposable tip to crush the polyacrylamide gel against the wall of the tube.

7. Transfer the tube to an agitation platform or a rotating wheel and incubate for ~3 to 16 h at room temperature for elution (see **Note 21**).

8. Centrifuge the sample at maximum speed in micro centrifuge and transfer the supernatant to a fresh 1,5 mL micro tube. Rinse out the polyacrylamide pellet with 150 µl of STE 1X buffer, centrifuge and transfer the supernatant again to the collect tube.

9. Pass the collected supernatant through a disposable spin filter column (e.g., Pall Laboratory Nanosep® columns) to remove any scraps of polyacrylamide.

10. Precipitate DNA from the flow-through with ethanol. Resuspend the purified substrates in 20 to 50 µl of TE buffer and store at -20°C until use.

3.2.3 Double-stranded synthetic oligonucleotides preparation

1. In a final volume of 40 µl, mix 500 picomoles of 5’ Cy3- or 5’ Cy5- labeled oligonucleotide with 750 picomoles of the complementary, unlabeled oligonucleotide in the 2X hybridization buffer (see **Note 22**).

2. Incubate the mix at 95 °C for 10 min and then slowly cool down to room temperature in order to conduct annealing under steady state equilibrium conditions (see **Note 23**).

3. Gel-purify the labeled substrates as in **Subheading 3.2.2**, from step 2 to 10.

3.2.4 Binding Reactions

1. Set up a binding reaction with 4 µL of 5 X binding buffer, 2 µL of BSA (50 $mg.mL^{-1}$), 1 µL of Poly d[I-C] (1 $mg.mL^{-1}$), 2 µL of fluorescently labeled DNA substrates (~ 50 nM), the purified protein (~ 50-100 nM final) or equivalent volume of protein dilution buffer, and water to adjust the volume to 20 µL. In order to visualize ternary protein-DNA complexes (or "sandwiches") in which the transposase binds to multiple DNA fragments, the binding reaction must be performed with the short and long substrate alone, and with both substrates mixed together at equimolar concentrations (Fig. 2B).

2. Incubate 30 min at 30°C (see **Note 24**).

3.2.5 Transposase-DNA complexes analysis by EMSA

1. Prepare a 4% (w/v) non-denaturing polyacrylamide gel (see **Notes 7** and **25**): for 50 mL, mix: 10 mL of 40% acrylamide:bis-acrylamide (19:1), 5 mL of TBE 10X and water up to a final volume of 50 mL. Add 300 µl of 10 % APS, 30 µl of TEMED and swirl the mixture gently. After adding APS and TEMED, pour the gel immediately. Once the gel is polymerized, clean the wells with a syringe (see **Note 17**).

2. Pre-run the gel for 1h prior loading the samples at constant voltage ($1V.cm^{-1}$) at 4°C (see **Note 18**).

3. Mix carefully the reaction mixtures with the 5 µl of 5X loading buffer and load the samples on the gel (see **Note 19**).

4. Run the gel at constant voltage ($1V.cm^{-1}$) at 4°C until the orange dye has reached the bottom of the gel (see **Note 5**).

5. Unmold the gel by taking care that it remains attached to the lower plate.

6. Scan the gel on a dedicated fluorescence scanner following the instructions of the manufacturer. Select appropriate parameters (excitation wavelength and filters) to selectively detect fluorescence emitted by Cy3 (550 nm, excitation; 570 nm, emission) and Cy5 (650 nm, excitation; 670 nm, emission) fluorophores, respectively. Choose the exposure time that gives the best signal-to-background ratio for each fluorophore. The individual scans selectively show specific complexes between the transposase and the Cy3- or Cy5-labeled substrates,

respectively. Higher-order complexes containing both DNA substrates are common to both profiles, and exhibit a mixed (i.e., yellow in Fig. 2B) fluorescence on the super-imposition of the Cy3 (green) and Cy5 (red) signals (Fig. 2B). In a typical "short-long" experiment performed with DNA substrates of different lengths, these complexes migrate at an intermediate position compared to the complexes assembled with the short and long substrate alone (Fig. 2B). The number and position of intermediate complexes reflect the different combinations where the short substrate becomes associated with the long substrate and reciprocally.

3.3 *In-gel* OP-Cu footprinting analysis of DNA-transposases complexes

3.3.1 [^{32}P]-labeled DNA substrate preparation.

Double stranded [^{32}P]-labeled DNA substrates are usually generated by PCR amplification, by labeling the 5' end of the primers (see **Note 8**).

1. Mix 10 pmol of DNA primer, 1 □L of 10X T_4 polynucleotide kinase buffer, 20 U of T_4 polynucleotide kinase, 50 μCi of [γ-^{32}P] ATP and water up to a final volume of 10 μL.

2. Incubate for 1h at 37 °C.

3. Heat inactivate by incubating 20 min at 65°C.

4. PCR-Amplify the desired DNA fragment, using the [^{32}P]-labeled primers (standard PCR procedures).

5. Gel-purify the labeled substrates as in **Subheading 3.2.2**, from step 2 to 10.

3.3.2 Transposase-DNA complexes separation by EMSA

Follow the steps in **Subheadings 3.2.4 and 3.2.5** as described above to separate transposase-DNA complexes from each other and from free DNA (see **Note 8**).

3.3.3 In-gel footprinting with 1,10-phenanthroline-copper (OP-Cu) (see **Note 9**)

1. Prepare solution A: mix 1 mL of the OP solution with 1 mL of the Cu^{2+} solution (the

mixture turns blue), add 18 mL of water and vortex.

2. Remove the one of the glass plates with care and immerse the wet gel (still attached to the lower plate) in 200 mL of 10 mM Tris–HCl pH 8 in a Pyrex dish (see **Note 26**).

3. Add 20 mL of solution A and shake the dish carefully.

4. Add 20 mL of solution B to initiate *In situ* digestion of the transposase–DNA complexes, shake quickly to distribute evenly and incubate 15 min at room temperature without shaking. The gel turns brownish.

5. Stop the reaction by the adding 20 mL of solution C and shake during 2 min. The gel turns yellowish.

3.3.4 Recovery of the treated DNA from the gel (see **Note 8**).

1. Remove all the liquid from the Pyrex dish using a glass pipette, rinse with water, and wrap the gel and plate with plastic wrap.

2. Localize free DNA and transposase-DNA complexes by autoradiography of the wet gel, excise them by cutting the corresponding gel band and place it in an 1.5 mL tube.

3. Recover the DNA from the gel by elution of the cut band in 300µl with crushing the gel in 0.5 M ammonium acetate, 1 mM EDTA. After ON elution at room temperature, centrifuge 10 min at 20.000g.

4. Extract DNA by phenol/chloroform followed by ethanol precipitation.

5. Resuspend the dried purified DNA in 5 µl of sequencing-gel loading buffer.

3.3.5 Preparation of the G + A sequencing ladder

Sequencing standards for footprinting reactions are prepared by dimethyl sulfate (DMS) modification and NaOH cleavage to produce a G + A ladder ***(37)***. The protocol described here is adapted from Bouet et al., 1999 (see **Note 27**).

1. Dilute the [^{32}P]-labeled DNA substrate in 200 µL of buffer A.

2. Add 5 µl of DMS: ethanol solution and incubate at room temperature for 1 minute (see **Note 10**).

3. Add 200 µL of quench solution.

4. Precipitate DNA with ethanol, and resuspend the pellet in 20 □L of buffer B.

5. Incubate 10 minutes at 90°C and then put on ice and spin down briefly.

6. Add 20 □L of 1N NaOH and incubate 5 minutes at 90°C.

7. Add 200 µL of Stop buffer to neutralize the reaction.

8. Precipitate with ethanol and resuspend the pellet with 5 µL of sequencing gel loading buffer. Store at -20°C until loading on the sequencing gel.

3.3.5 Resolution of the footprint patterns in a denaturing polyacrylamide gel

1. Prepare a denaturing 8% (w/v) denaturing polyacrylamide gel: for 100 mL, mix 20 mL of 40% (w/v) acrylamide:bis-acrylamide (19:1), 10 mL of TBE 10X, 42 g of urea, and add water up to a final volume of 100 mL. Add 1 mL of 10 % APS and 180 µL of TEMED and swirl the mixture gently. After adding APS and TEMED, pour the gel immediately and wait until the polymerization is complete (see **Notes 7** and **17**) and then clean the wells with the running buffer using a syringe.

2. Pre-run the gel for about 30 minutes at constant power (50 W) (see **Note 18**).

3. Heat the samples, including the cleavage products obtained for each transposase-DNA complex and unbound substrate and the G + A sequencing ladder, at 95°C for 5 min.

4. Load the samples and run the gel at constant power (50 W) (see **Note 19**).

5. Dry the gel on Whatman DE81 paper and exposed to a phosphor screen.

6. Scan the cleavage patterns for each complex and for unbound DNA using a PhosphorImager and quantify using an image analysis software (e.g. Image J ***(38)*** or ImageQuant^{TM}, GE Health care).

4. Notes

1. A whole range of fluorophores for labeling customized oligonucleotides at the 5' or 3' end is currently proposed by manufacturers. We routinely use Cy3- and Cy5- labeled oligonucleotides because they are amongst the least expensive ones and because their spectroscopic parameters are compatible with most basic fluorescence scanning setups. 5' labeled oligonucleotide is necessary to be used as a PCR primer.
2. Binding conditions may be optimized to favor the stability and specificity of DNA-protein interactions, such as the ionic strength and pH of the binding buffer, the presence/absence of glycerol, carrier proteins (*e.g.*, BSA), divalent cations (e.g., Mg^{2+}, Ca^{2+}, Mn^{2+} or Zn^{2+}), the concentration and type of competitor DNA, as well as the temperature and time of the binding reaction.
3. The addition of an excess of non-specific DNA competitor to the binding reaction mixture prevents or reduces non-specific DNA-protein interactions with the labeled substrate. One of the most widely used is the very simple co-polymers Poly-d[I-C]. However, other synthetic DNA such as Poly(dAdT)-poly(dAdT), or sheared salmon sperm or calf thymus DNA can be used as well. The competitor DNA must be incubated with the protein before adding the labeled DNA target.
4. The protein dilution buffer is used to dialyse the transposase at the end of purification.
5. Dye-containing loading buffer is a pre-mixed buffer containing one or more tracking dyes for gel electrophoresis. We recommend the use of Orange G since, migrating approximately the same as a 50 bp fragment of DNA, it will not appear in gel photographs.
6. In the case of poor fixation, other buffers, of low or high conductivity, can be tested for migration to stabilize complexes. The most common other buffers are Tris-Glycine-EDTA (10X: 500 mM Tris, 25 mM EDTA, 4 M Glycine) and Tris-acetate-EDTA (10X: 400 mM Tris-acetate, 10 mM EDTA).
7. Safety precaution: acrylamide is a powerful central and peripheral nervous system toxicant and has been classified as probably carcinogenic in humans. It is readily absorbed through skin, which requires wearing gloves, eye protection and lab coats to avoid direct contact with the skin when manipulating acrylamide solution and also polyacrylamide gels, as they may still contain low levels of unpolymerized acrylamide.
8. Safety precaution: manipulating radioactive materials such as [γ-^{32}P] ATP requires special procedures, and the reader has to refer to the guidelines on manipulation in effect at his research facility. Special precaution should be taken when manipulating

preparative gels to prepare labeled substrates.

9. Safety precaution: Solution preparation for in-gel footprinting with OP-Cu requires safety caution. Wear gloves, eye protection and dust mask when weighting all chemicals (1,10-phenanthroline, copper (II) sulfate anhydrous and neocuproine). Mercaptoproprionic acid is a toxic reagent that causes burns in contact with skin and eyes and requires wearing gloves, eye protection and lab coats and careful handling.
10. Safety precaution: DMS is an extremely toxic and carcinogenic chemical. It is readily absorbed through the skin and also very toxic by inhalation, so wear gloves and eye protection and work in a fume hood at all times when working with DMS.
11. A recurrent issue in protein production in bacterial cells is the loss of the expression vector during growth, due to segregational instability. This is notably favored if the expressed protein has some toxic or deleterious effect (which is often the case of transposases), or if the selection marker of the expression vector encodes an enzyme that inactivates the antibiotic in the medium (as is the case for □-lactamase, conferring resistance to ampicillin). To avoid this problem, we routinely centrifuge the pre-culture and resuspend the cells in fresh LB to wash the inoculum from secreted resistance enzymes prior to start the culture, and add fresh antibiotics 1 or 2 hours (at $OD_{600nm} \approx 0.3$-0.4) before induction. An alternative is to start the culture with a pool of ≈ 500 bacterial colonies grown on solid medium collected in 1 mL of fresh LB.
12. Fast chilling of the culture is important to prevent cells to adapt to lower temperature, thereby reducing the intracellular level of heat-shock chaperones.
13. Incubation may be performed at room T° if more convenient. The optimal time of induction depends on the protein and the expression system used. Incubation time giving the highest yield of soluble protein can be determined by scoring the amount of protein remaining in the supernatant of cell extracts after centrifugation.
14. The lysis buffer is the buffer in which bacterial cells will be lysed during the first step of protein purification and will depend on the purification protocol.
15. Simple linear DNA substrates are generally generated by PCR, while annealing is used for smaller linear substrates (e.g. in the range of 60 bp or less) or for more complex substrates like DNA forks.
16. The lengths of the short and long end substrates are chosen so as to optimize protein-DNA complexes separation in standard EMSA and possibly identify hybrid complexes (e.g., paired-end complexes containing both DNA substrates) with intermediate electrophoretic mobility. Typically, a length difference of ~60 to 80 bp between the

short and long substrate gives an adequate spatial distribution of complexes within the gel.

17. Ensure that gel plates, spacers and comb are clean and that all components are dry before pouring the gel. Avoid bubble formation and insert the well-forming comb immediately.
18. Pre-running electrophoresis is important to equilibrate the gel by removing small charged contaminants such as persulfate and ammonium ions and acrylic acid impurities.
19. Sample are preferably loaded using dedicated drawn-out micropipette tips. Up to ~100 µl of DNA sample can be loaded per well depending on the comb slots. Load identical samples onto adjacent wells if necessary. On the contrary, space out different substrates in separate areas of the gel to avoid cross-contaminations during purification.
20. When loaded in sufficient amount, Cy5- and Cy3-labeled DNA fragments are readily discernible to the naked eye within the gel, giving turquoise blue- and fuchsia-colored bands, respectively. Otherwise, briefly expose the gel to the fluorescent scanner and print a full-scale image of the result. Align the image with the gel by placing it underneath the glass plate. Use easily recognizable details such as the wells and the borders of the gel to properly align DNA bands with the corresponding gel segments.
21. Small DNA fragments (< 200 bp) reach maximal elution after 1 to 3 h, while larger fragments require more time.
22. Oligonucleotides are taken from a 100 µM stock solution. Higher concentrations of oligonucleotides can be used to produce larger amounts of substrates. A slight (1.5) molar excess of unlabeled oligonucleotide is added to the annealing reaction to titrate out the fluorescently-labeled partner in the duplex. Excess of un-annealed oligonucleotide is discarded by gel-purification of the substrate.
23. Slow cool down is important to conduct oligonucleotides annealing under steady state conditions to avoid formation of miss-paired products.
24. 30 min incubation of the binding reaction at 30°C is usually sufficient to reach the equilibrium.
25. The concentration of the polyacrylamide gel used in EMSA is chosen as a function of both the size of the labeled DNA substrates and the resolution of the resulting DNA-protein complexes. Usual concentrations vary from 4% for large labeled DNA fragments (150 bp or longer) to 8% for smaller DNA fragments. Increasing the

concentration may resolve DNA-protein complexes that would appear as a single diffuse complex at a low concentration, but on the other hand, high-concentration gels can also destabilize DNA-protein complexes.

26. Clean rigorously the pyrex dish before use (never use a plastic dish) and rinse extensively your gloves to remove the talk powder.
27. Dimethyl sulphate (DMS) is a chemical that methylates guanines (at the N-7 position) and adenines (at the N-3 position). These methylated bases are subsequently cleaved by NaOH.

Acknowledgments

We thank M. Chandler, P. Rousseau and F. Boudsocq for careful reading and suggestions concerning this manuscript. Research in the BH lab is supported by grants from the 'Fonds National de la Recherche Scientifique' (FRS-FNRS) [grants # 2.5496.04, 1.9226.04 and 1.5.225.10]. Research in the BTH lab is supported by Agence National pour la Recherche (ANR-12-BSV8-0009-01). C.G. research is supported by a grant from Toulouse III University (IDEX). E.N. held a research assistant fellowship from the FNRS.

5. References

1. Hickman A.B. and Dyda F. (2016) DNA Transposition at Work. *Chem Rev*. **116**(20):12758-12784. DOI:10.1021/acs.chemrev.6b00003.
2. Mizuuchi K. (1983) *In vitro* transposition of bacteriophage Mu: a biochemical approach to a novel replication reaction. *Cell*. **35**(3 Pt 2):785-94. PMID: 6317201.
3. Morisato D. and Kleckner N. (1987) Tn*10* transposition and circle formation *in vitro*. *Cell*. **51**(1):101-11.PMID: 2820584.
4. Bainton R.J., Kubo K.M., Feng J.N. and Craig N.L. (1993) Tn*7* transposition: target DNA recognition is mediated by multiple Tn*7*-encoded proteins in a purified *in vitro* system. *Cell*. **72**(6):931-43. PMID: 8384534.
5. Goryshin I.Y. and Reznikoff W.S. (1998) Tn*5 in vitro* transposition. *J Biol Chem.* **273**(13):7367-74. PMID: 9516433.
6. Ton-Hoang B., Polard P. and Chandler M. (1998) Efficient transposition of IS*911* circles *in vitro*. *EMBO J.* **17**(4):1169-81. PMID:9463394. DOI: 10.1093/emboj/17.4.1169.

7. Leschziner AE, Griffin TJ 4th, Grindley ND. (1998) Tn*552* transposase catalyzes concerted strand transfer *in vitro*. *Proc Natl Acad Sci U S A*. **95**(13):7345-50. PMID: 9636151.
8. Ton-Hoang B., Turlan C. and Chandler M. (2004) Functional domains of the IS*1* transposase: analysis *in vivo* and *in vitro*. *Mol Microbiol.* **53**(5):1529-43. PMID: 15387827.
9. Guynet C., Hickman A.B., Barabas O., Dyda F., Chandler M. and Ton-Hoang B. (2008) *In Vitro* Reconstitution of a Single-Stranded Transposition Mechanism of IS*608*. *Mol Cell* **29**:1-11.
10. Nicolas E., Oger C.A., Nguyen N., Lambin M., Draime A., Leterme S.C., Chandler M. and Hallet B.F. (2017) Unlocking Tn3-family transposase activity in vitro unveils an asymetric pathway for transposome assembly. *Proc Natl Acad Sci* USA :**114**(5):E669-E678. DOI: 10.1073/pnas.1611701114.
11. Kleckner N. (1990) Regulation of transposition in bacteria. *Annu Rev Cell Biol.* **6**:297-327. Review. DOI:10.1146/annurev.cb.06.110190.001501.
12. Duval-Valentin G. and Chandler M. (2011) Cotranslational control of DNA transposition: a window of opportunity. *Mol Cell.* **44**(6):989-96. DOI: 10.1016/j.molcel.2011.09.027.
13. Chalmers R.M. and Kleckner N. (1994) Tn*10*/IS*10* transposase purification, activation, and *in vitro* reaction. *J Biol Chem.* 269(11):8029-35. PMID:8132525.
14. Nolivos S., Pages C., Rousseau P., Le Bourgeois P. and Cornet F. (2010) Are two better than one? Analysis of an FtsK/Xer recombination system that uses a single recombinase. *Nucleic Acids Res.* **38**(19):6477-89. PMID: 20542912. DOI: 10.1093/nar/gkq507.
15. Lambin M., Nicolas E., Oger C.A., Nguyen N., Prozzi D. and Hallet B. (2012) Separate structural and functional domains of Tn4430 transposase contribute to target immunity. *Mol Microbiol.* **83**(4):805-20. PMID: 22624153. DOI: 10.1111/j.1365-2958.2012.07967.x.
16. Messing S.A., Ton-Hoang B., Hickman A.B., McCubbin A.J., Peaslee G.F., Ghirlando R., Chandler M. and Dyda F. (2012) The processing of repetitive extragenic palindromes: the structure of a repetitive extragenic palindrome bound to its associated nuclease. **40**(19):9964-79. PMID: 22885300. DOI: 10.1093/nar/gks741.
17. Carra S., Alberti S., Arrigo P.A., Benesch J.L., Benjamin I.J., Boelens W., Bartelt-Kirbach B., Brundel B.J., Buchner J., Bukau B., Carver J.A., Ecroyd H., Emanuelsson

C., Finet S., Golenhofen N., Goloubinoff P., Gusev N., Haslbeck M., Hightower L.E., Kampinga H.H., Klevit R.E., Liberek K., Mchaourab H.S., McMenimen K.A., Poletti A., Quinlan R., Strelkov S.V., Toth M.E., Vierling E. and Tanguay R.M. (2017) The growing world of small heat shock proteins: from structure to functions. *Cell Stress Chaperones* **2017**. DOI: 10.1007/s12192-017-0787-8.

18. Maleki F., Khosravi Z., Nasser A., Taghinejad H. and Azizian M. (2016) Bacterial Heat Shock Protein Activity. *J Clin Diagn Res.* **10**(3):BE01-3. DOI: 10.7860/JCDR/2016/14568.7444.
19. Costa S., Almeida A., Castro A. and Domingues L. (2014) Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Front Microbiol.***19**;5:63. DOI: 10.3389/fmicb.2014.00063.
20. Arya R., Sabir J.S., Bora R.S. and Saini K.S. (2015) Optimization of culture parameters and novel strategies to improve protein solubility. *Methods Mol Biol.* 1258:45-63. DOI: 10.1007/978-1-4939-2205-5_3.
21. Jia B. and Jeon C.O. (2016) High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open Biol.* **6**(8). pii: 160196. DOI: 10.1098/rsob.160196.
22. Dey B., Thukral S., Krishnan S., Chakrobarty M., Gupta S., Manghani C. and Rani V. (2012) DNA-protein interactions: methods for detection and analysis. *Mol Cell Biochem*. **365**(1-2):279-99. DOI: 10.1007/s11010-012-1269-z.
23. Carey M.F., Peterson C.L. and Smale S.T. (2012) Experimental strategies for the identification of DNA-binding proteins. *Cold Spring Harb Protoc.* 2012(1):18-33. DOI: 10.1101/pdb.top067470.
24. Lane D, Prentki P and Chandler M. (1992) Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol Rev*. **56**(4):509-28. Review. PMID:1480106.
25. Ruscher K., Reuter M., Kupper D., Trendelenburg G., Dirnagl U., and Meisel A. (2000) A fluorescence based non-radioactive electrophoretic mobility shift assay. *J Biotechnol.* **78**(2):163-70. PMID:10725539.
26. Garner M.M. and Revzin A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**:3047-3059. PMID:6269071.

27. Fried M. and Crothers D.M. (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505-6525. PMID:6275366.
28. Wu H.M. and Crothers D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*. **308**(5959):509-13.PMID:6323997.
29. Bouet J.Y. and Lane D. (2009) Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex. *J Biol Chem*. **284**(1):165-73. DOI: 10.1074/jbc.M802752200. PMID:19001378.
30. Papavassiliou A.G. (2009) Footprinting DNA-protein interactions in native polyacrylamide gels by chemical nucleolytic activity of 1,10-phenanthroline-copper. *Methods Mol Biol*. **543**:163-99. DOI: 10.1007/978-1-60327-015-1_13.
31. Normand C., Duval-Valentin G., Haren L. and Chandler M. (2001) The terminal requirements for synaptic complex assembly and activity. *J Mol Biol.* **18**; 308(5):853-71. DOI: 10.10006/jmbi.2001.4641.
32. He S., Guynet C., Siguier P., Hickman A.B., Dyda F., Chandler M. and Ton-Hoang B. (2013) IS200/IS605 family single-strand transposition: mechanism of IS608 strand transfer. *Nucleic Acids Res.* 41(5):3302-13.
33. Bouet J.-Y. and Funnell B. E. (1999) P1 ParA interacts with the P1 partition complex at parS and an ATP-ADP switch controls ParA activities. *EMBO J.* **18,** 1415–1424. PMID: 10064607.
34. Spassky A., Rimsky S., Buc H. and Busby S. (1998) Correlation between the conformation of Escherichia coli -10 hexamer sequences and promoter strength: use of orthophenanthroline cuprous complex as a structural index. *EMBO J.* **7**(6):1871-9. PMID:3049079.
35. Sigman, S. R., Kuwabara, M. D., Chen, C. H., and Bruice, T. W. (1991) Nuclease activity of 1,10-phenanthroline-copper in study of protein– DNA interactions. *Methods Enzymol.* **208:** 414–433. PMID:1779842.
36. Rosano G.L. and Ceccarelli E.A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol.* **5**:172. DOI: 10.3389/fmicb.2014.00172. Review.
37. Sambrook J., Fritsch E.F. and ManiatisT. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

38. Schneider C.A., Rasband W.S. and Eliceiri K.W. (2012) "NIH Image to ImageJ: 25 years of image analysis". Nature Methods **9**: 671-675.

39.

Figure legends

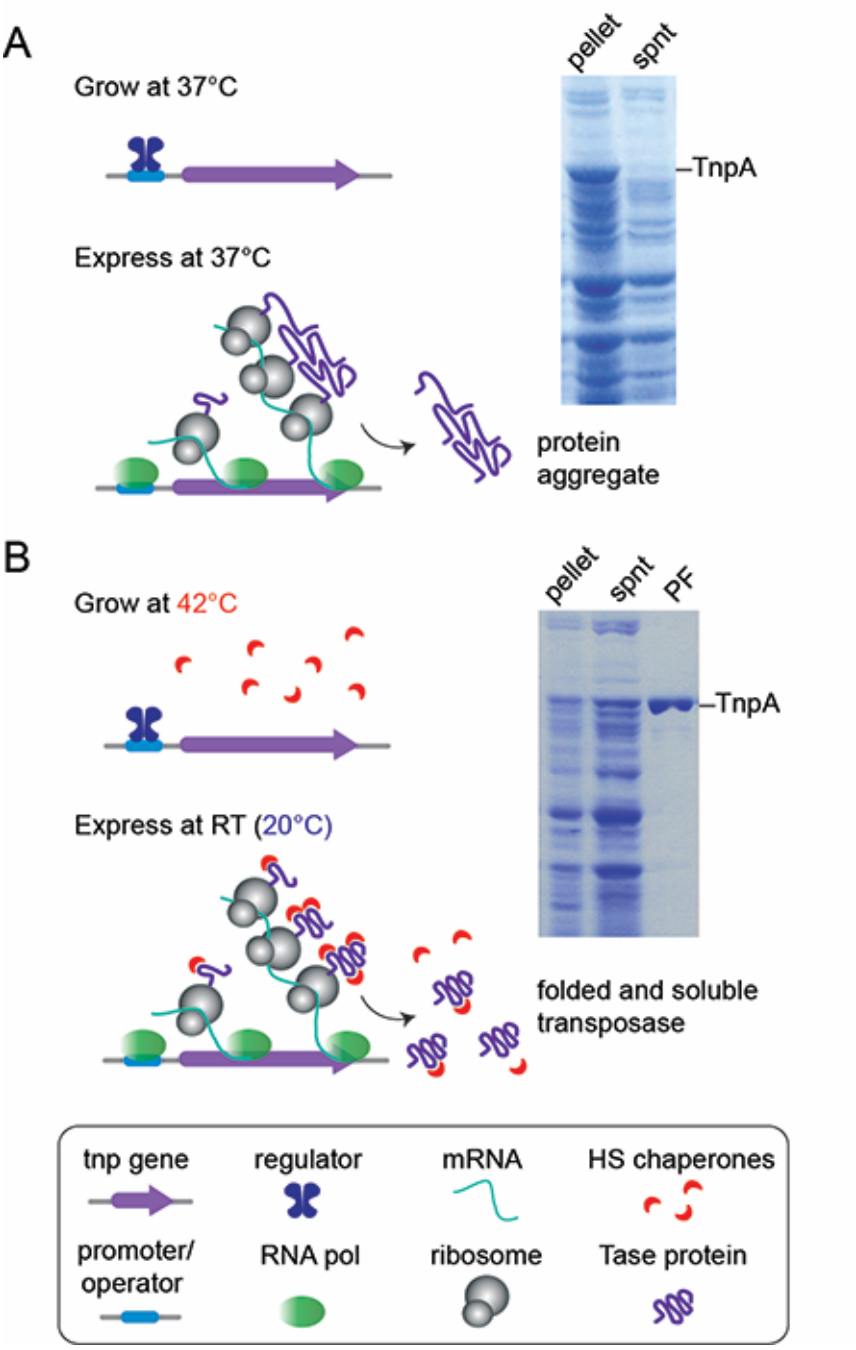

Figure 1: Principle of the "thermal shift induction" method for active transposase expression and purification. (A) In classical protein expression protocols, *E. coli* cells harbouring the transposase expression vector are grown under repressed conditions at the optimal temperature (37°C) until mid-log phase prior to start expression. With poorly soluble proteins, the procedure often leads to the formation of aggregates due to co-precipitation of unfolded

proteins. As shown for the TnpA transposase of Tn*4430*, the majority of the expressed transposase is then retained in the pellet after cell lysis and centrifugation. (B) In the thermal shift procedure, cells are first grown at elevated temperature (42°C) in order to load them with heat-shock (HS) protein chaperones prior to chill the culture to room temperature (RT° or ~20°C) and add the inducer. Low thermal agitation together with the presence of chaperones facilitate protein folding during expression. For Tn*4430* TnpA and other transposases, the procedure increased the yield of soluble and active proteins in the supernatant of cell lysates (spnt), which allowed their purification by simple affinity chromatography. PF, purified fraction. Coomassie-stained gel of (B) was extracted from ***(10)*** with permission.

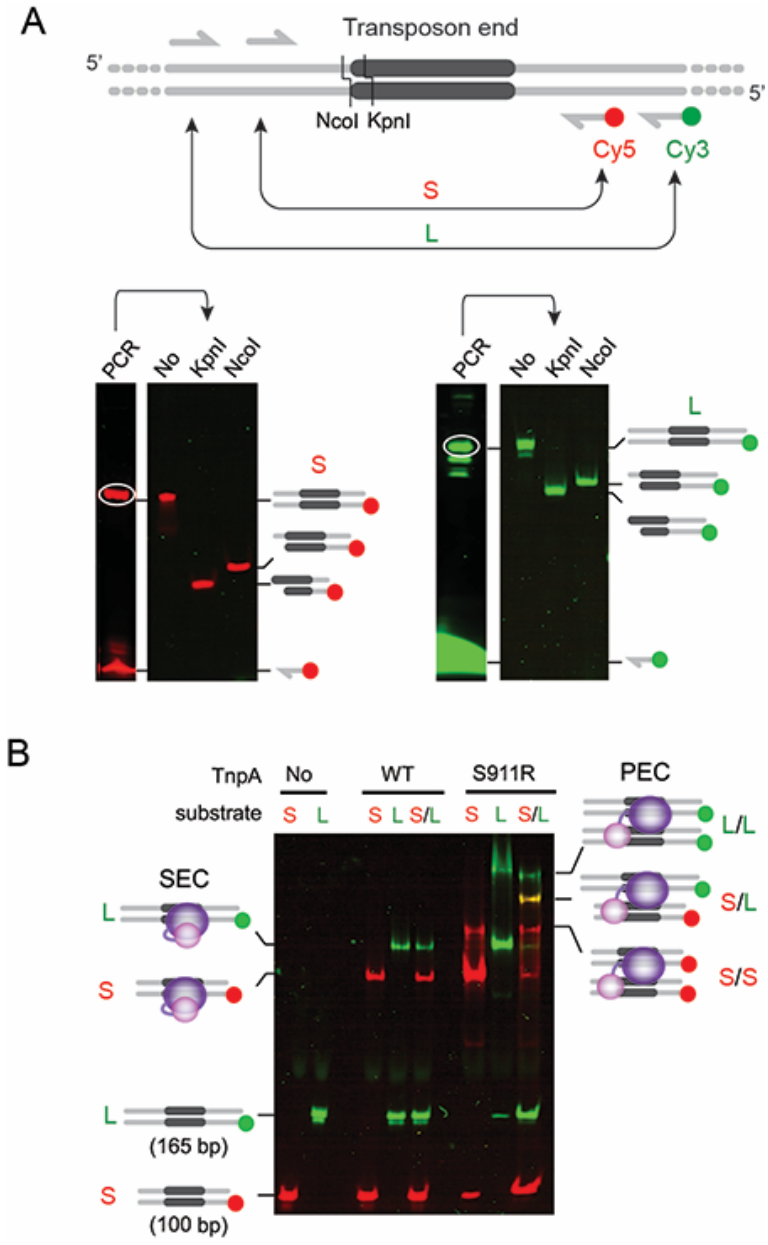

Figure 2: Short-long EMSA experiment with differentially-labeled fluorescent DNA probes. (A) The figure illustrates the case where fluorescently labeled DNA substrates are produced by PCR. A short (S; *e.g.*, ~80-120 bp) and a long (L; *e.g.*, ~160-200 bp) DNA fragment containing the transposon end are PCR-amplified using a 5'Cy5-labeled (red dot) or a 5'Cy3-labeled (green dot) specific primer on one side and unlabeled primer on the opposite side. PCR products are run on a 4% to 8% polyacrylamide gel and analysed by fluorescence scanning using appropriate excitation/detection filters. DNA substrates are purified from the gel and checked by digestion with appropriate restriction enzymes (*e.g.*, *Kpn*I and *Nco*I in the example) followed by gel electrophoresis. (B) Example of a short-long experiment performed with the wild type TnpA transposase (WT) and the TnpA^{S911R} mutant (S911R) of Tn*4430*.

The proteins were incubated with a short (S, 100 bp) and a long (L, 165 bp) DNA substrate separately, and with both substrates together in the same reaction (S/L). The reactions were analysed under standard EMSA condition and the gel was scanned to specifically detect Cy3 and Cy5 fluorescence (green and red bands, respectively). Wild type TnpA formed a single-end complex (SEC) on both the short and the long substrate. In addition to SEC, the deregulated $TnpA^{S911R}$ mutant formed an additional, slower migrating paired end complexe (PEC) in which the protein binds to two transposon ends simultaneously. In the mixed reaction, this resulted in the formation of an intermediate complex containing a short and a long DNA substrate (yellow band in the overlay). Fluorescent scan of (B) was extracted from ***(10)*** with permission.