Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 7570–7574, November 1985
Biochemistry

A defined system for the DNA strand-transfer reaction at the initiation of bacteriophage Mu transposition: Protein and DNA substrate requirements

(protein HU/DNA recombination)

ROBERT CRAIGIE*, DONNA J. ARNDT-JOVIN†, AND KIYOSHI MIZUUCHI*

*Laboratory of Molecular Biology, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20205; and †Abteilung Molekulare Biologie, Max-Planck Institut für Biophysikalische Chemie, Postfach 2841, D-3400 Göttingen, Federal Republic of Germany

Communicated by Gary Felsenfeld, July 23, 1985

ABSTRACT **An early step in the transposition of bacteriophage Mu DNA *in vitro* is a DNA strand-transfer reaction that generates an intermediate DNA structure in which the Mu donor DNA and the target DNA are covalently joined. DNA replication, initiated at the DNA forks in this intermediate, generates a cointegrate product; simple insert products can also be formed from the same intermediate by degradation of a specific segment of the structure, followed by gap repair. This DNA strand-transfer reaction requires ATP, magnesium, the Mu A and Mu B proteins, and a factor supplied by an *Escherichia coli* cell extract. We have now shown that the host protein factor requirement can be satisfied by purified protein HU. The defined system has been used to determine the DNA substrate requirements for the reaction. The reaction requires the two Mu ends, located on the same DNA molecule, in the same relative orientation to one another as in the phage Mu genome. To participate in the strand-transfer reaction efficiently the mini-Mu plasmid, used as the transposon donor, must be supercoiled; the target DNA molecule may be supercoiled, relaxed circular, or linear.**

Bacteriophage Mu is a temperate phage of *Escherichia coli* and many other species of enterobacteria (for a recent review, see ref. 1). Phage Mu uses a replicative transposition reaction to replicate its genome during the lytic cycle. Transposition of Mu is highly efficient relative to that of other prokaryotic transposons (for a recent review of prokaryotic transposons, see ref. 2), making Mu an attractive system for biochemical analysis of the transposition reaction. The Mu transposition reaction requires, in addition to the two Mu ends on the same molecule and in proper orientation, the Mu A protein (3, 4) and an unknown number of host proteins; the Mu B protein enhances the efficiency of the reaction by a factor of about 100 (4–6). The development of an *in vitro* system for Mu transposition (7) has greatly facilitated biochemical analysis of the reaction.

The *in vitro* system (7) we have exploited for studying the mechanism of Mu transposition uses a mini-Mu plasmid (pMK108), containing the Mu left and right ends in the same relative orientation as in the Mu genome, as the transposon donor and ϕX174 replicative form DNA (ϕXRF DNA) as the target of transposition; intramolecular transposition of pMK108 and transposition of pMK108 into another pMK108 molecule also occur, but, under the conditions used, the major products result from the intermolecular reaction involving pMK108 and ϕXRF DNA. The Mu A and Mu B proteins are provided in the form of cell extracts or purified proteins; host proteins are provided by a cell extract. The complete *in vitro* Mu transposition reaction produces both cointegrate and simple insert products. Cointegrates consist of fused donor and target molecules with duplicated copies of the transposon at the junctions, whereas simple inserts are insertions of a single copy of the transposon into a target molecule. In both cases a short target site sequence is duplicated. We have recently succeeded in trapping an intermediate in Mu transposition and determining its structure (8). This intermediate has the structure predicted by Shapiro (9), with the previously predicted (10) polarity of the DNA strand-transfer step. Both cointegrates and simple inserts are generated from this intermediate *in vitro* by reactions that require only *E. coli* host proteins (8). Cointegrates are generated by DNA replication, initiated at the DNA branches in the intermediate; simple inserts can be made by degradation of the DNA that flanked the transposon part of the donor molecule, followed by gap repair (10).

Many prokaryotic transposons, like Mu, can make both simple insert and cointegrate transposition products. Such transposons may share a common mechanism of transposition but differ in their bias toward simple insertion or cointegration at a branch point in the transposition pathway. This bias is likely to be influenced by their respective transposases and, where present, accessory proteins. These proteins will affect other cellular components that are involved either in successful initiation of DNA replication of the intermediate, which leads to cointegration, or in degradation of the DNA that flanked the transposon part of the donor molecule, which leads to simple insertion.

The Mu transposition intermediate is formed by a pair of single-strand DNA transfers, one from each end of the transposon, to the DNA target site. For high efficiency, this DNA strand-transfer reaction requires ATP, magnesium, the Mu A and Mu B proteins, and host cell extract (8). Here, we report the purification and identification of the host protein factor and the DNA substrate requirements for this step of the transposition reaction.

MATERIALS AND METHODS

Nucleic Acids and Proteins. The mini-Mu plasmid pMK108 contains the two Mu ends in proper orientation and has been described previously (7, 10). pMK34 is the same as pMK108 except that the *Eco*RI fragment that contains the Mu right end is in inverted orientation. pMK20 differs from pMK108 in that the *Bam*HI fragment that contains the Mu left end is deleted. Similarly, pMK35 is the same as pMK108 except that the *Eco*RI fragment that contains the Mu right end is deleted. ϕXRF DNA was purchased from Bethesda Research Laboratories. The Mu A protein was purified as described previously (11); the Mu B protein was purified as described by Chaconas *et al.* (12). Type I topoisomerase-containing ex-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: ϕXRF DNA, ϕX174 replicative form DNA.

tract (from chicken erythrocytes) was provided by Mary O'Dea of this laboratory. Pronase was from Calbiochem. Protein HU, originally from A. Kornberg (Stanford), was provided by H. Nash (National Institutes of Health). Antibody to protein HU, originally from J. Rouvière-Yaniv (Pasteur Institute), was also provided by H. Nash.

Chromatography Media. DEAE-Sepharose and Sephadex G-100 (superfine) were purchased from Pharmacia, phosphocellulose (cellulose phosphate P-11) was from Whatman, and hydroxylapatite (Bio-Gel HTP) was from Bio-Rad.

Mu DNA Strand-Transfer Reaction. The reaction conditions were simplified from those reported previously (8). Mixtures (25 μl) contained 25 mM Tris·HCl (pH 8), 140 mM NaCl, 10 mM $MgCl_2$, 2mM ATP, 1 mM dithiothreitol, bovine serum albumin (25 μg/ml), pMK108 DNA (10 μg/ml), φXRF DNA (10 μg/ml), Mu B protein (4 μg/ml), Mu A protein (5 μg/ml), and protein HU or an aliquot of a fraction to be assayed for host factor, as noted in the figure legends. Reactions were started by addition of the Mu A protein and incubated at 30°C for 1 hr, then stopped by the addition of an equal volume of 0.6 M NaOAc (pH 7.0), 40 mM EDTA, and tRNA (100 μg/ml), precipitated with 3 volumes of ethanol, and dried. The samples were resuspended in 50 μl of 10 mM Tris·HCl (pH 8), 0.5 mM EDTA, 50 mM NaCl, 0.2% $NaDodSO_4$. Pronase was added to a concentration of 100 μg/ml and the samples were incubated at 37°C for 1 hr. After ethanol precipitation, the Pronase-treated samples were suspended in 10 mM Tris·HCl (pH 8), 0.5 mM EDTA.

Electrophoresis. The DNA products were electrophoresed in 0.7% high-gelling-temperature agarose (SeaKem Laboratories, Rockland, ME) in TBE buffer (90 mM Tris base, 90 mM boric acid, 2.5 mM EDTA) for 2 hr at 5.5 V/cm and stained with ethidium bromide. Protein samples were electrophoresed in 15% acrylamide/0.4% bisacrylamide $NaDodSO_4$/polyacrylamide gels (13), and stained with Coomassie brilliant blue.

Purification of the Host Protein Factor. Purification of the host factor, from *E. coli* N100 (25 g of cell paste), was as described by Dixon and Kornberg (14) for protein HU, including 1 M KCl in the lysis step. In our hands, after the hydroxylapatite step the HU still contained several contaminant proteins. These contaminants were removed by two additional steps. The pooled hydroxylapatite fractions were dialyzed against 20 mM Tris·HCl (pH 8), 1 mM EDTA, 1 mM dithiothreitol, 60 mM KCl, 10% (wt/vol) glycerol and passed over a DEAE-Sepharose column (1-ml bed volume) equilibrated with the same buffer. Protein HU was found in the flow-through fraction as expected (15). This fraction was concentrated by precipitation with ammonium sulfate and suspended in 200 μl of 25 mM Hepes (pH 7.5), 100 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol, 5% glycerol, and applied to a Sephadex G-100 column (0.75 cm × 50 cm) equilibrated with the same buffer. The column was run at a flow rate of 0.9 ml/hr and 240-μl fractions were collected. Protein concentrations were determined using the Bio-Rad protein assay, with bovine serum albumin as the standard.

RESULTS

Purification of the Host Protein Factor. Craigie and Mizuuchi (8) have shown that a fused donor-target DNA structure is generated in a reaction requiring ATP, magnesium, the Mu A and B proteins, and an *E. coli* host cell extract. The structure of this transposition intermediate is depicted in Fig. 1. We have now used this reaction as an assay to purify the host protein factor.

Preliminary experiments suggested that this factor was a low molecular weight DNA binding protein. Initially, *E. coli* integration host factor for phage λ integration was considered as a candidate; however, a requirement for integration host factor was excluded because cell extracts made from *E. coli* strains carrying deletion mutations in the genes coding for both integration host factor subunits support the reaction (data not shown). Another likely candidate was protein HU, and the data presented below show that HU does indeed satisfy the host protein factor requirement.

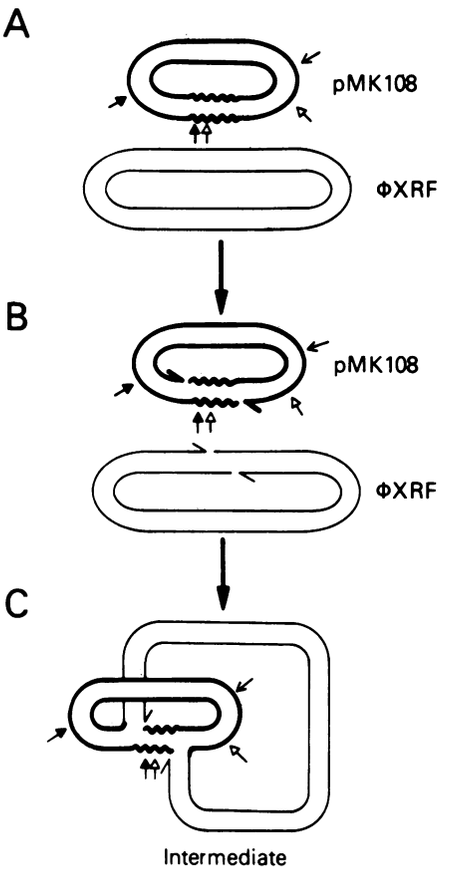

FIG. 1. Structure of an intermolecular transposition intermediate. (*A*) The mini-Mu plasmid pMK108 with φXRF DNA used as the transposition target. Mini-Mu part of pMK108, thick lines; DNA flanking the mini-Mu part of pMK108, wavy lines; φXRF DNA, thin lines. (*B*) A staggered cut in the target molecule and nicks on opposite strands at each end of the mini-Mu DNA, coupled with a pair of single-strand DNA transfers, generate the intermediate structure shown in *C*. Half arrows denote the 3′ end of each strand. The locations of the *Eco*RI (open arrows), *Bam*HI (solid arrows), and *Pvu* I (line arrows) restriction sites are indicated.

The possible requirement for protein HU was tested by purifying protein HU essentially as described by Dixon and Kornberg (14), assaying for complementing activity for the Mu strand-transfer reaction at each stage of the purification. Activity was always coincident with the HU-containing fractions. Fig. 2*A* shows an assay of fractions eluted from the hydroxylapatite column; a Coomassie-stained gel of the same fractions is shown in Fig. 2*B*. The contaminating proteins remaining after the hydroxylapatite step were removed by passing the pooled HU-containing fractions over DEAE-Sepharose and then chromatographing on Sephadex G-100. The activity eluted from the Sephadex G-100 column with the protein HU peak (Fig. 3). At this stage no contaminants were detected by Coomassie staining (Fig. 3*B*). The purified Mu A and Mu B proteins and the pooled HU peak after Sephadex G-100 chromatography are the only proteins required for efficient formation of the transposition intermediate (Fig. 4). In this purified system, as in the semipurified system (8), the reaction requires ATP and magnesium (data not shown).

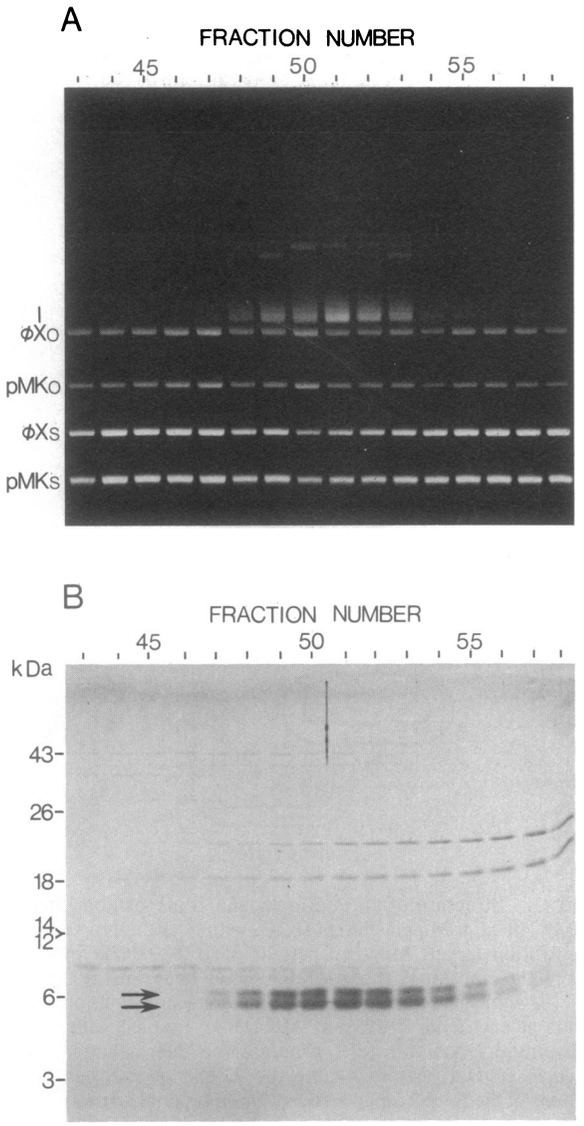

FIG. 2. Assay and $NaDodSO_4$/polyacrylamide gel electrophoresis of fractions after the hydroxylapatite purification step. (*A*) Equal volumes of each fraction eluted from the hydroxylapatite column were assayed and the products were electrophoresed in agarose as described; the gel was stained with ethidium bromide. The following bands are labeled: pMK108 (pMK); φXRF DNA (φX); intermolecular (pMK108 and φXRF DNA) transposition intermediate (I). The subscripts "s" and "o" denote the supercoiled and open circle species, respectively. (*B*) Aliquot of the same fractions were electrophoresed in a 15% $NaDodSO_4$/polyacrylamide gel. The gel was stained with Coomassie brilliant blue. Protein HU, which is just resolved as a doublet, is indicated by arrows. Since protein HU stains poorly with Coomassie blue (15), this staining method underrepresents the abundance of HU relative to contaminants. Ovalbumin (43 kDa), α-chymotrypsinogen (26 kDa), β-lactoglobulin (18 kDa), lysozyme (14 kDa), cytochrome *c* (12 kDa), bovine trypsin inhibitor (6 kDa), and insulin (3 kDa) were coelectrophoresed as standards.

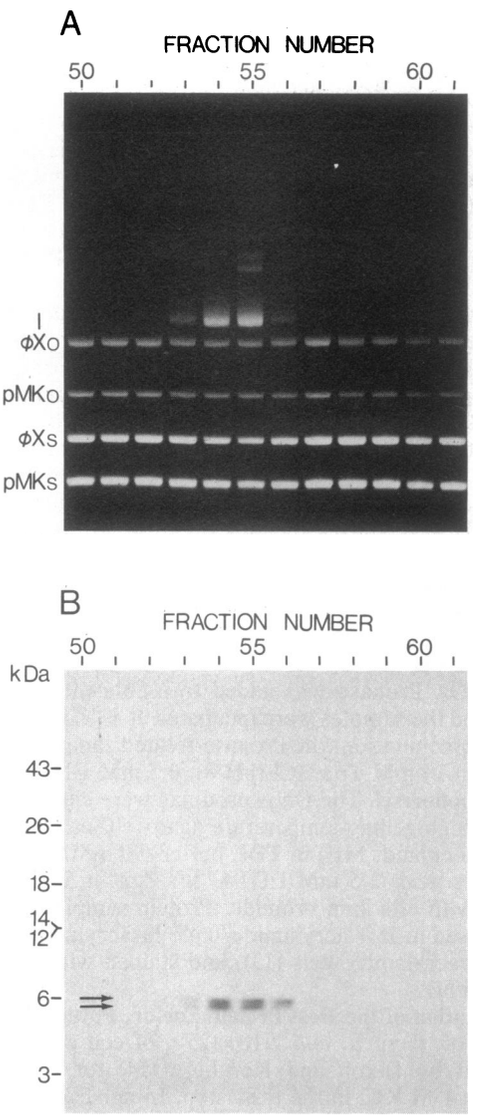

FIG. 3. Assay and $NaDodSO_4$/polyacrylamide gel electrophoresis of fractions after the Sephadex G-100 purification step. An aliquot of each of the fractions eluted from the Sephadex G-100 column was assayed for activity (*A*) and analyzed by $NaDodSO_4$/polyacrylamide gel electrophoresis (*B*). Labeling is as in Fig. 1.

The identification of this protein factor as HU was checked in several ways. The protein comigrated with protein HU in $NaDodSO_4$/polyacrylamide gel electrophoresis (data not shown) and, throughout the purification, activity eluted from each column as expected for HU. Protein HU, isolated as a stimulatory factor for *oriC*-dependent DNA replication *in vitro* (14), supports the Mu reaction (Fig. 5). The reaction is inhibited by antibody to HU (lane b), but this inhibition is overcome by addition of excess HU (lanes c–e). Similarly, reactions using our purified factor are inhibited by antibody to HU and this inhibition is overcome by the addition of excess factor (data not shown). Protein HU isolated from *Anabaena*, which can substitute for *E. coli* protein HU in the *oriC*-dependent DNA replication assay (14), also supports the Mu reaction (data not shown). These data confirm that the factor is protein HU.

DNA Substrate Requirements. The DNA supercoiling requirements for the reaction were investigated using pairwise combinations of supercoiled or relaxed mini-Mu plasmid DNA and φXRF target DNA as the DNA substrates for the reaction. The intermolecular intermediate product, which is the major product under the reaction conditions used, is formed with either supercoiled (Fig. 6, lane a) or relaxed (lane b) φXRF DNA as the target, provided that the mini-Mu plasmid is supercoiled. No product is observed with either supercoiled (lane c) or relaxed (lane d) φXRF DNA as the target if the mini-Mu plasmid is relaxed. Since relaxed DNA is a competent target, it was anticipated that the reaction would work using a linear target DNA. With linear φXRF DNA as the target, a new band is indeed seen (lane e) that is not formed in the absence of Mu A protein (lane f). This product contains both mini-Mu plasmid DNA and φXRF DNA (data not shown), as expected; however, we have not

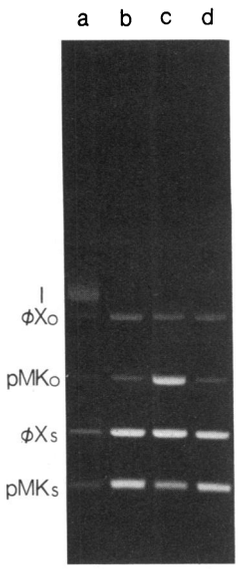

FIG. 4. Protein requirements for formation of the intermediate. Labeling of bands is as in Fig. 2*A*. The complete reaction mixture (lane a) contained Mu A protein at 5 μg/ml, Mu B protein at 4 μg/ml, and HU protein at 2 μg/ml. The protein HU used was the pooled Sephadex G-100 fractions. Other lanes: b, Mu A protein was omitted; c, Mu B protein was omitted; d, protein HU was omitted.

analyzed its structure in detail. The only topological requirement for the reaction is therefore that the mini-Mu DNA be supercoiled.

No product is formed using a mini-Mu plasmid that has the Mu ends in the wrong relative orientation (lane g); lane h is a control reaction omitting the Mu A protein. Similarly, a mixture of plasmids, each containing either a single Mu left end or a single Mu right end does not serve as a substrate (lanes i and j).

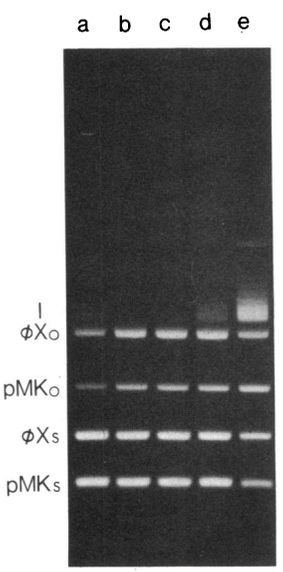

FIG. 5. Intermediate formation is inhibited by antibody to protein HU. Lane a: the complete reaction mixture (see *Materials and Methods*), except for the Mu A and Mu B proteins, was preincubated at 30°C for 15 min; lanes b–e: as for lane a, except that each reaction mixture contained an equal quantity of antibody to protein HU. Protein HU (from A. Kornberg) was used at 2 μg/ml (lanes a and b), 4 μg/ml (lane c), 8 μg/ml (lane d), or 20 μg/ml (lane e). After preincubation, the Mu B protein and the Mu A protein were added to final concentrations of 4 and 5 μg/ml, respectively, and the mixtures were incubated for 1 hr at 30°C. Labeling of the bands is as in Fig. 2*A*.

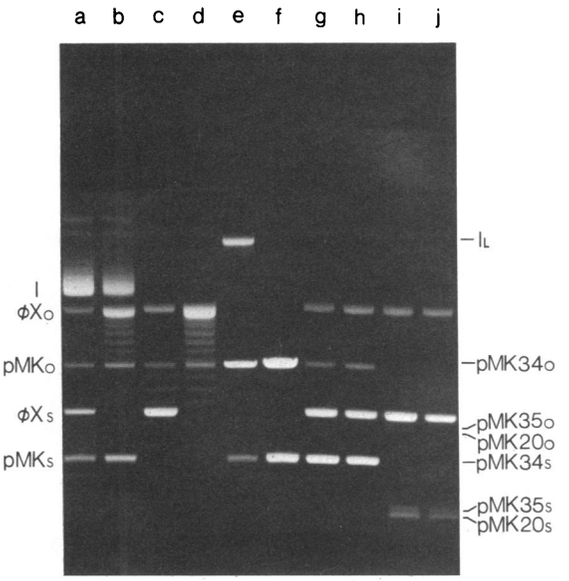

FIG. 6. DNA substrate requirements for formation of the intermediate. Reaction mixtures contained the Mu A protein at 5 μg/ml, the Mu B protein at 4 μg/ml, and protein HU (pooled fractions after Sephadex G-100 chromatography) at 2 μg/ml, except that the Mu A protein was omitted from the reactions shown in lanes f, h, and j. The DNA substrates for the reactions were as follows: lane a, supercoiled pMK108 plus supercoiled ϕXRF; lane b, supercoiled pMK108 plus relaxed ϕXRF; lane c, relaxed pMK108 plus supercoiled ϕXRF; lane d, relaxed pMK108 plus relaxed ϕXRF; lanes e and f, supercoiled pMK108 plus linear ϕXRF; lanes g and h, supercoiled pMK34 plus supercoiled ϕXRF; lanes i and j, supercoiled pMK35 plus supercoiled pMK20 plus supercoiled ϕXRF. pMK34 contains the two Mu ends in the wrong orientation. pMK35 contains a single Mu left end, and pMK20 contains a single Mu right end. The supercoiled and open-circle forms of these plasmids are denoted by the subscripts "s" and "o", respectively. Other bands are labeled as in Fig. 2*A*; an additional band (I_L) corresponding to the intermolecular intermediate structure generated with linear ϕXRF DNA as the target is also labeled.

DISCUSSION

The DNA strand-transfer reaction is a central step in Mu transposition. In addition to the Mu A and Mu B proteins, the reaction requires only the host protein HU. We cannot, however, exclude the possibility that other host proteins may affect the efficiency of the reaction. The Mu A protein recognizes the ends of Mu (16) and provides the specificity factor; both protein HU (15) and the Mu B protein (ref. 12; unpublished data) bind to DNA nonspecifically. Since Mu transposition occurs *in vivo* in the absence of the Mu B protein (4–6) albeit at much lower efficiency, the DNA cutting and joining required to generate the intermediate is most likely to be mediated by the Mu A protein; protein HU is not expected to have such activity. We have not yet determined whether a low level of intermediate is produced in reactions omitting the Mu B protein. It is therefore still possible that the low level of transposition observed *in vivo* in the absence of the Mu B protein may require an additional host protein that can substitute inefficiently for the Mu B protein. Protein HU, together with the Mu A protein, might be involved in organizing the tertiary structure of the two Mu end DNA segments and possibly in the formation of a complex between the two Mu ends and the DNA target site that is likely to precede the reaction.

The reaction requires the two Mu ends, located on the same DNA molecule, in the same relative orientation as in the Mu genome. This structural requirement, which is necessary for the complete Mu transposition reaction *in vivo* (17, 18) and *in vitro* (7), is required for this early step in the

transposition reaction. Several other DNA recombination systems—e.g., resolution of cointegrates by the transposon γδ (19) and Tn*3* (20) resolvase proteins—also share this type of requirement. Although attempts have been made to explain this phenomenon in terms of tracking models (ref. 20; see also ref. 21) in which a protein initially binds to one DNA site and then finds a second site by one-dimensional "tracking" along the DNA, direct experimental support is lacking. The other DNA substrate requirement is that the mini-Mu plasmid must be negatively supercoiled. This requirement will be satisfied *in vivo* for prophage induction, but infecting phage Mu DNA is linear and does not form covalently closed circular structures. However, a circular form of Mu DNA, in which the two ends are held together by a protein bridge, has been observed after infection (22). Further, this circular DNA has been reported to undergo supercoiling (22). The free energy of supercoiling and ATP hydrolysis are both possible energy sources for driving the Mu DNA strand transfer reaction; their roles remain to be elucidated.

The requirements for this recombination reaction have turned out to be rather simple. The ability to study this reaction in a defined system should make it easier to investigate this step of the Mu transposition reaction in detail.

We are grateful to Howard Nash for his advice and helpful discussions on this work. We thank Arthur Kornberg and Robert Haselkorn for gifts of protein HU and Josette Rouvière-Yaniv for providing antibody to protein HU. We also thank Martin Gellert for critical reading of this manuscript and Mary Lou Miller for careful typing.

1. Toussaint, A. & Résibois, A. (1983) in *Mobile Genetic Elements*, ed. Shapiro, J. A. (Academic, New York), pp. 105–158.
2. Kleckner, N. (1981) *Annu. Rev. Genet.* **15**, 341–404.
3. Wijffelman, C. & Lotterman, B. (1977) *Mol. Gen. Genet.* **151**, 169–174.
4. O'Day, K. J., Schultz, D. & Howe, M. M. (1978) in *Microbiology 1978*, ed. Schlessinger, D. (Am. Soc. Microbiol., Washington, DC), pp. 48–51.
5. Faelen, M., Huisman, O. & Toussaint, A. (1978) *Nature (London)* **271**, 580–582.
6. Groenen, M. A. M., Timmers, E. & van de Putte, P. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2087–2091.
7. Mizuuchi, K. (1983) *Cell* **35**, 785–794.
8. Craigie, R. & Mizuuchi, K. (1985) *Cell* **41**, 867–876.
9. Shapiro, J. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1933–1937.
10. Mizuuchi, K. (1984) *Cell* **39**, 395–404.
11. Craigie, R. & Mizuuchi, K. (1985) *J. Biol. Chem.* **260**, 1832–1835.
12. Chaconas, G., Gloor, G. & Miller, J. L. (1985) *J. Biol. Chem.* **260**, 2662–2669.
13. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
14. Dixon, N. E. & Kornberg, A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 424–428.
15. Rouvière-Yaniv, J. & Gros, F. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3428–3432.
16. Craigie, R., Mizuuchi, M. & Mizuuchi, K. (1984) *Cell* **39**, 387–394.
17. Howe, M. M. & Schumm, J. W. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 337–346.
18. Schumm, J. W. & Howe, M. M. (1981) *Virology* **114**, 429–450.
19. Reed, R. R. (1981) *Cell* **25**, 713–719.
20. Krasnow, M. A. & Cozzarelli, N. R. (1983) *Cell* **32**, 1313–1324.
21. Abremski, K., Hoess, R. & Sternberg, N. (1983) *Cell* **32**, 1301–1311.
22. Harshey, R. M. & Bukhari, A. I. (1983) *J. Mol. Biol.* **167**, 427–441.