Proc. Nat. Acad. Sci. USA
Vol. 72, No. 11, pp. 4399–4403, November 1975
Biochemistry

Influence of insertions on packaging of host sequences covalently linked to bacteriophage Mu DNA

(*Escherichia coli*/agarose gel electrophoresis/DNA heteroduplexes/restriction endonucleases/phage Mu genetics)

A. I. BUKHARI*† AND A. L. TAYLOR‡

* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724; and ‡ Department of Microbiology, University of Colorado Medical Center, Denver, Colo. 80220

Communicated by J. D. Watson, September 9, 1975

ABSTRACT **Insertions in bacteriophage Mu DNA have been identified. These insertions are responsible for at least seven *X* mutations, all of which eliminate essential Mu functions. The insertions are about 800 base pairs long and are located to the left of the cleavage site of restriction endonuclease *Eco*RI, near the immunity end of Mu DNA. We have found that such insertions cause a reduction in the length of nonhomologous terminal sequences which are seen as split ends in denatured and renatured Mu DNA molecules. These heterogeneous sequences apparently arise from packaging of host DNA from maturation precursors in which Mu and host DNA are covalently linked. We infer that a single Mu genome length is too short to be cut during morphogenesis, and thus some host DNA is packaged into mature virions. Since the insertions increase the length of Mu DNA, they decrease the amount of host DNA needed for packaging.**

The temperate bacteriophage Mu is characterized by the highly promiscuous integration of its DNA into the genome of its host bacterium *Escherichia coli* (1–3). The prophage Mu DNA can be excised precisely from the different integration sites, and as a result wild-type function of the gene into which Mu was inserted can be restored (4). This reversal of Mu integration can be seen if the prophages contain the *X* mutations, which block the lethal and essential phage functions.

Mu *X* mutants revert to wild type at a frequency of about 10^{-8}. However, no *X* amber mutants have been found. Thus, we proposed that the *X* mutations are insertions (4). We have investigated the nature of the *X* mutations by agarose-gel electrophoresis of Mu DNA fragments produced by specific endonucleases, and by electron microscopic examination of Mu DNA heteroduplexes. These studies have shown that the *X* mutations are insertions near the immunity (*c*) end of Mu DNA. In one *X* mutant, the insertion is demonstrated to be of about 720 base pairs, located 4 kilobases (kb) away from the *c* end.

The isolation of insertions in Mu DNA has allowed us to analyze further some aspects of Mu DNA structure. Mu DNA, a double-stranded linear duplex of about 37 kb (5), has some interesting features. When it is denatured and reannealed, two types of homoduplex molecules are obtained (Fig. 1). All the molecules show single-stranded tails or split ends, measuring about 4% of the molecular length. The split ends, which apparently represent sequences picked up randomly from the *E. coli* chromosome, are always found at the *S* gene end, called here the SE end, of Mu DNA (6–8). Some molecules show another nonrenaturable region, termed the G bubble, near the SE end (9). Recently, we observed that the *c* end of Mu DNA is also not fixed, and probably varies in length by about 100 base pairs (10, 11).

Abbreviation: kb, kilobase.

† To whom correspondence should be directed.

One plausible model for the origin of the heterogeneity of the SE end is that Mu DNA is packaged into phage heads from maturation precursors which contain both Mu DNA and host sequences. This hypothesis implies that the Mu genome alone is too short to be cut during the maturation process and therefore some host DNA is packaged along with the phage DNA.

The headful packaging model leads to the prediction that if some foreign DNA is inserted into Mu DNA there will be a decrease in the length of the split end corresponding to the size of the insertion. By examining Mu DNA containing an *X* insertion, we have found that this prediction is fulfilled. We report here that the presence of an insertion in Mu DNA causes a reduction in the length of the split end by an amount equal to the length of the insertion.

METHODS

Genetic Procedures. The bacterial strains, all derivatives of *E. coli* K12, the bacteriophage Mu strains, and the isolation and characterization of the *X* mutants of prophage Mu were reported recently (4).

Biochemical Procedures. To prepare Mu particles, the phage lysates were precipitated with polyethylene-glycol, the precipitates were resuspended in Mu buffer (100 mM NaCl, 20 mM Tris-base, 0.25 $MgSO_4$, 0.1 mM $CaCl_2$, and 0.1% gelatin pH 7.5), and the phage particles were purified by cesium chloride density-gradient centrifugation (10, 11). To analyze Mu DNA molecules with the specific restriction endonucleases, 1–2 μg of the phenol-extracted DNA molecules were digested with the enzymes, and the fragments generated were resolved by electrophoresis through agarose-slab gels in the presence of 0.5 μg of ethidium bromide per ml, as described by Sharp *et al.* (12). The enzymes, *Hin*dIII (from *Hemophilus influenzae* Rd) and *Eco*RI (from *E. coli* RY13), were prepared by R. J. Roberts' group at the Cold Spring Harbor Laboratory.

Electron Microscopic Procedures. Denaturation and renaturation of Mu DNA, and mounting of renatured DNAs by the formamide method, were done essentially as de-

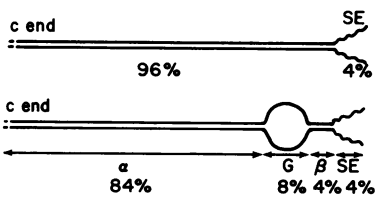

FIG. 1. The two types of molecules observed after denaturation and self-renaturation of Mu DNA. The split end is designated SE. For measurements of different segments of Mu DNA, see Table 1.

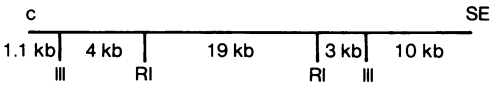

FIG. 2. Map of *Hin*dIII and *Eco*RI cleavage sites in Mu DNA. Molecular lengths of the fragments are given in kilobases (10).

scribed by Sharp *et al.* (13). Micrographs were taken at a magnification of 10,000 with an RCA model EMU4. Length measurements were made from tracings of photographic negatives enlarged with a Scherr-Tumico optical comparator. All measurements are given in kilobases (kb), a length corresponding to 1000 bases or base pairs, and with reference to double-stranded phage PM2 DNA (9.54 kb) and single-stranded phage ϕX174 DNA (5.15 kb) as internal calibration standards.

RESULTS

Isolation of phage Mu particles with *X* mutations

The *X* mutants of Mu are unable to grow in Mu-sensitive bacteria. These mutants are readily obtained by plating *E. coli* cells carrying Mu *cts*62, a heat-inducible derivative of Mu, at temperatures above 42°. At these temperatures most of the lysogenic cells are killed because of induction of Mu *cts*62, which presumably makes a thermosensitive repressor owing to a mutation in the immunity gene *c*. As we showed earlier, some of the survivors at high temperatures contain Mu *cts*62 prophages with the *X* mutations (4). The Mu *cts*62 *X* lysogens can be recognized by their ability to be cured of Mu DNA. However, these lysogens do not yield any plaque-forming phage and it is not possible to isolate and assay the Mu *X* particles normally. Therefore, we tried to rescue the *X* mutants with the wild-type Mu *cts*62 phage. The F′ *pro*$^+$ *lac* episomes containing Mu *cts*62 *X* prophages in the *Z* gene of the *lac* operon were transferred to strain BU165, which has a deletion of the *proA,B-lac* region and contains a Mu *cts*62 X^+ prophage. These strains were thus dilysogenic for Mu, having one copy of Mu with the *X* mutation and one without. The dilysogens were heat-induced and the phage particles were purified. The phage preparations obtained from the dilysogens would be mixtures of the *X* mutant particles and the X^+ particles, if the *X* mutants can grow normally with the helper phage.

Phage particles from these preparations banded as a single species in cesium chloride density gradients. However, the restriction endonuclease analysis showed that DNA molecules of these particles were different from wild-type Mu DNA.

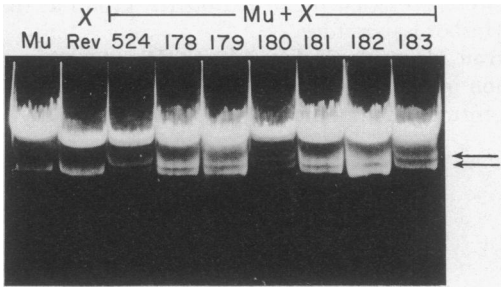

FIG. 3. Electrophoresis of *Eco*RI digests of Mu DNA on 1.4% agarose gels. The DNAs of phage preparations from Mu$^+$:Mu *X* dilysogens, identified by numbers, show two bands in the region where Mu$^+$ DNA alone and an X^+ revertant DNA (obtained from mutant *X*5004 in strain BU524) give only a single band. Heteroduplex studies reported below were done on BU180 DNA.

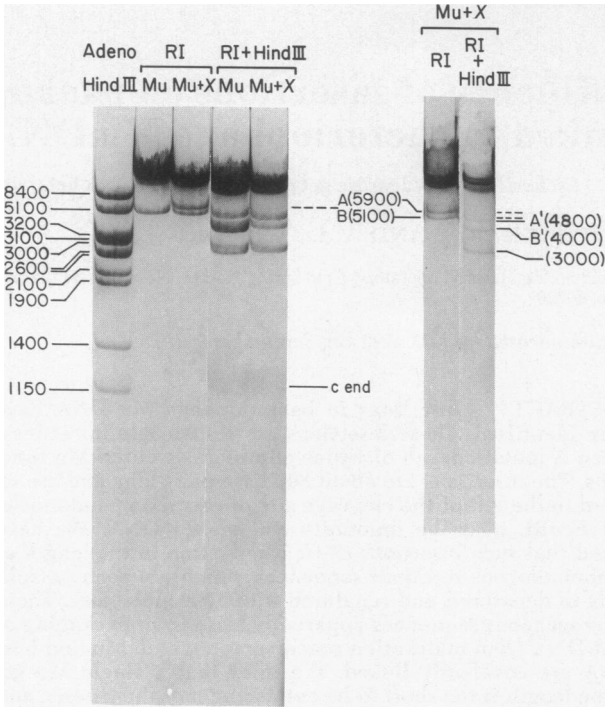

FIG. 4. Electrophoresis of *Eco*RI and *Hin*dIII digests of Mu DNA on 1.4% agarose gels. The adenovirus (Ad2) DNA fragments produced by *Hin*dIII were used as markers to calculate the sizes of Mu fragments. Molecular weights of the 12 adeno fragments, some of which are not resolved on this gel, were determined by R. J. Roberts (personal communication). Samples labeled Mu + *X* are from preparations of BU524, a dilysogen for Mu *cts*62 and Mu *cts*62 *X*5004. The diffuse immunity-end fragment, 1100 ± 100 base pairs, produced by *Hin*dIII, is labeled *c* end. The digestion of the 5100 base pair *Eco*RI fragment by *Hin*dIII is incomplete in the gel at the left. At right, cleavage of the *c*-end *Eco*RI fragments, from a Mu$^+$:Mu *X* preparation, by *Hin*dIII is shown more clearly. *Hin*dIII reduces the size of A and B, shown by broken lines, by 1100 base pairs, generating fragments A′ and B′. The diffuse 1100 base pair fragment is too low in concentration to be seen in this case.

Analysis of Mu DNAs with restriction endonucleases

The specific endonuclease *Eco*RI cleaves Mu DNA at two places, generating a relatively small fragment of about 5000 base pairs and two larger fragments (10). These cleavage sites are shown in Fig. 2. The small fragment, which comes from the *c* end of Mu DNA, is cleanly separated from the other two fragments on agarose gels. Mu$^+$:Mu DNA molecules from seven independent *X* mutants were digested with *Eco*RI. It is shown in Fig. 3 that *Eco*RI digests of these Mu$^+$: Mu *X* DNA preparations had two, instead of one, small fragments. One migrated on gels as the *c* end *Eco*RI fragment does, whereas the other had significantly lower mobility. This indicated that the Mu *X*:Mu$^+$ preparations had two species of the *Eco*RI fragment of the *c* end of Mu DNA. One originated from the Mu$^+$ DNA molecules, whereas the larger molecular weight species, with slower mobility, arose from the Mu *X* molecules. This interpretation was confirmed by further cleavage of the DNA molecules with the endonuclease *Hin*dIII. *Hin*dIII cuts Mu DNA at a site closer to the immunity end, releasing heterogeneous fragments of about 1100 base pairs (Fig. 2). As shown in Fig. 4, *Hin*dIII cleaved both *Eco*RI fragments, generating two fragments which were shorter than the original fragments by about 1100 base pairs and the heterogeneous 1100 base pair frag-

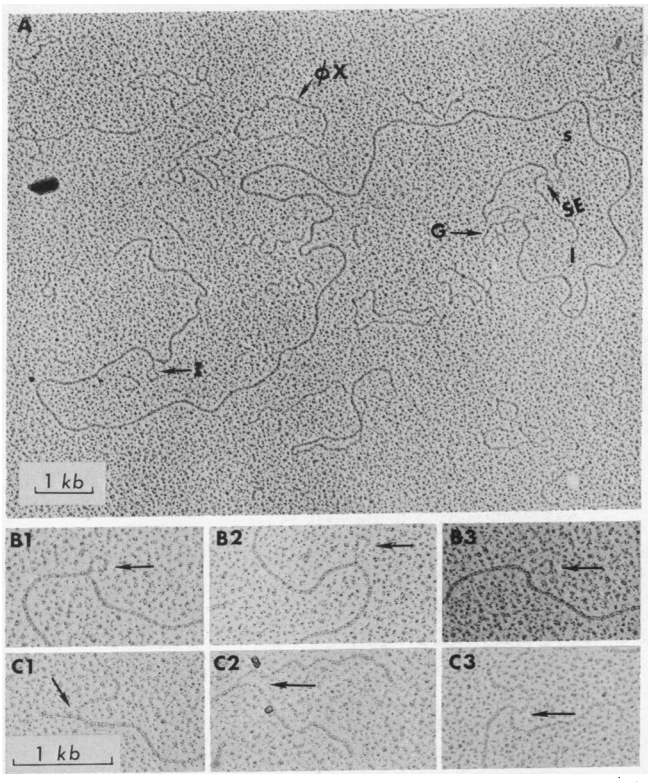

FIG. 5. Electron micrographs showing different structures of Mu DNA. (A) Mu$^+$:Mu *X* heteroduplex showing insertion loop (I), a somewhat twisted G bubble, and the split ends (SE). The long and short arms of the split end are labeled l and s respectively. A φX174 single-strand circle is also shown. (B) Some examples of the *X* insertion loops. A double-stranded stem can be seen in B1. (C) The split ends seen in renatured molecules of Mu$^+$:Mu *X* preparations. C1 shows a short split end, presumably in a Mu *X*:Mu *X* homoduplex; C2 shows a long split end, presumably a Mu$^+$:Mu$^+$ homoduplex; and C3 shows the split end in a Mu$^+$:Mu *X* heteroduplex. Frames B1 to C3 are shown at the same magnification.

ment from the *c* end. A comparison of the Mu *Eco*RI fragments with the marker adeno DNA fragments of known molecular length showed that the Mu$^+$ *c*-end *Eco*RI fragment was of 5100 base pairs and the Mu *X* fragment was of about 5900 base pairs. It was inferred, therefore, that the Mu *X* mutants contained an insertion of about 800 base pairs to the left of the *Eco*RI cut near the *c* end of the Mu DNA.

Electron microscopic examination of *X* insertions

The DNA molecules obtained from the Mu$^+$:Mu *X* preparations were denatured, allowed to renature, and then examined electron microscopically. An insertion loop could be readily seen in some molecules near the *c* end, which can be recognized by the absence of the single-stranded tails of the split end. These molecules were thus Mu$^+$:Mu *X* heteroduplexes. The molecules in which no insertion loop could be detected were presumably Mu$^+$:Mu$^+$ or Mu *X*:Mu *X* homoduplexes. The insertion loops in the Mu$^+$:Mu *X* heteroduplexes are shown in Fig. 5. Molecular lengths of the different structures in Mu$^+$:Mu$^+$ homoduplexes and Mu$^+$:Mu *X* heteroduplexes, originating from the mutant *X* of strain BU180, were systematically measured. The results given in Table 1 for the α, G, β, and *c* to SE segments agree with earlier measurements (7) and thus confirm the physical dimensions of these structures. Measurement of 37 heteroduplexes showed the insertion to have a mean length of 720 ± 110 base pairs. The insertion was located almost exactly 4000 base pairs away from the *c* end. Thus the insertion was about 1000 base pairs to the left of the *Eco*RI cleavage site, which is consistent with the endonuclease data.

Table 1. Length of structures found in heteroduplex and homoduplex molecules of Mu$^+$ and Mu *X* DNA

DNA segment*	No. measured	Mean length (kb)
α	23	30.7 ± 1.1
G	32	2.9 ± 0.4
β	24	1.6 ± 0.1
c to SE†	12	34.9 ± 1.2
SE‡	—	—
c to I	72	4.0 ± 0.2
I	37	0.72 ± 0.11

* The symbol I denotes the insertion in Mu *X*5004. All other symbols are identified in Fig. 1.
† Measured in molecules that lack a G inversion bubble.
‡ SE measurements are presented separately in Table 2 and Fig. 6.

Analysis of the split ends

The two arms of the single-stranded tails of the split end of a Mu DNA molecule are generally not equal in length. We measured both arms of the split end in a large number of molecules from wild-type Mu DNA preparations and from Mu$^+$:Mu *X* DNA preparations. These measurements are presented in Table 2 and Fig. 6. We found that the split end in wild-type Mu DNA can be as short as 500 base pairs or as long as 3200 base pairs. The mean length of the split ends in wild type is 1450 ± 500 base pairs. This large deviation in mean length is in agreement with the deviation reported by Daniell *et al.* (7). As shown in the histogram in Fig. 6, the distribution of the split-end lengths is clearly different in Mu$^+$:Mu *X* preparations. The difference in the lengths of the split-end arms of confirmed Mu$^+$:Mu *X* heteroduplexes, in which an insertion loop can be seen, is very sharply defined. One arm in these cases can be as short as 200 base pairs, and it is often much shorter than the other arm (Fig. 5, C3). The ratio of long/short split-end lengths was >2.0 in 56% of heteroduplex molecules (50 measured), as compared to 6% in Mu$^+$:Mu$^+$ homoduplexes (64 measured). The mean

Table 2. Molecular lengths of single-stranded split ends in renatured Mu DNA

DNA	Split ends measured*	No. measured	Range of lengths (kb)	Mean length (kb)†
Mu$^+$:Mu$^+$ homoduplex	s	64	0.5–2.4	1.22 ± 0.42
	l	64	0.9–3.2	1.69 ± 0.49
	s and l	128	0.5–3.2	1.45 ± 0.51
Mu$^+$:Mu *X*‡ heteroduplex	s	50	0.2–2.3	0.80 ± 0.46
	l	50	1.0–3.0	2.00 ± 0.52
	s and l	100	0.2–3.0	1.40 ± 0.77
Mu$^+$:Mu$^+$, Mu$^+$:Mu *X*, and Mu *X*:Mu *X*	s	104	0.2–2.3	
	l	104	0.4–3.3	
	s and l	208	0.2–3.3	

* The symbols s and l denote the shorter and longer arms of each split-end pair, respectively.
† Standard deviations of the means are indicated. Mean lengths are not computed for the mixed population of homoduplex and heteroduplex molecules.
‡ Mu *X*5004 from strain BU180.

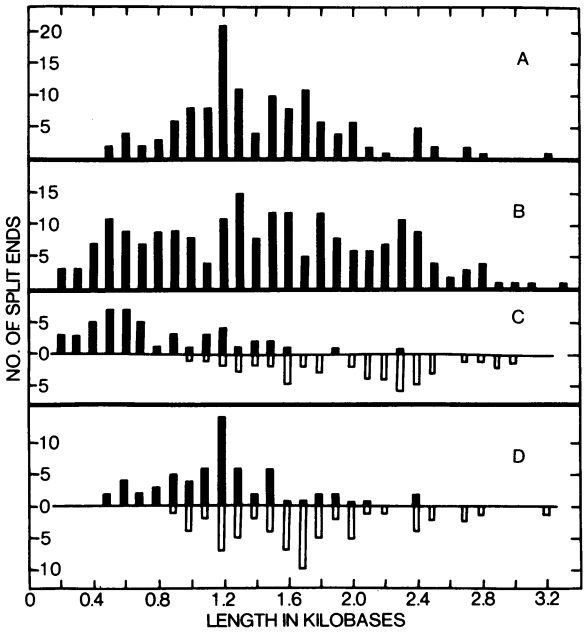

FIG. 6. Histograms of split-end length distribution in renatured Mu DNA molecules. (A) Lengths of the split-end arms of wild-type Mu *cts*62, the progenitor of the *X* mutants. (B) Lengths of the split-end arms of DNA populations containing both Mu *cts*62 (a strain carrying nonlethal mutations *gov* and *mom*) and Mu *cts*62 *X*5004 molecules from strain BU180. (C) Lengths of the split ends in known Mu^+:Mu *X* heteroduplexes, in which insertion loops can be seen. Each split end was arbitrarily divided into a long arm and a short arm, according to length. The long arms, measured independently in each split end, are represented by open bars; the short arms are shown by solid bars. (D) Lengths of the split ends in Mu *cts*62 molecules. As in C, the long arms are shown by open bars and the short arms by solid bars.

length of the small arms in the Mu^+:Mu *X* heteroduplexes was found to be 800 base pairs, about 650 base pairs shorter than the mean length of the split ends in Mu^+:Mu^+ homoduplexes. This difference in mean lengths approaches the length of the *X* insertion, which is about 720 base pairs.

DISCUSSION

We have demonstrated that the *X* mutations of mutator phage Mu are caused by insertions. These insertions are located to the left of the *Eco*RI cleavage site near the immunity end, the *c* end, of Mu DNA. The *Eco*RI fragment from the *c* end consists of about 5100 base pairs. This fragment is 5900–6000 base pairs long in the *X* mutants. Thus there is an addition of about 800 base pairs near the *c* end in these mutants. In one of the *X* mutants, the size of the insertion has been determined by electron microscopic analysis of heteroduplexes between the mutant and the wild-type Mu DNA. This insertion is about 720 base pairs long, and the distance of the insertion site from the *c* end is 4000 base pairs, or 11% of the molecular length of Mu DNA.

The *X* mutants of Mu have an interesting phenotype. Mu prophages with *X* mutations can be precisely excised from the host chromosome, and cells cured of Mu DNA can be obtained (4). The seven *X* mutations examined so far are all insertions of approximately the same size, apparently located in the early regulatory region of the Mu genome. However, we have evidence that all of the insertions are not located at the same site. We can detect differences between the *X* mutants with *Hpa*II endonuclease from *H. parainfluenzae* (4). Two *X* mutant DNAs have been heteroduplexed with each other and show insertion loops which are several hundred base pairs apart. In at least some heteroduplex molecules a double-stranded stemlike structure can be seen at the base of the insertion loops. This may mean that the *X* insertions are flanked by inverted repeat sequences, which specifically pair with each other. It is also possible that the *X* insertions are related to the IS1 insertion element, which is about 800 base pairs long and occurs spontaneously in *E. coli* (17, 18).

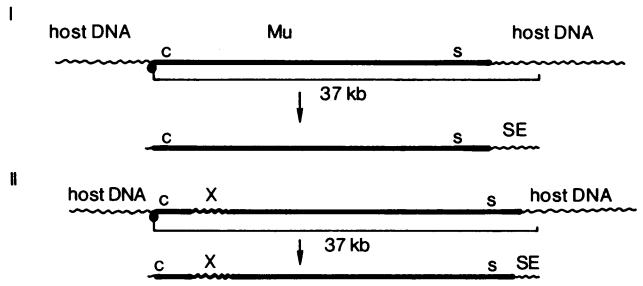

FIG. 7. The DNA packaging model for the origin of heterogeneous host sequences in Mu DNA. Packaging reaction starts at the *c* end, and folding of Mu DNA occurs. The product of the reaction is a molecule of about 37 kb, slightly larger than the actual Mu genome length. The consequence of an insertion within Mu DNA would be shortening of the host sequences packaged.

Isolation of insertions in Mu DNA has allowed us to test the DNA packaging model for the origin of nonhomologous sequences which are seen as split ends in denatured and renatured Mu DNA molecules. The basic postulates of this model, shown diagrammatically in Fig. 7, are that (*1*) mu DNA is packaged into virion heads from maturation precursors which have covalently linked Mu DNA and host DNA; (*2*) the packaging starts specifically at the *c* end of Mu DNA in a maturation precursor and proceeds unidirectionally toward the other end of Mu; and (*3*) the Mu genome alone cannot fill the heads adequately and therefore some host DNA, covalently linked to Mu DNA, is packaged during morphogenesis. It is known that heterogeneous covalently closed circles containing Mu and host sequences are formed during the productive cycle of Mu (14, 15). It has been pointed out that nonhomologous sequences in mature virions could be derived from these circles (15). However, the mechanism by which these circles are generated and the exact function of these circles in the Mu development are not clear. They could be the replicative intermediates giving rise to the maturation precursors, or they could be the maturation precursors themselves. Genetic evidence indicates that Mu DNA can be attached to different host DNA setments during lytic growth (16). Multiple integration of Mu during lytic growth is probably related to the formation of circles with different host sequences. The circles have been found to increase in number shortly before lysis, and electron microscopic heteroduplex studies show that Mu DNA is connected to host DNA which differs in amount and sequence from circle to circle (B. Waggoner and A. L. Taylor, unpublished results). Thus, the heterogeneous covalently closed circles meet all the criteria for being the source from which the heterogeneity of the SE end of Mu originates.

If Mu maturation involves packaging of a fixed length of DNA, slightly larger than the Mu genome, from precursors which contain both Mu and host sequences, then an insertion within Mu would reduce the length of host DNA needed for efficient packaging. As depicted in Fig. 7, this would cause the packaging reaction to terminate earlier with respect to host DNA, and hence host sequences in mature virions

would be smaller. This prediction is borne out by the measurements reported in this paper. In Mu^+:Mu *X* heteroduplexes, which show an insertion loop, one arm of the split end is often much smaller than the other arm. The mean length of the small arms in the heteroduplexes is less than the mean length of the split ends in wild-type Mu. The difference amounts to about 650 base pairs, approximately equaling the length of the insertion. It can be inferred that the main postulates of the DNA packaging model are essentially correct. We have recently further confirmed the model by examining Mu DNA molecules which contain insertions about 2800 base pairs long. No split ends can be seen in these molecules, and the packaging cut is made in the β region within Mu DNA (A. I. Bukhari and L. Chow, manuscript in preparation). This also helps to explain why deletion mutants of Mu have not been isolated so far. The methods for isolating deletion mutants have been based on detection of differences in the DNA content of phage particles. Since the DNA content of Mu virions seems to remain constant, as implied by the DNA packaging model, deletions would merely increase the contribution of host DNA without changing the density of the phage particles.

A wide variation is found in the length of the split ends in wild-type Mu DNA. We have established that the heterogeneous sequence can be as short as 500 base pairs and as long as 3200 base pairs, although most of the split ends hover around 1450 base pairs. This 6-fold range of split-end lengths suggests that the packaging mechanism is not absolutely precise. A small change in the size of Mu DNA—an addition, a deletion, or a duplication—could also affect the length of the split ends. This possibility has to be considered when different strains of Mu are used.

We acknowledge with appreciation the help of R. J. Roberts, who provided us with endonucleases and other materials. We are very thankful to D. Botstein, E. Daniell, V. Chapman, and S. DeLong for help with the electron microscopic work, to D. Zipser for interesting discussions, and to Susan Froshauer for technical assistance. This work was supported by grants to A.I.B. from the National Science Foundation (GB-43280), the Cystic Fibrosis Foundation, and The Jane Coffin Childs Memorial Fund for Medical Research, and also by grants to A. L.T. and D. Zipser from the National Institutes of Health. A.I.B. holds a Career Development Award of the National Institutes of Health (1-K4-GM-00127-01).

1. Taylor, A. L. (1963) *Proc. Nat. Acad. Sci. USA* **50**, 1043–1051.
2. Bukhari, A. I. & Zipser, D. (1972) *Nature New Biol.* **236**, 240–243.
3. Daniell, E., Roberts, R. & Abelson, J. (1972) *J. Mol. Biol.* **69**, 1–8.
4. Bukhari, A. I. (1975) *J. Mol. Biol.* **96**, 87–99.
5. Martuscelli, J., Taylor, A. L., Cummings, D. J., Chapman, V. A., DeLong, S. S. & Cañedo, L. (1971) *J. Virol.* **8**, 551–563.
6. Daniell, E., Kohne, E. E. & Abelson, J. (1975) *J. Virol.* **15**, 739–743.
7. Daniell, E., Abelson, J., Kim, J. S. & Davidson, N. (1973) *Virology* **51**, 237–239.
8. Daniell, E., Boram, W. & Abelson, J. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 2153–2156.
9. Hsu, M. & Davidson, N. (1974) *Virology* **58**, 229–239.
10. Bukhari, A. I. & Allet, B. (1975) *Virology* **63**, 3–39.
11. Allet, B. & Bukhari, A. I. (1975) *J. Mol. Biol.* **92**, 529–540.
12. Sharp, P., Sugden, B. & Sambrook, J. (1973) *Biochemistry* **12**, 3055–3063.
13. Sharp, P., Hsu, M. T., Ohtsubo, E. & Davidson, N. (1972) *J. Mol. Biol.* **71**, 471–497.
14. Waggoner, B., Gonzales, N. S. & Taylor, A. L. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 1255–1259.
15. Schröder, W., Bade, E. G. & Delius, H. (1974) *Virology* **60**, 534–542.
16. Razzaki, T. & Bukhari, A. I. (1975) *J. Bacteriol.* **122**, 437–442.
17. Fiandt, M., Szybalski, W. & Malamy, M. H. (1972) *Mol. Gen. Genet.* **119**, 223–231.
18. Starlinger, P. & Saedler, H. (1972) *Biochimie* **59**, 177–185.