Correct Integration of Retroviral DNA
In Vitro

Patrick O. Brown,† Bruce Bowerman,†
Harold E. Varmus,*† and J. Michael Bishop**†
* Departments of Microbiology and Immunology
University of California, San Francisco
San Francisco, California 94143
† Departments of Biochemistry and Biophysics
University of California, San Francisco
San Francisco, California 94143
† G. W. Hooper Foundation
University of California, San Francisco
San Francisco, California 94143

Summary

We have developed a cell-free system for studying the integration of retroviral DNA. In our assay, amber mutations in a bacteriophage λ genome that serves as the target for integration are suppressed by integration of an MLV derivative that carries the E. coli supF gene. The structure of the reaction products is that expected from an authentic MLV integration reaction. Linear viral DNA from the cytoplasm of infected cells serves as a precursor, though not necessarily the immediate precursor, to the provirus integrated in vitro. The viral DNA in the infected cell appears to be tightly associated with the enzymatic machinery required for its integration. Supercoiling, chromatin structure, transcription, and replication are not required of the target DNA. Since no high-energy cofactor is necessary, the DNA breakage and joining steps in the integration reaction are probably coupled.

Introduction

Retroviruses replicate via a complex life cycle, during which the viral genome undergoes a remarkable series of information-conserving structural transformations (Varmus and Swanstrom, 1985). Two essential steps of this cycle are the synthesis of a DNA copy of the RNA viral genome and integration of this viral DNA into a chromosome of the infected cell. The resulting provirus can thereafter be transmitted and expressed as a stable genetic element in the cell is established by infection. For example, introduction of naked viral DNA into a permissive host is not sufficient for legitimate proviral integration (Luciw et al., 1984). Even when all the viral proteins necessary for integration are provided by superinfection, plasmid DNA containing the MLV circle junction sequence, when introduced by means other than retroviral infection, does not integrate by a retroviral mechanism (Kriegler and Botchan, 1983; P. Brown, unpublished data). This apparent confinement of integration activity to authentic viral reverse transcripts is consistent with a model in which subviral structures maintain a high degree of order during replication. In particular, the product of the 3′ end of the viral pol gene, which has been shown to be required for integration (Donohower and Varmus, 1984; Schwartzberg et al., 1984; Panganiban and Temin, 1984b), might be active only on DNA in a specific subviral nucleoprotein complex. Indeed, most of the unintegrated viral DNA produced by reverse transcription in MLV-infected cells can be isolated in stable fast-sedimenting nucleoprotein particles that also contain virally encoded proteins (B. Bowerman,
unpublished data). No host cell functions have yet been identified as essential for integration.

There is accumulating evidence for a close similarity between retroviral integration and the intracellular transposition of retrotransposons, a class of eukaryotic transposable elements that includes the yeast Ty elements and the copia family of elements in Drosophila (Varmus, 1983; Boeke et al., 1985; Shiba and Saigo, 1983; Flavell, 1984). Compared to retroviral integration after a synchronized inactivation of the terminal 2 bp thought to be present at each end of the unintegrated linear precursor but is otherwise colinear with the precursor, and it is flanked by a precise 4 bp duplication of the sequence at the target site. The target DNA is otherwise unmodified and unarranged.

Active Precursors for Integration In Vitro

In vivo, proviral integration takes place in the nucleus. Furthermore, results from experiments with SNV support the
Integration of Retroviral DNA In Vitro

Infect 3T3 cells with MLV supF

Harvest cells, lyse with digitonin

Cytoplasmic extract

Nuclear extract

Linear LTR

MIX with λ gtWES concatemers

Integration

Recover DNA, package in vitro

Plate on CC114

Plate on KM392

Recover recombinants only

Recover unreacted λ gtWES as well as recombinants

The simplest explanation for this result is that it follows from the greater abundance of unintegrated viral DNA molecules in the cytoplasm than in the nucleus. As shown in Table 1, integration activity is proportional to the quan-

Figure 1. Flowchart of the In Vitro Integration Protocol

The supF gene is indicated by the shaded area in the LTR of MLV supF. As discussed in Results, the cytoplasmic extract contains only the linear form of viral DNA, whereas the nuclear extract contains the 1-LTR and 2-LTR circular forms as well as the linear form.

Figure 2. Assay for Integration of MLV DNA In Vitro

Integration of an MLV supF provirus into the λ gtWES DNA provides a suppressor for the three amber mutations in the target phage genome. This allows the recombinant phage to make plaques on the normally restrictive supF host. The plaques are blue on this indicator plate because an amber mutation in the host lacZ gene is also suppressed by the supF gene in the provirus.

The idea that the 2-LTR circular form of unintegrated viral DNA is the immediate precursor to the integrated provirus (Panganiban and Temin, 1984a). Circular viral DNA has been detected only in the nucleus of infected cells (Guntaka et al., 1976; Varmus and Swanstrom, 1985). We therefore expected that the nuclear extract would be more active than the cytoplasmic extract in our assay. We were surprised to find that, in fact, the cytoplasmic extract was more active (Table 1). In mixing experiments (Table 2), supplementation of the nuclear extract from MLV supF-infected cells with cytoplasmic extract from uninfected cells failed to enhance the activity of the nuclear extract. Nor did nuclear extract from uninfected cells inhibit the activity of cytoplasmic extract from MLV supF-infected cells. These results argue that the greater activity of the cytoplasmic extract was not simply due to the presence of stimulatory factors in the cytoplasm or inhibitory factors in the nuclear extract.
Figure 3. Sites and Orientation of MLV Proviruses Integrated In Vitro
Flags indicate orientation of proviruses, and point from 5' to 3' relative to the MLV (+) strand. Numbered flags correspond to proviruses whose junctions with λ DNA were sequenced (Figure 4). The heavy lines indicate regions essential for λ growth on a λcA' host. The positions of several λ genes are indicated as landmarks.

Figure 4. Sequences at the λ MLV Junctions
The sequences shown correspond to the plus strand of the MLV DNA. The sequences of the ends of the unintegrated linear form of MLV DNA are shown at the top, for reference. The boxes enclose the 4 bp sequence, present in a single copy in the target DNA, that was duplicated upon MLV integration. Note that in the integrated provirus, the terminal 2 bp of the linear precursor have been lost.

Table 1. Comparison of Nuclear and Cytoplasmic Extracts

<table>
<thead>
<tr>
<th>Linear DNA molecules</th>
<th>Cytoplasmic Extract</th>
<th>Nuclear Extract</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-LTR circles</td>
<td>8 × 10^2</td>
<td>2 × 10^7</td>
</tr>
<tr>
<td>1-LTR circles</td>
<td>None detected</td>
<td>1.7 × 10^6</td>
</tr>
<tr>
<td>PFu on KM392</td>
<td>4.4 × 10^8</td>
<td>2.6 × 10^6</td>
</tr>
<tr>
<td>PFu on CC114</td>
<td>5.7 × 10^2</td>
<td>1.4 × 10^6</td>
</tr>
<tr>
<td>Recombinants</td>
<td>2.5 × 10^3</td>
<td>4.7 × 10^4</td>
</tr>
<tr>
<td>Recombinants - linear molecule</td>
<td>3.1 × 10^3</td>
<td>6.7 × 10^6</td>
</tr>
<tr>
<td>Recombinants - 2-LTR circle</td>
<td>&gt;6.2 × 10^1</td>
<td>3.3 × 10^3</td>
</tr>
<tr>
<td>Recombinants - 1-LTR circle</td>
<td>&gt;6.2 × 10^1</td>
<td>3.9 × 10^2</td>
</tr>
</tbody>
</table>

Table 1: Comparison of Nuclear and Cytoplasmic Extracts

Nuclear and cytoplasmic extracts (100 µl) were prepared from 4 × 10⁷ cells, assayed for unintegrated forms of viral DNA (as shown in Figure 5), and used for integration reactions as described in Experimental Procedures. Calculation of the number of recombinants per reaction was as described in Experimental Procedures. In the calculations for this table, the correction factor for packaging and plating efficiency (designated "E" in the formula shown in Experimental Procedures) was 0.13.

Figure 4: Sequences at the λ MLV Junctions
The sequences shown correspond to the plus strand of the MLV DNA. The sequences of the ends of the unintegrated linear form of MLV DNA are shown at the top, for reference. The boxes enclose the 4 bp sequence, present in a single copy in the target DNA, that was duplicated upon MLV integration. Note that in the integrated provirus, the terminal 2 bp of the linear precursor have been lost.

Figure 3: Sites and Orientation of MLV Proviruses Integrated In Vitro
Flags indicate orientation of proviruses, and point from 5' to 3' relative to the MLV (+) strand. Numbered flags correspond to proviruses whose junctions with λ DNA were sequenced (Figure 4). The heavy lines indicate regions essential for λ growth on a λcA' host. The positions of several λ genes are indicated as landmarks.

Requirements and Optimal Conditions for Integration
A time course of the integration reaction (Figure 6) demonstrates that the yield of recombinant products increases progressively with continued incubation for about 1 hr. The apparent decline in the rate of the reaction with time is probably at least partly due to thermolability of the activity, since activity diminishes appreciably when the cellular extract is preincubated at 37°C (data not shown).

Table 3 presents the results of our preliminary characterization of the conditions that affect MLV integration in vitro. EDTA abolished activity, implying that a divalent cation is essential; we have not yet tested other divalent cations in place of Mg²⁺. The optimal K⁺ concentration was about 150 mM; activity was decreased very little at 75 mM KCl, and was about 10-fold lower at a KCl concentration of 300 mM. Polyethylene glycol at 5% stimulated activity 4-fold, but was not required. The pH optimum was broad, extending at least from pH 6.6 to pH 8.6. Roughly equivalent activity was observed at an incubation temperature of 30°C or 37°C, but no detectable integration occurred in a 1 hr incubation at 0°C. The yield of recombinants was not affected by heating the DNA recovered from a reaction to 60°C for 10 min.
Table 2. Test for Interaction between Nuclear and Cytoplasmic Components

<table>
<thead>
<tr>
<th>Source of Extracts</th>
<th>Pfu on KM392</th>
<th>Pfu on CC114</th>
<th>Recombinants %</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 μl Infected Cytoplasm + 50 μl Uninfected Cytoplasm</td>
<td>4.2 x 10⁶</td>
<td>2601</td>
<td>1.1 x 10⁵</td>
</tr>
<tr>
<td>50 μl Infected Cytoplasm + 50 μl Uninfected Nuclei</td>
<td>3.4 x 10⁶</td>
<td>4789</td>
<td>2.4 x 10⁴</td>
</tr>
<tr>
<td>50 μl Infected Nuclei + 50 μl Uninfected Nuclei</td>
<td>1.3 x 10⁶</td>
<td>305</td>
<td>4.0 x 10⁴</td>
</tr>
<tr>
<td>50 μl Infected Nuclei + 50 μl Uninfected Cytoplasm</td>
<td>2.1 x 10⁶</td>
<td>418</td>
<td>3.4 x 10⁴</td>
</tr>
</tbody>
</table>

Preparation of extracts, integration reactions, and calculation of the number of recombinants per reaction were as described in Experimental Procedures, except that uninfected cellular extracts were prepared from NIH 3T3 cells that were cultivated in the absence of MLV supF producer cells. For each reaction, extracts were mixed as indicated before addition to the remainder of the reaction mixture. In the right-hand column, results are expressed as a percent of the number of recombinants obtained with cytoplasmic extracts alone.

Figure 5. Quantity and Structure of Viral DNA in Nuclear and Cytoplasmic Extracts

Viral DNA was prepared from nuclear and cytoplasmic extracts of NIH 3T3 cells that were acutely infected by cocultivation with an MLV supF producer line. DNA samples were resolved by agarose gel electrophoresis, then transferred to a nitrocellulose filter. MLV DNA was identified by hybridization, using as a probe the plasmid p8.2 (a complete, permuted clone of the MLV 1-LTR circle), labeled by nick translation. Lanes: (1) viral DNA from 10 μl of cytoplasmic extract; (2) viral DNA from 40 μl of nuclear extract; (3-8) 160, 80, 40, 20, 10, and 5 μl, respectively, of p8.2 plasmid DNA digested with HindIII to generate 8.2 and 4.3 kb fragments as standards for quantitation. Note that 1 μg of DNA is equivalent to 1 x 10⁵ copies of full-length viral DNA. The intensities of the viral DNA bands were compared to the standards in lanes 3-8. The results, adjusted to represent an extract volume of 100 μl, are presented in Table 1. In the gel shown here, viral DNA was overloaded with respect to linears to facilitate quantitation of the circular forms.

Pretreatment of the cell extract for 5 min at 37°C with SDS and proteinase K or with 5 mM N-ethylmaleimide abolished activity, whereas pretreatment with RNAase A had little effect. Thus protein, but probably not RNA, is required for activity.

Assessment of the Nucleotide Requirements for Activity and Partial Purification of the Active Nucleoprotein Complex

To assess the role of nucleoside triphosphates in integration, we used Bio-gel A5m (Bio-Rad) gel-exclusion chromatography to remove the endogenous triphosphates from our cell extracts. Using a luciferase assay (Cheer et al., 1974), we determined that the residual ATP concentration in the column-purified extract was about 1.5 x 10⁻³ M. To further assess the completeness of nucleotide removal, [³²P]dCTP was mixed with the extract before loading on the column. Less than one part in 5 x 10⁵ was detectable in the active, excluded volume pool. This degree of purification also translates into an ATP concentration of about 1.5 x 10⁻³ M, yet no additional ATP was required for integration when this pool was assayed (Table 3). Furthermore, no other nucleoside triphosphate (NTP) or deoxynucleoside triphosphate (dTTP) was required for integration when this fractionated extract was used.

Gel-exclusion chromatography through a Bio-gel A5m column also provided a substantial purification of the in-
<table>
<thead>
<tr>
<th>Condition</th>
<th>% of Standard Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>100</td>
</tr>
<tr>
<td>- Mg. + EDTA</td>
<td>&lt;0.2</td>
</tr>
<tr>
<td>75 mM KCl</td>
<td>89</td>
</tr>
<tr>
<td>300 mM KCl</td>
<td>11</td>
</tr>
<tr>
<td>- PEG</td>
<td>24</td>
</tr>
<tr>
<td>pH 6.5, PIPES</td>
<td>119</td>
</tr>
<tr>
<td>pH 6.5, Trit</td>
<td>0.02</td>
</tr>
<tr>
<td>30°C incubation</td>
<td>107</td>
</tr>
<tr>
<td>0°C Incubation</td>
<td>&lt;0.2</td>
</tr>
<tr>
<td>+ Proteinase K, SDS</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>+ 5 mM N-ethylmaleimide</td>
<td>&lt;2</td>
</tr>
<tr>
<td>+ RNase A</td>
<td>78</td>
</tr>
<tr>
<td>Heated to 68°C after reaction</td>
<td>93</td>
</tr>
<tr>
<td>- ATP, rNTP, creatine phosphate</td>
<td>216</td>
</tr>
<tr>
<td>- ATP, rNTP, dNTP, creatine phosphate</td>
<td>200</td>
</tr>
<tr>
<td>- pSPCJ, 10^5 copies</td>
<td>78</td>
</tr>
<tr>
<td>Naked MLVsupF DNA</td>
<td></td>
</tr>
<tr>
<td>+ wt MLV-infected cell extract</td>
<td>0.04*</td>
</tr>
<tr>
<td>MLVsupF producer cell extract alone</td>
<td>&lt;0.02</td>
</tr>
</tbody>
</table>

*Except for the indicated modifications, reaction conditions were as described in experimental procedures. For most of the conditions tested, reactions were performed in duplicate or triplicate, in parallel with duplicate or triplicate reactions using standard conditions. The number of recombinant plaques scored from each set of standard reactions ranged from 338 to 6272, and except where the yield was less than 5% of standard, between 66 and 7340 recombinant plaques were scored from each of the reactions using modified conditions.

Table 3. Conditions Affecting MLV Integration

Discussion

We have developed an extremely sensitive in vitro system for studying retroviral integration. Starting with a supercoiled or otherwise modified DNA molecule might be a preferred target for integration. The lack of a requirement for dNTPs or NTPs implies that neither transcription nor extensive DNA replication per se is required for covalent linkage of the provirus to the target DNA. Clearly, generation of the flanking 4 bp repeat requires DNA synthesis, but formation of a stable linkage between provirus and target DNA does not. Thus we propose that the initial covalent linkage at each junction is on one strand only and that its formation is not intrinsically coupled with DNA synthesis (Figure 7). The DNA synthesis required to produce the mature double-stranded structure of the flanking repeat can presumably take place in E. coli after infection with the initial recombinant phage.

Characteristics of the DNA Target for Provirial Integration

Since naked, linear (relaxed) DNA could act as a target for integration, using an extract that contained no detectable histones (data not shown), the DNA target for integration need not be supercoiled nor assembled as chromatin. It remains possible, however, that a supercoiled or otherwise modified DNA molecule might be a preferred target for integration. The lack of a requirement for dNTPs or NTPs implies that neither transcription nor extensive DNA replication per se is required for covalent linkage of the provirus to the target DNA. Clearly, generation of the flanking 4 bp repeat requires DNA synthesis, but formation of a stable linkage between provirus and target DNA does not. Thus we propose that the initial covalent linkage at each junction is on one strand only and that its formation is not intrinsically coupled with DNA synthesis (Figure 7). The DNA synthesis required to produce the mature double-stranded structure of the flanking repeat can presumably take place in E. coli after infection with the initial recombinant phage.

It is interesting that experiments in vivo show that integration occurs at least preferentially in cells that are actively replicating their DNA (Varmus et al., 1977). Perhaps, in the cell, replication physically exposes the DNA to the integration machinery—unnecessary with our naked DNA target. Alternatively, factors present only during S phase may be required. Since our extracts from unsynchronized starting with an extract from 4 x 10^6 cells, containing about 5 x 10^7 viral DNA molecules, yields more than 10^5 recombinants, of which over 1000 can be recovered as blue plaques. Our assay is highly specific. Under our standard reaction conditions, the only recombination event we have seen giving rise to a blue plaque in this assay produces a provirus whose structure accurately duplicates that of an in vivo provirus.

The molecular structure of the junction between an MLV provirus and host DNA is distinctive (Dhar et al., 1980, Shoemaker et al., 1981). All of the in vitro recombinants that we analyzed showed this same characteristic structure. It seems quite unlikely that an artificial mechanism could account for this consistent pattern. We therefore conclude that retrovirus integration in our in vitro system accurately mimics the in vivo integration mechanism.

The yield of recombinant products increases progressively during incubation of λ gtWES DNA with the active cell extract. When the cell extract is pretreated with SDS and proteinase K or N-ethylmaleimide, or if the incubation is performed on ice or in the presence of EDTA, integration activity is abolished. Thus the recombination event occurs during the incubation of the λ gtWES DNA target with the cell extract and is dependent upon protein and a divalent cation. The recombinant product is stable to extensive proteolytic digestion, followed by phenol and chloroform extraction and heating to 68°C in water for 10 min. The junction formed between provirus and λ DNA during the incubation with cell extract is therefore protein-independent and probably covalent, presumably a normal 3'--3' phosphodiester bond.

Such a model is also supported by the apparent inability of the putative integration machinery to interact with analogues of the unintegrated viral DNA (table 3). Naked MLVsupF DNA, isolated from either the cytoplasm or nucleus of infected cells, was not efficiently integrated even when mixed with cytoplasmic extract from wild-type MLV-infected cells. Furthermore, there was little or no inhibition of integration by up to 10^6 copies of a plasmid, pSPCJ, containing the MLV circle junction sequence.

Discussion

We have developed an extremely sensitive in vitro system for studying retroviral integration. Starting with 3 x 10^6 target genomes of bacteriophage λ DNA, we can generally detect a signal (i.e., blue plaques) from any reaction that yields 100 or more MLV λ recombinants. The efficiency of the in vitro reaction under standard conditions is such that between 0.2% and 1% of the retroviral DNA molecules in a cellular extract become integrated into λ DNA during a 1 hr incubation. Thus a typical experiment
cells are competent for integration, the apparent requirement for DNA synthesis in vivo cannot be a reflection of inhibitory factors specific to other phases of the cell cycle. The in vitro reaction showed no apparent target sequence specificity. No consistent features could be identified in the sequences surrounding the integration sites. Furthermore, the integration sites appear to be randomly distributed throughout the dispensable portion of the λ gtWES genome. This observation is similar to the results others have reported from analysis of in vivo proviruses (Varmus, 1983). However, it has been suggested that, in vivo, retroviral DNA preferentially integrates in the upstream portion of actively transcribed genes or near DNAase-hypersensitive sites (Rohdewohld et al., 1987). Moreover, there is evidence that not all regions of the genome are equally favored for integration in vivo (King et al., 1985). The bias in target-site selection in vivo may reflect differences in accessibility of DNA in vivo, not duplicated with the naked target DNA in our in vitro system. Alternatively, there may be a subtle sequence selectivity not detected in our analysis. The possibility that nicks or gaps or other modifications of the target DNA are required for integration remains to be investigated.

The Precursor to the Integrated Provirus

Only the linear form of viral DNA was detectable in the cytoplasm, yet the cytoplasm was a better source of active proviral precursors than the nucleus. Furthermore, integration activity in various cell extracts paralleled the quantity of linear viral DNA in the extracts, rather than the quantity of either of the circular forms. Thus the linear MLV DNA appears to be active as a precursor for integration in vitro, just as it is in vivo. We have not, however, determined whether the linear DNA structure is the immediate precursor to the integrated provirus, or whether it must first be circularized.

The results of Panganiban and Temin (1984a) with SNV would imply that integration can occur in vivo via a 2-LTR circular intermediate. Since circular forms are found only in the nucleus (Figure 5 and Table 1), a factor or condition restricted to the nucleus may be required for circularization to occur. However, in our in vitro system, integration was efficiently carried out by the cytoplasmic extract and was not greatly stimulated by addition of nuclear extract. A simple explanation for this result would be that circularization of the viral DNA is not required for integration in our system. Alternatively, if circularization is required, our findings could indicate contamination of the cytoplasmic extract with a small but sufficient amount of nuclear extract. It is also possible that the apparent requirement for entry into the nucleus before circularization can occur might actually reflect a requirement for target DNA as a cofactor in this process. Such a requirement would be satisfied in our in vitro system, without the need to supply nuclear extract. Experiments are in progress to determine whether circularization of the linear DNA occurs in our in vitro system.
What is the Source of Energy for Formation of the New Covalent Bonds in Integration?

Integration necessarily depends on the formation of new covalent bonds linking the provirus to the host DNA. Many precedents can be found among simple recombination mechanisms for a concerted breakage and rejoining reaction, requiring no external energy source (e.g., Reed and Grindley, 1981; Vetter et al., 1983). However, formation of the bacteriophage Mu transposition intermediate—of the biochemically characterized recombination reactions, perhaps the most similar to retroviral integration—appears to require ATP (Mizuuchi, 1983).

Our results strongly suggest that MLV integration does not require added ATP or indeed any other energy-providing cofactor. Phosphodiester bond exchange via a protein–DNA intermediate in a concerted DNA breakage and joining event could obviate any requirement for an external energy source in integration per se (Reed and Grindley, 1981; Craig and Nash, 1983; Gronostajski and Sadowski, 1985). We therefore favor this concerted mechanism for integration over one involving uncoupled DNA cleavage and ligation steps.

The source of energy for formation of the new bonds when the linear viral DNA is converted into a 2-LTR circle is less apparent, since in this conversion there is a net increase in phosphodiester bonds (Figure 7A). Thus if the linear viral DNA can be integrated only via a 2-LTR circular intermediate, we have a paradox. We are examining four hypotheses that might resolve the paradox. First, the linear DNA rather than the 2-LTR circle could be the direct precursor to the integrated MLV provirus (Figure 7B). This may seem unlikely in view of the previously cited results of Panganiban and Temin (1984a), but it is worth noting that their results do not exclude the possibility that the linear form might be a direct precursor. Indeed, as these two forms differ only in the two bonds directly at the circle junction, they might not be distinguished by the integration machinery. Second, the linear DNA precursor might not be flush-ended as previously postulated, but might bear single-strand extensions that are lost by phosphodiester bond exchange in forming the 2-LTR circle. Third, the ends of the linear DNA, as isolated in our cell extracts, could already be "charged" by a high-energy bond, for example, by adenylation or by a protein–DNA bond. Fourth, the energy for ligation could be stored in a preformed ligase–adenylate intermediate, as has been demonstrated for other DNA ligase reactions (Weiss and Grindley, 1981; Craig and Nash, 1983; Gronostajski and Sadowski, 1985).

Applications

The availability of the in vitro system described here allows us now to begin to isolate and purify the components required for retroviral integration. We are currently investigating the structure of a subviral nucleoprotein complex that appears to be an active intermediate in MLV integration (B. Bowe et al., unpublished data). We have already been able to obtain a 100-fold purification (in activity per mg of protein) by gel-exclusion chromatography. As an alternative approach to identifying or confirming the identity of essential activities, antibodies or chemical inhibitors directed against defined viral or cellular proteins can be tested for their effects on integration in vitro.

Conversely, our assay can be used to screen for agents that specifically block the integration (or circularization) of retroviral DNA. Inhibitory factors of interest include the cellular components that mediate Fv-1 restriction of MLV host range (Jolicoeur, 1979) or interferon-induced blockade of provirus establishment (Morris and Burke, 1979), as well as synthetic compounds. Indeed, application of this assay to pathogenic retroviruses—for instance, the human immunodeficiency virus (HIV)—may be a useful approach to screening for a new class of antiviral agents.

We are currently working to develop a direct, physical assay for retroviral integration. Such an assay would be more rapid and economical than the one we have described here and, more importantly, would allow the direct detection of intermediates in the integration process. A physical assay appears quite feasible, at least for MLV, given the high efficiency of integration in vitro.

Experimental Procedures

Strains

E. coli KM392 (supF, hsdR2, hsdM2), a lac- derivative of LE392, was obtained from T. St. John. E. coli CC114 (sup2, lacZam, recA1, hsdR2, hsdM2) and MLVsupF (called MLVinSil1 in Lobe et al., 1985) were provided by Steve Goff.

Target DNA

A gtWES phage was purified by banding in a CsCl density gradient, and DNA was isolated as described in Maniatis et al. (1982). DNA was concentrated by ligation at 37°C overnight at a DNA concentration of 150-200 μg/ml, using T4 DNA ligase (IBI) in the reaction mix provided by the manufacturer. The reaction was stopped with EDTA and heated to 65°C for 15 min to inactivate the ligase. For experiments in which ATP-free DNA was required, E. coli DNA ligase (NEB) was substituted, and the ligation mix was as recommended by the manufacturer. Alternatively, ATP was removed (to <1 mM), after ligation with T4 DNA ligase, by dialysis, followed by phenol extraction and ethanol precipitation.

Cellular Extracts

Cells (106) of an MLVsupF-producing NIH 3T3 line were plated together with 4 x 106 uninfected NIH 3T3 cells in a 10 cm dish, in DMEM with 10% fetal calf serum and 6 μg/ml polybrene. Twenty-four hours after plating, the cells were harvested by trypsinization, washed once with buffer A (10 mM Tris-HCl, pH 7.4; 225 mM KCl; 5 mM MgCl2; 1 mM DTT; 20 μg/ml aprotinin), and then lysed in 250 μl per dish of buffer B (0.5 M sucrose, 50 mM Tris-HCl, pH 7.4; 10 mM MgCl2; 1 mM DTT; 20 μg/ml aprotinin). The lysate was centrifuged at 1000 g for 3 min. The nuclear pellet from this spin was resuspended in 250 μl per dish of buffer A, gently broken in a ball-bearing homogenizer, and then centrifuged at 8000 rpm in a JA-20 rotor for 20 min to pellet chromatin. The supernatant, designated “nuclear extract,” contained unintegrated linear and circular viral DNA, and was largely free of chromosomal DNA.

The supernatant from the initial low-speed spin of the cell lysate was centrifuged at 8000 rpm in a JA-20 rotor for 20 min. The supernatant from this spin, designated “cytoplasmic extract,” contained full-length linear viral DNA as well as subgenomic intermediates in DNA synthesis. Extracts retained activity for at least several weeks when stored at 4°C, frozen in liquid nitrogen, and stored at -70°C.

For some experiments, gel-exclusion column chromatography was used to remove ATP and other small molecules from the extract. Two milliliters of extract was loaded onto a 46 ml (20 cm) column of Bio gel A5m (Bio-Rad), equilibrated with buffer A, and chromatographed at a
Integration of Retroviral DNA In Vitro

flow rate of 0.5 ml/min. The excluded volume absorbance peak was pooled. ATP in the extract was assayed using firefly lantern extract (Sigma), as described in Cheer et al. (1974).

Integration Reactions
Integration reactions were carried out in a total volume of 150 μl in 1.5 ml microfuge tubes, and contained, except where noted (including the components of buffer A, provided by the cell extract), 100 μl cell extract (cytoplasmic or nuclear); 1.4 μg concentrated λ gtWES DNA; 20 mM HEPES, pH 7.5; 6 mM Tris-HCl, pH 7.4; 150 mM KCl; 5 mM MgCl₂; 1.25 mM ATP; 50 μM each of UTP, CTP, and GTP; 100 μM each of dATP, dGTP, dCTP, and dTTP; 10 mM creatine phosphate; 1 mM DTT; 13 μg/ml aprontin; and 5% polyethylene glycol 8000. The order of addition was the following: 15 μl 10× cocktail; 10 μl target DNA solution, 100 μl cell extract, mixed well; and 25 μl 30% polyethylene glycol. Reactions were incubated at 37°C for 1 hr, then stopped by addition of 12 μl of a solution of 250 mM EDTA, 0.3% proteinase K.

Scoring Integration
After incubation, reaction mixtures were subjected to protolysis at 37°C for 24 hr, then diluted with 160 μl of 0.4% SDS, incubated at 37°C for an additional 24 hr, and finally diluted with 160 μl of a solution containing 10 mM Tris-HCl (pH 7.4), 1 mM EDTA, and 600 mM NaCl. Ex extractions were performed gently with CHCl₃, then with phenol, then twice with CHCl₃:phenol:1:1, then with CHCl₃. The DNA was then recovered by ethanol precipitation, washed twice with 70% ethanol, resuspended overnight in 4 μl water, and packaged into phage particles, using Gigapack Gold packaging extract (Stratagene). To select for provirus-containing recombinants, the phage particles were plated on a lawn of CC114. Recombinant phage containing the MLVsupF provirus produce plaques on IPTG-I-X-Gal plates. Since MLV DNA has an unmodified EcoRI restriction site, the plating efficiency for the recombinants on the hsdR' strain LG75 is about three times higher than on the hsdR- strain LG114. DNA recovery and packaging efficiency for the recombinants on the hsdR- strain CC114 is about three times higher than on the hsdR' strain LG75. DNA recovery and packaging efficiencies were determined by examining the genetic map of λ gtWES to be 0.27. To control for differences in packaging and plating efficiencies between the provirus-containing recombinant phage genomes (50 kb) and uncombined λ gtWES (40 kb), defined amounts of DNA from λ gtWES and an MLVsupF-λ gtWES recombinant were ligated together, then packaged and plated in parallel with experimental samples. E is the ratio between the efficiencies with which unreacted λ gtWES and the recombinants were recovered as plaques.

DNA Analysis
Plaque amplification, λ DNA isolation, Southern blotting, and subcloning were performed using standard methods (Maniatis et al., 1982). Filter hybridizations followed the method of Church and Gilbert (1984). With repeated amplification of the provirus-containing recombinant phage, deletion derivatives are recovered in which the unique portion of the provirus is lost by recombination between the LTRs. Sequence analysis was by the dideoxy method (Sanger et al., 1977).

Acknowledgments
We thank Larry Donehower and Kimiko HagnYo-Yamagishi for their generous advice. Sue Klapholz for her helpful comments on the manuscript, and Steve Goff for providing plasmids and strains. This work was supported by funds from the National Institutes of Health, the G. W. Hooper Foundation, and the Lucille P. Markey Charitable Trust. P. O. B. is a Lucille P. Markey Scholar. H. E. V. is an American Cancer Society Research Professor.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 9, 1987; revised March 2, 1987.

References


