Retroviral recombination during reverse transcription

( Harvey murine sarcoma virus/Moloney murine retrovirus/polymase chain reaction/viral protein requirement/identical recombinants in vitro and in vivo )

DAVID W. GOODRICH* AND PETER H. DUESBERG

Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

Contributed by Peter H. Duesberg, December 21, 1989

ABSTRACT After mixed infection, up to half of related retroviruses are recombinants. During infection, retroviral RNA genomes are first converted to complementary DNA (cDNA) and then to double-stranded DNA. Thus recombination could occur during reverse transcription, by RNA template switching, or after reverse transcription, by breakage and reunion of DNA. It has not been possible to distinguish between these two potential mechanisms of recombination because both single-stranded cDNA and double-stranded proviral DNA exist in infected cells during the eclipse period. Therefore we have analyzed for recombinant molecules among cDNA products transcribed in vitro from RNA of disrupted virions. Since recombinants from allelic parents can only be distinguished from parental genomes by point mutations, we have examined the cDNAs from virions with distinct genetic structures for recombinant-specific size and sequence markers. The parents share a common internal allele that allows homology-directed recombination, but each contains specific flanking sequences. One parent is a synthetically altered Harvey murine sarcoma virus RNA that lacks a retroviral 3' terminus but carries a Moloney murine retrovirus-derived envelope gene (env) fragment 3' of its transforming ras gene. The other parent is intact Moloney virus. Using a Harvey-specific 5' primer and a Moloney-specific 3' primer, we have found recombinant cDNAs with the polymerase chain reaction, proving directly that retroviruses can recombine during reverse transcription unassisted by cellular enzymes, probably by template switching during cDNA synthesis. The recombinants that were obtained in vitro were identical with those obtained in parallel experiments in vivo.

Upon simultaneous infection, up to 50% of the allelic sequences of distinct retroviruses recombine, indicating efficient homologous recombination (1–3). Even illegitimate recombination among retroviruses has been reported to be “surprisingly efficient” (4). Illegitimate recombination during reverse transcription has been proposed to explain the characteristic ability of retroviruses to transduce cellular sequences (4–11, 28). This hypothesis was based on the efficient regeneration of Harvey murine sarcoma virus (HaSV) from a truncated provirus that lacked a 3' long terminal repeat (LTR) upon transfection in the presence of helper Moloney murine retrovirus (MoV) (4). However, reinvestigation of this experimental system indicated that recombination during reverse transcription is efficient only when sequence homology exists between the parental genomes (12, 13). Efficient HaSV regeneration from the truncated provirus was instead due to tandem recombinations of proviral DNA during transfection (12, 13).

Since retroviral RNA genomes in infected cells are converted first to complementary DNA (cDNA) by reverse transcriptase and then to double-stranded proviral DNA, recombination could occur by template switching during cDNA synthesis or by breakage and reunion of proviral DNA (14–16). As both cDNA and double-stranded proviral DNA exist simultaneously in the cell, it is not possible to distinguish between these two mechanisms in vivo (14, 15). One previous study (16) investigating recombination during reverse transcription in vitro by electron microscopy showed largely double-stranded DNA structures with single-stranded crossovers between molecules. These structures have been interpreted as the results of fragmented plus-strand DNAs involving gapped regions of double-stranded DNA. Therefore it was proposed that such structures serve as precursors for conventional DNA recombination in the cell (16).

Here we have investigated cDNAs transcribed in vitro from distinct parental RNAs for recombinant structures. For this purpose it was critical to develop a system that allowed biochemical detection of recombinants without the help of allelic markers that are biologically distinguishable in vivo (1–3). Since homologous recombinations between allelic parents would produce a series of recombinants that would differ only in the parental origins of point mutations and, hence, be extremely tedious if not impossible to sort out biochemically, we studied homologous recombination between parents that shared a homologous sequence flanked by specific heterologous sequences. Crossovers within the shared homology would thus generate recombinants with termini from each parent. Specifically, we searched for recombinants in which a truncated 3' terminus of Harvey sarcoma provirus was replaced by a 3' terminus from Moloney provirus by a crossover within an internal envelope gene (env)-specific homology. Detection of recombinant molecules was based on selective amplification with the polymerase chain reaction (PCR), using one primer that was specific for the HaSV parent and another that was specific for the MoV parent. By this method recombinant cDNAs were shown to be generated directly during reverse transcription. These results confirm the template-switching model (14) but not the DNA strand-invasion model (16) for homologous recombination upon reverse transcription.

MATERIALS AND METHODS

Transcription of Harvey and Moloney Virus RNAs to cDNAs in Vitro. Purified virus (17, 18) at 2–4 mg-protein equivalents per ml was incubated for 12–24 hr in a solution containing the four deoxynucleoside triphosphates (6 μM each), 10 mM MgCl2, 50 mM KCl, 20 mM dithiothreitol, 50 mM Tris/HC1 at pH 8, and 0.01% nonionic detergent (Nonidet P-40) (19–21). The cDNA was isolated from the reaction mixture by phenol extraction. In some experiments the viral RNA was eliminated from the cDNA by treatment for 4 hr at 40°C in 0.3 M NaOH (22).

Abbreviations: HaSV, Harvey murine sarcoma virus; MoV, Moloney murine retrovirus; LTR, long terminal repeat; PCR, polymerase chain reaction.

*Present address: Department of Pathology, School of Medicine, University of California at San Diego, La Jolla, CA 92039.

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.
Hybrid Selection of ras-Specific cDNA. The total cDNA obtained by reverse transcription of HaSV–MoV complexes was subjected to hybrid selection for ras-containing species. The ras DNA used for hybrid selection was a plasmid, termed p1.7, that contained a 1.7-kilobase (kb) DNA fragment from Harvey provirus bordered by BamHI and Xba I restriction enzyme sites (23). Hybridization between cDNA and nitrocellulose filter-bound plasmid DNA was carried out in 5 × SSC (1 × SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7) with 50% (vol/vol) formamide at 42°C for 12–24 hr, as described (12). The filters were washed three times at room temperature in 2 × SSC and once at 65°C in 0.1% SSC (12) before the bound cDNA was eluted by incubation in water at 95°C for 10 min. After washing, the filters typically retained about 0.5% of the total cDNA made in the reaction, as measured with radioactive tracer.

Specific Amplification of Recombinant cDNA by the PCR. The cDNA from about 0.1 A260 unit (=40 μg) of purified virus was incubated in 50 μl of 50 mM KCl/10 mM Tris, pH 8.3/2.5 mM MgCl2 containing 10 μmol of each deoxynucleoside triphosphate, 0.05 μmol of each of a set of two recombinant-specific primers (shown in Fig. 1), 10 μg of gelatin, and 1 unit of Thermus aquaticus (Taq) DNA polymerase. The mixture was heated to 94°C and then carried through 30–60 cycles of denaturation for 1 min at 94°C, annealing for 2 min at 54°C–62°C, and polymerization for 3 min at 72°C, as described (24, 25). An aliquot of the reaction mixture was then analyzed for size and sequence content by agarose gel electrophoresis.

RESULTS

The System Used to Study Recombination During Reverse Transcription in Vitro. To obtain recombinant cDNA molecules that could be physically and chemically distinguished from parental cDNAs, two viral RNA sequences were chosen that shared an internal region of homology but differed from each other in specific flanking regions. One parent was the RNA of an artificially altered HaSV and the other was the RNA of MoV. The two RNAs were obtained as a defective HaSV–helper MoV complex, termed R5ΨΨ I virus, that was harvested from cells transformed by a synthetically altered Harvey provirus and superinfected with MoV (12). The HaSV RNA of R5ΨΨ I virus carried an artificial 1.2-kb env sequence, derived from sequence positions 6025–7227 of MoV RNA (23), 3' of the ras coding region to allow homology-directed recombination with MoV (Fig. 1 and ref. 13). Crossovers between HaSV and MoV RNAs via this homology would define the sizes of the recombinant cDNA molecules (see Fig. 1 and below). Moreover, the HaSV RNA-packaging signal between the Xba I and Sac II restriction sites of the Harvey provirus (Fig. 1) was replaced via an Xba I linker by that of MoV (23), which is termed Ψ in Fig. 1. The purpose of this was to avoid a possible packaging bias of the MoV against HaSV RNA and thus a possible reduced yield in recombination during reverse transcription (13).

Other modifications of the Harvey provirus studied here had been introduced for parallel experiments studying recombination in vivo. (i) The untranscribed region of the 5' LTR of the Harvey provirus was truncated to prevent recovery of replicating HaSV by recombination among input HaSV DNAs during transfection (12, 13). (ii) The 3' terminus of the Harvey provirus was truncated. This prevented efficient synthesis of Harvey cDNA by reverse transcription, since it eliminated the terminally repeated RNA sequences necessary to transfer the cDNA, initiated from a primer near the 5' end of the retroviral RNA, to the 3' terminus (23). The 3' truncation thus forced selection for recovery of the 3' LTR.

Fig. 1. The proviral structures of a synthetically altered HaSV (pR5ΨΨ Ψ) and of MoV examined for recombination during reverse transcription. The boxes represent proviral DNA sequences, of which the LTRs are shaded. The two vertical lines within the LTR delineate the R region, from which transcription of viral RNA originates (23). The thick horizontal lines depict sequences from the plasmid pBR322. EcoRV, Xba I, HindIII, Nco I, and BamHI restriction endonuclease sites in the proviruses are indicated. The 5' LTR of the Harvey provirus is truncated up to an EcoRV site, designated by a jagged line. Ψ represents a 714-base MoV-derived RNA-packaging region bordered on the 5' side by the Xba I site of the LTR and on the 3' side by a Pst I site joined with the Sac II site of HaSV with an Xba I linker. ras designates the coding region of the Harvey sarcoma virus transforming gene, and env is a 1.2-kb Nco I-bordered region derived from the envelope gene of MoV (23). The diagram shows the 3' half of MoV, including part of the reverse transcriptase gene (pol) and all of the env gene, aligned with HaSV based on their common 1.2-kb env region. The pBR322-derived, pR5ΨΨ I-specific nucleotide sequence 5' of the common env region and the MoV-specific sequence 3' of it were used as primers in PCRs carried out to detect HaSV–MoV recombinant cDNAs.
by recombination with MoV in the biological assay system (4).

**Identification of Recombinant cDNAs Made in Vitro.** Re-
combination between the altered HaSV and MoV RNAs of the R5%IH 1° virus complex during reverse transcription was examined by analysis of recombinant Harvey ras-containing cDNAs made in vitro after selective amplification by the PCR. Before amplification by the PCR, Harvey ras-
containing cDNA was separated from MoV cDNA by hybrid selection using ras DNA bound to a nitrocellulose filter. This selection precluded recombination between HaSV and MoV cDNAs during the subsequent PCR, since only MoV-specific sequences that were already covalently linked to ras sequences would survive the selection. Further, the hybrid selection eliminated PCR “noise” that could arise from nonspecific priming by MoV cDNAs and possibly nonviral cDNAs. The yield of HaSV cDNA from our HaSV–MoV complex (R5%IH 1° virus) was only 0.5% of the total cDNA. This low yield was consistent with the truncated structure of the Harvey RNA in this virus, which prevents efficient cDNA synthesis prior to recombination (see below).

The ras-specific cDNA of the R5%IH 1° virus was then specifically amplified for HaSV–MoV recombinant mole-
ecules by PCR using the Harvey-specific 5′ primer and the Moloney-specific 3′ primer, which flank the 1.2-kb env homology region. Based on the positions of the primers used, a recombinant DNA of 1.2 kb was expected (Fig. 1). A DNA of the expected size was indeed detected after agarose gel electrophoresis of the PCR products whether or not viral RNA was removed from cDNA by NaOH (Materials and Methods) prior to amplification (Fig. 2A, lane 2). In addition, this reaction and other reactions, described below, yielded DNAs of about 1 kb and 0.5 kb (Fig. 2A), which appear to be artifacts of the PCR.

The following controls indicated that the 1.2-kb DNA was generated by recombination between the RNA genomes of HaSV and MoV during reverse transcription. The 1.2-kb DNA band was not amplified when either of the primers was left out of the reaction (data not shown). Further, no 1.2-kb DNA was generated when the two primers were used to amplify cDNA from in vitro transcription of MoV RNA in a detergent-disrupted virus, although the nonspecific 0.5-kb product was generated (Fig. 2A, lane 1). In addition, recom-
binant RNA was shown not to have preexisted in R5%IH 1° virus by two methods. (i) Genomic DNA of R5%IH 1° virus-producing cells was isolated and examined by PCR for the 1.2-kb recombinant DNA; no such DNA was found (Fig. 3, lanes 1–3). (ii) When purified RNA of R5%IH 1° virus was transcribed in vitro with commercial reverse transcriptase using random primers (Amersham), and then amplified by the PCR with the 5′ HaSV-specific 5′ primer and the MoV-
specific 3′ primer, no 1.2-kb recombinant DNA was found (Fig. 2A, lane 4). Further, no 1.2-kb DNA was detected if this cDNA was mixed with MoV cDNA prepared in detergent-
disrupted virus prior to amplification by the PCR (Fig. 2A, lane 5). Likewise, no recombinant RNA or proviral DNA was found by direct analysis of cellular RNAs and DNAs from R5%IH 1° virus-producing cells in a parallel study (13). It followed that the 1.2-kb HaSV–MoV recombinant DNA generated during reverse transcription of detergent-disrupted R5%IH 1° virus originated from homology-directed recombi-
nation involving the 1.2-kb env homology shared by the proviral RNAs (Fig. 1).

The recombinant nature of the 1.2-kb DNA fragment from R5%IH 1° virus cDNA was confirmed by two methods. (i) Digestion of the env sequence within the 1.2-kb DNA frag-
ment with restriction endonuclease BamHI should produce two fragments 710 and 530 bases long, based on the locations of BamHI and Nco I sites in the env gene of MoV (ref. 23 and Fig. 1). Fragments of these sizes were observed, along with the disappearance of the 1.2-kb fragment, upon BamHI digestion of the amplified cDNA products of R5%IH 1° virus (Fig. 2A, lane 3). (ii) The 1.2-kb DNA amplified from R5%IH 1° virus cDNA hybridized with two 32P-labeled oligonucleotides from sequence positions 6480–6500 and 7247–7270 in MoV (23) (Fig. 2B, lane 2). The 710- and 530-base BamHI restriction fragments also hybridized with oligonucleotides 6480 and 7247, respectively (Fig. 2B, lane 3). In addition, the above-mentioned 1- and 0.5-kb fragments, which were not digested by BamHI, hybridized with these oligonucleotides, suggesting that they are env-related sequences (Fig. 2B, lanes 2 and 3). As expected, no specific signals were observed upon hybridization of DNA amplified from MoV cDNAs with the recombinant DNA-specific primers (Fig. 2B, lane 1).

**Conditions for Recombination During Reverse Transcrip-
tion in Vitro.** In an effort to define the requirements of recombination during reverse transcription in detergent-disrupted virus, we set out to reconstruct a recombination system from purified components in vitro. For this purpose, phenol-extracted RNA from R5%IH 1° virus (which contains both parental RNAs) was incubated with purified reverse transcriptase but otherwise under the conditions described...
Fig. 3. Cellular and proviral DNA from cells transformed by truncated Harvey provirus (pR5ΨH) and then infected with MoV, and from cells transformed by virus harvested from Harvey provirus-transformed cells. Lanes 1–3: genomic DNA (4, 2, and 1 μg, respectively) from NIH mouse 3T3 cells transformed by pR5ΨH Harvey provirus (Fig. 1) and superinfected with MoV was subjected to PCR amplification using the HaSV-specific 5' and MoV-specific 3' primers described for Fig. 1; the total reaction product was electro- phoresed as described for Fig. 2 and stained with ethidium bromide. Lane 4: genomic DNA (1 μg) of cells transformed by virus harvested from the Harvey provirus-transformed cells analyzed in lanes 1–3 was subjected to amplification and analyzed as for lanes 1–3. DNA sizes were estimated based on HindIII-resistant λ phage DNA size markers.

above for detergent-disrupted virus. The cDNA was subse- quently amplified with the HaSV-specific 5' and MoV- specific 3' primers and examined for the characteristic 1.2-kb recombinant DNA. No 1.2-kb DNA was found (Fig. 2A, lane 4). Likewise, no 1.2-kb recombinant DNA was produced when the same cDNA template was amplified in the presence of MoV cDNA generated by reverse transcription in detergent-disrupted virus (Fig. 2A, lane 5). It appears that unde- fined viral structural proteins, together with the proximity of two parental genomes within a viral particle (26, 27), are essential for crossing-over during reverse transcription in vitro. Such proteins have also been postulated to be neces- sary for in vitro transcription of full-length cDNA from viral RNA (19). The absence of a 1.2-kb fragment from cDNA transcribed from purified R5ΨH 1° viral RNA with purified reverse transcriptase, even in the presence of MoV cDNA generated by disrupted virus, also indicated that recombinant cDNA obtained by reverse transcription of detergent-disrupted R5ΨH 1° virus was not an artifact of the PCR technique.

In Vivo Recombination Generates the Same Products as Recombination During Reverse Transcription in Vitro. If the same recombination that we observed in vitro were to occur during reverse transcription in vivo, infection with R5ΨH 1° virus (which contains the two parental RNAs) would be expected to generate recombinant cDNA. Further, R5ΨH 1° virus-infected cells would be expected to produce recombi- 
nant virus.

Indeed, in vitro transcription with purified reverse transcrip- 
tase of phenol-extracted RNA from virus produced by R5ΨH 1° virus-infected cells generated the expected 1.2-kb recombinant DNA fragment after amplification (Fig. 2B, lane 8). As expected, this 1.2-kb DNA contained the 710- and 530-base BamHI-resistant fragments (Fig. 2B, lane 9). (Not enough DNA was applied to visualize the DNAs by ethidium bromide staining in Fig. 2A, lanes 8 and 9.) By contrast, no recombinant DNA was obtained by reverse transcription of the purified RNAs from parental R5ΨH 1° virus (Fig. 2, lanes 4). Moreover, the genomic DNA of R5ΨH 1° virus-infected cells contained 1.2-kb recombinant proviral DNA, as evidenced by PCR amplification (Fig. 3, lane 4).

When cDNA made by detergent-disrupted virus from cells infected by R5ΨH 1° virus was amplified with the HaSV- specific 5' and MoV-specific 3' primers and analyzed by electrophoresis, the 1.2-kb recombinant DNA was detected by staining with ethidium bromide (Fig. 2A, lane 6) and by hybridization with the cDNA-specific oligonucleotides (Fig. 2B, lane 6). This DNA also contained the expected 710- and 530-base BamHI-resistant fragments (Fig. 2A and B, lane 7). However, the background of nonspecific cDNAs was higher in this cDNA than in the cDNA prepared from R5ΨH 1° virus, presumably because this cDNA was not hybrid-selected for ras-specific species prior to amplification.

It appears that the same recombination that was observed in vitro during reverse transcription also occurred in vivo—possibly by the same mechanism.

DISCUSSION

Recombinant molecules that did not exist prior to reverse transcription were generated by homologous recombination during reverse transcription within disrupted virions in vitro. Due to the distinct genetic structures of the parental and recombinant genomes in our system, such recombinant mol- ecules could be amplified selectively by the PCR among primary cDNA transcripts. Similar recombinant molecules were also observed in vivo in cells infected by the same virus used for reverse transcription in vitro. Hence, retroviral RNA genomes can presumably recombine during reverse trans- 

cription in vivo.

Several observations indicate that most, if not all, recombinations observed here in vitro occurred by RNA template switching during cDNA synthesis rather than during or after synthesis of double-stranded DNA. Since the HaSV RNA studied here lacks a retroviral 3' terminus, cDNA can be initiated by the retroviral primer near the 5' end but cannot be extended from the 5' repeat (R)-region (Fig. 1) to the missing 3' repeat. Thus the 3' truncated HaSV RNA will not be converted to cDNAA, except by chance via random primers or by template switching of cDNA initiated from MoV RNA. It is consistent with this view that only about 0.5% of the total cDNA generated in vitro by our R5ΨH 1° virus, which contains both MoV RNA and 3' truncated HaSV RNA, was HaSV cDNA—most of which presumably represents HaSV- MoV cDNA recombinants. This is not due to a shortage of HaSV RNA templates, because we have shown previously that HaSV (MoV) complexes with 3' truncated HaSV RNAs like those described here carry just as much HaSV RNA as wild-type HaSV (12). The observation that 3'-truncated HaSV RNA in the absence of homology 3' of ras is not infectious and not subject to recombination with helper MoV (12) is also entirely consistent with the proposal that it is not an adequate template for reverse transcription.

Our results are at variance with a previous study investigat- ing recombination during reverse transcription (16). That study concluded that recombination is initiated only during synthesis of the second DNA strand. The second strand was reported to invade double-stranded DNA molecules at hom- 

ologous gapped regions and the resulting structures were postulated to initiate conventional DNA recombination in the cell (16). Recombination during reverse transcription may have been overlooked because RNA-DNA hybrids were not distinguishable from DNA-DNA hybrids by the electron microscopy and because physically indistinguishable paren- 

tal genomes were analyzed (16). Other studies of retrovirus recombination did not distinguish between recombination mechanisms that involved RNA-cDNA complexes and dou-
ble-stranded proviral DNAs, because the assay depended on biological selections of recombinant viruses that appeared only after one or several rounds of virus replication in vivo, rather than on direct selection of recombinant molecules (1–3, 14).

We have not attempted to analyze illegitimate recombination in our system for two reasons. (i) The sizes of recombinants that might arise by illegitimate recombination in vitro would be unpredictable and heterogeneous; hence, it would be difficult, if not impossible, to distinguish such recombinants from the background noise generated by the PCR. Potential recombinants would thus have to be cloned and analyzed individually. (ii) Since there is no evidence for efficient or reproducible illegitimate recombination in vitro between nonhomologous sequences of retroviruses (12, 13), any illegitimate recombination detected in vitro might have no biological relevance.

We thank G. S. Martin and Ren-Ping Zhou for critical comments. This work was supported by Outstanding Investigator Grant 5-R35-CA39915-04 from the National Cancer Institute.
