

Kinetics of expression of multiply spliced RNA in early human immunodeficiency virus type 1 infection of lymphocytes and monocytes

(alternative splicing/reverse transcriptase-PCR/regulatory genes/human retrovirus)

MARY E. KLOTMAN*[†], SUNYOUNG KIM*[‡], ABY BUCHBINDER*[§], ANITA DEROSI*[¶], DAVID BALTIMORE^{||}, AND FLOSSIE WONG-STAAAL*^{††}

*Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20815; [‡]New England Deaconess Hospital and Harvard Medical School, Boston, MA 02215; and ^{||}The Rockefeller University, New York, NY 10021

Contributed by David Baltimore, March 14, 1991

ABSTRACT The genome of human immunodeficiency virus type 1 (HIV-1) encodes at least six proteins involved in regulation as well as the structural proteins Gag, Pol, and Env. The interplay of the various regulators generates early and late transcriptional phases in the HIV-1 life cycle; the earliest RNA is enriched in subgenomic species, and the genomic transcript appears at the later stage of infection. We investigated the nature of the mRNAs expressed in the early stages of infection when the 2 kilobase subgenomic species predominate. RNA was analyzed in the early phase of a one-step growth cycle of HIV-1 infection in T-lymphoid and monocytic cell lines by using PCR amplification of *in vitro*-synthesized viral cDNAs. In both cell lines, expression of Tat-, Rev-, and Nef-specific messages appeared simultaneously and could be detected within 8–12 hr of infection but in different amounts with a predominance of Nef-specific message. The Env-specific message could be detected as early as the Rev-specific message, indicating that expression of at least small amounts of the singly spliced message could occur before the accumulation of Rev.

The human immunodeficiency type 1 (HIV-1) genome contains at least six genes that may be involved in regulation in addition to the structural genes common to all retroviruses, *env*, *pol*, and *gag*. There is an emerging picture of a complex pattern of regulation of the virus mediated by at least two of the transregulatory proteins, Tat and Rev. While Tat increases expression from the HIV-1 long terminal repeat, resulting in enhanced expression of all viral genes (1), Rev differentially enhances the expression of the unspliced or singly spliced messages coding primarily for structural genes and genomic RNA (2–6). It has been shown that there is a controlled temporal pattern of expression of HIV RNAs upon acute infection; the earliest RNA is enriched in subgenomic species, and the genomic transcripts appear at the later stage of infection (7). This same pattern has been noted when chronically infected low-expression cell lines are stimulated with agents that activate the HIV-1 long terminal repeat (8). It is not yet known which specific genes are being expressed at the early stages of HIV-1 infection and activation. Such a pattern of transcriptional phases has been well described in other viruses, such as simian virus 40, polyoma virus, and herpesviruses, where early and late phases of transcription can be differentiated (9). In these instances the genes are regulated by unique promoters, whereas in HIV-1 a single promoter regulates expression of all genes.

The mRNAs of all HIV-1 proteins are processed from the same precursor RNA with alternative splicing allowing for protein diversity. The mRNAs coding for the regulatory

proteins Tat, Rev, and Nef are doubly spliced with identical 5' and 3' exons, differing in the splice acceptor of the second exon. The determination of splice-site use provides an important additional means of posttranscriptional regulation (10), particularly in light of the different regulatory functions associated with the proteins encoded by these transcripts. To better understand the control of early events that occur upon acute infection we examined the kinetics of splicing involved in the expression of viral genes, particularly the regulatory genes. We studied the early events during a one-step growth cycle in the T-lymphoid and monocytic cell lines. PCR amplification of *in vitro*-synthesized cDNAs with subsequent oligonucleotide hybridization was used for the specific and sensitive detection of the spliced HIV messages. Unlike conventional RNA (Northern) blot analysis, this method allowed us to distinguish the specific splicing events involved in the expression of Tat, Rev, and Nef mRNA as well as of Vif and Env. Increased synthesis of specific mRNAs for regulatory proteins could be detected 8–12 hr after infection in both lymphocyte and monocyte cell lines. The patterns of splicing and relative abundance of the regulatory gene messages were identical in both of these cell lines. Although the regulatory mRNAs appeared simultaneously, the doubly spliced Nef mRNA was always more abundant relative to the doubly spliced mRNA coding for Rev and Tat. The singly spliced RNA capable of coding for Env could be detected as early as the regulatory messages by using this sensitive means of detection.

MATERIALS AND METHODS

Virus and Infection. Host cells for HIV-1 infection were H9 and U937, T-lymphoid and monocytic cell lines, respectively. The virus strain used was HIV-1_{W13}, which grows preferentially in T cells. A high-titer viral stock was made by the shaking method (7). The same number of H9 and U937 cells was infected with equal amounts of virus as described (7). At appropriate time points, samples were taken for RNA isolation. RNA from uninfected H9 cells as well as from H9 cells chronically infected with HIV-1_{IIIIB} was analyzed.

RNA Preparation. Preparation of the total cellular RNA was as described by Kim *et al.* (7).

cDNA Synthesis and PCR Amplification. Synthetic oligonucleotide primers and probes were used for cDNA synthe-

Abbreviation: HIV-1, human immunodeficiency virus type 1.

[†]To whom reprint requests should be addressed.

[§]Present address: Departments of Medicine and Pathology, Manhattan Veterans Administration Hospital, New York, NY.

[¶]Present address: Institute of Oncology, University of Padova, Padova, Italy.

^{||}Present address: Departments of Medicine and Biology, University of California at San Diego, San Diego, CA 92093.

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.

sis, PCR amplification, and hybridization detection of specific HIV cDNAs. The cDNAs were synthesized in a standard reverse transcriptase reaction. The reaction buffer containing 50 mM Tris-HCl (pH 8.0), 3.0 mM MgCl₂, and 75 mM KCl mixed with 200 ng of primer and 2 μg of total cellular RNA was heated to 65°C and slowly cooled to 24°C to allow primer annealing. Sixty units of avian myeloblastosis virus reverse transcriptase (Super Reverse Transcriptase; Molecular Genetics Resources, Tampa, FL) was added along with a mixture of dNTPs and dithiothreitol to a final concentration of 1 mM and 8 mM, respectively. The reaction was done at 40°C for 1 hr. Five microliters of the cDNA reaction was used directly in the amplification reaction.

Each antisense primer was paired with the same-sense primer, designated US, which is located upstream from the first major splice donor at the 5' end of the genome, for amplification of specific cDNAs. The PCR included the sense and antisense primers at a final concentration of 0.5 μM, reaction buffer of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, and 0.01% gelatin, 200 μM of each dNTP, 2.5 units of *Taq* polymerase (Perkin-Elmer/Cetus), and 5 μl from the reverse transcriptase reaction as the cDNA template. The samples were amplified through 30 cycles that included the following parameters: denaturation at 92°C for 1 min, primer annealing at 55°C for 2 min, and primer extension at 72°C for 2 min. The final cycle included a 7-min extension at 72°C.

Oligonucleotide Primers and Probes. Oligonucleotide primers and probes were made on an Applied Biosystems synthesizer. The location of the primers and probes used for cDNA synthesis and detection, as well as their sequences, is indicated in Figs. 1 and 2A, respectively. The primer designated ART7 is derived from the third exon of the regulatory genes *tat*, *rev*, and *nef*, the primer ART5 is downstream from the splice acceptor for *vif*, and the primer ART2 is downstream from the initiation codon of *env*. In each case, the primer pairs used for amplification spanned known splice junctions for the regulatory and *env* genes illustrated below the primers in Fig. 1 and were anticipated to yield specific products, as predicted by the known HIV-1 splice sites (Fig. 2B). The combination of primers ART7/US was used to amplify the doubly spliced mRNAs for the regulatory proteins Tat, Rev, and Nef; the primer pair ART5/US used to amplify the first splice junction of *Vif* mRNA and ART2/US

was used to amplify spliced *Env* mRNA. The viral probe designated S1 was used to detect *tat*-specific cDNA and is located in the second exon upstream from the *rev* and *nef* splice acceptors. The probe S2 was used to detect both *tat*- and *rev*-specific cDNAs and is located in the second exon of *rev* upstream from the *nef* splice acceptor. The probe S3 located at the 3' end of the second exon common to *Tat*, *Rev*, and *Nef* messages was used to detect these cDNAs as well as *env* cDNA. A specific probe (S4) spanning 14 bases on either side of the first splice junction of *vif* was used to detect this message. Additional oligonucleotide probes derived from the two small exons upstream from *tat* (12), ART9 and ART8, and from the third exon of the mRNAs coding for the regulatory proteins, S5, were used to identify additional amplified cDNAs.

Gel Electrophoresis and Hybridization. The PCR products (25 μl of the reaction) were resolved on 5% nondenaturing polyacrylamide gels, denatured with 0.5 M NaOH/1.5 M NaCl, neutralized in 0.5 M Tris-HCl (pH 8.0)/1.5 M NaCl, and transferred electrophoretically onto Nytran membranes (Schleicher & Schuell) using 20 V over 18 hr in 1× TBE buffer (0.89 M Tris/0.89 M boric acid/0.02 M EDTA, pH 8.3). The nucleic acid was bound to the membrane by UV irradiation using the UV Stratlinker 1800 (Stratagene). The membranes were prehybridized in 6× NET (20× NET is 3 M NaCl/20 mM EDTA/0.3 M Tris-HCl, pH 8.0)/0.1% SDS/5× Denhardt's solution/tRNA at 150 μg/ml. Oligonucleotide probe end-labeled with [³²P]ATP by standard methods using T4 polynucleotide kinase was added to the prehybridization solution at 1 × 10⁶ cpm/ml of solution. After overnight hybridization, the membrane was washed in 6× NET/0.1% SDS at room temperature twice for 15 min and in the same solution at 55°C. After autoradiography, each probe was stripped from the filter by treatment with 0.5 M NaOH/1.5 M NaCl before hybridization with another probe.

RESULTS

To investigate the structure of the RNA species expressed during chronic infection, cDNAs were synthesized and amplified from RNA prepared from chronically infected H9 cells. When the PCR products generated with primer pair ART7/US were hybridized with the S1 probe, there was a

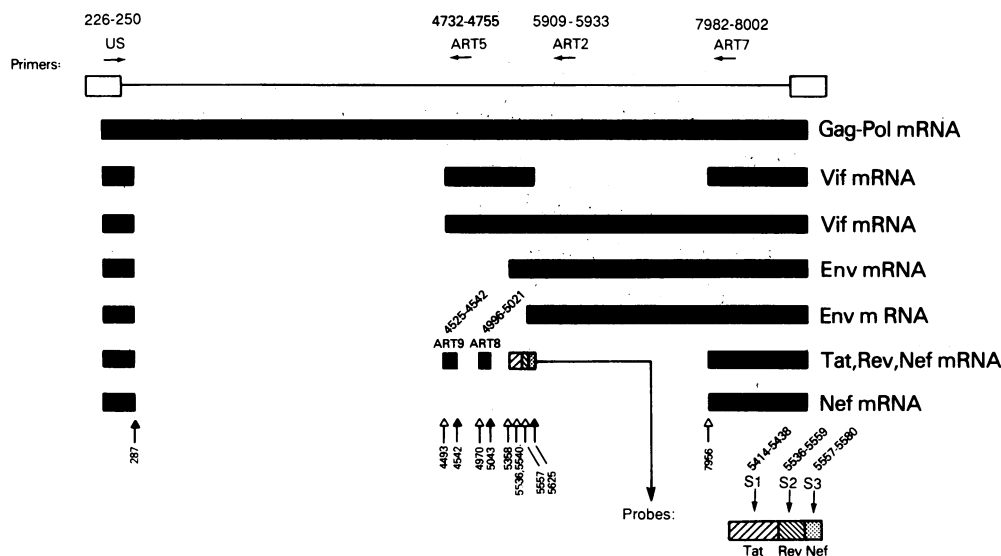


FIG. 1. Structures of unspliced and spliced mRNAs coding for HIV-1 proteins. Primers used for cDNA synthesis and PCR amplification are indicated by arrows. Direction of arrow indicates primer orientation. Those primers designated ART are antisense primers, and those designated US are sense primers. Location of oligonucleotides used as probes is indicated above the appropriate exon. The thin line represents genomic DNA with the flanking long terminal repeat, and thick lines represent exons. Closed arrowheads at the bottom indicate splice donors; open arrowheads indicate splice acceptors. Nucleotide numbering is according to Ratner *et al.* (11).



FIG. 2. (A) Nucleotide sequences of primers and probes. Several primers have the restriction sites, as indicated by the box, with additional 5' sequence to allow for cloning of PCR products. The first 5' nucleotide of the specific HIV-1 sequence is indicated. (B) Specific primer pairs and probes used to detect the indicated PCR products.

predominant band of 402 bases representing the doubly spliced message capable of coding for Tat (Fig. 3A, lane 14). The probe S2 detected two prominent bands of 402 and 219/225 representing *tat* and *rev* cDNAs, respectively (Fig. 3B, lane 14). Cloning and sequencing of *rev*-specific amplified cDNAs revealed a mixture of two previously reported splice acceptors for Rev-specific message (13–16). When probe S3 was used on the same gel, an additional prominent band was detected below the 219/225-base-pair (bp) band, and this was as predicted, 203 bp representing the spliced cDNA capable of coding for Nef (Fig. 3C, lane 14). Several of the less prominent bands in lanes 14 in Fig. 3 also hybridized to probe ART8 or ART9 and, therefore, represent Tat-, Rev-, and Nef-specific messages with additional upstream exons.

In lymphocytes, the specific Tat message could be detected as early as 4 hr after infection with significant increases seen between 12 and 16 hr in both cell lines (Fig. 3A). The Rev-specific 219/225-bp message could also be faintly de-

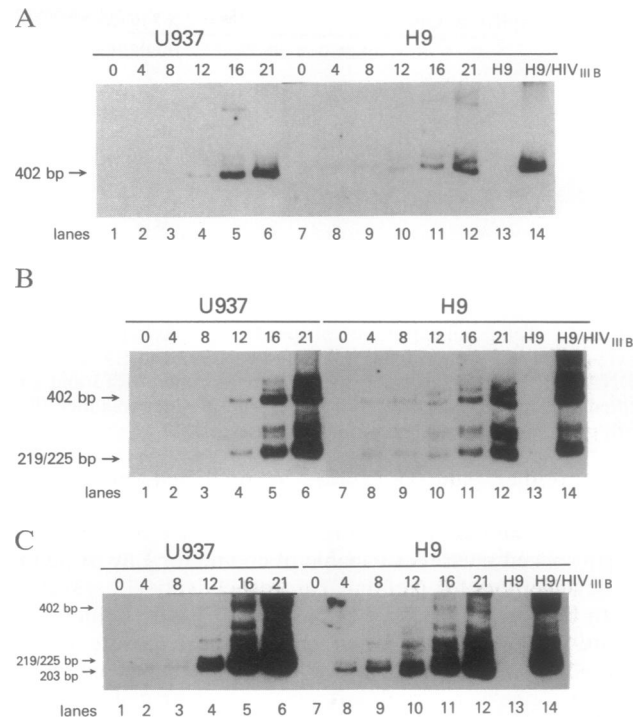


FIG. 3. Temporal expression of specific mRNAs for regulatory proteins from HIV-1-infected U937 and H9 cell lines. cDNA was synthesized by using the antisense primer ART7 and amplified with the primer pair ART7/US. Each autoradiogram is a 24-hr exposure. The numbers above each lane indicate time of RNA extraction after viral infection. H9/HIV_{III B} represents amplified cDNA synthesized from chronically infected H9 cells, whereas time 0 and H9 represent amplified cDNA synthesized from uninfected cells. (A) Temporal expression of *tat*-specific mRNA detected with probe S1. (B) Temporal expression of Tat- and Rev-specific mRNA detected with probe S2. (C) Temporal expression of Tat-, Rev-, and Nef-specific mRNA detected with probe S3.

tected at 4 hr in the lymphocytes but did not increase significantly until 12–16 hr in both cell lines (Fig. 3B). The 203-bp band seen with probe S3 in Fig. 3C representing a *nef*-coding cDNA was detected as early as the other regulatory messages in the lymphocyte cell line and significantly increased by 12 hr postinfection. (It is possible that the faint signals seen at 4 hr in the RNA extracted from lymphocytes could represent spliced mRNA in the viral preparation used to inoculate the cells; however, it was not seen in the 4-hr monocyte preparation.) Because the same primer pair and PCR reaction were used to identify all three regulatory messages, the messages could be quantitatively compared when probe S3 was used to detect all three cDNAs. In both cell lines, the *nef*-specific cDNA was more abundant at all time points than the doubly spliced messages coding for Rev and Tat. In H9 and U937 cell lines, the pattern of expression and relative abundance of the cDNAs coding for the regulatory proteins were identical, although detectable expression of the regulatory messages appeared to be delayed by 4 hr in the monocytes.

Expression of message using the splice site for Vif mRNA was not readily detected before 16 hr after infection in both cell lines (Fig. 4); however, because a different primer pair and probe were used for amplification and detection, a quantitative comparison relative to the other regulatory mRNAs could not be made. This splice site is used for the expression of both singly and doubly spliced mRNA coding for Vif as well as by a small noncoding exon that has been found in multiply spliced variants of Tat-, Rev-, and Nef-specific mRNA (Fig. 1) (15).

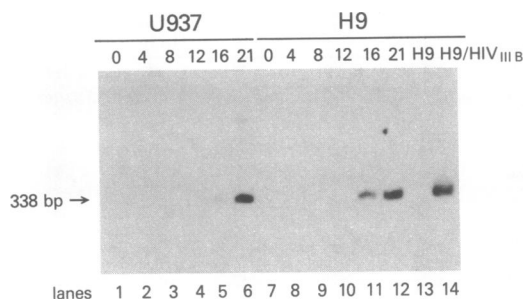


FIG. 4. Temporal expression of specific mRNAs using the splice acceptor for Vif protein from HIV-infected U937 and H9 cell lines. cDNA was synthesized using the antisense primer ART5 and amplified with the primer pair ART5/US. The product was identified with probe S4. Lanes are as indicated in Fig. 3.

The primer pair ART2/US, as predicted, amplified a predominant product of 466 bases and a less prominent larger band that was 665 bases (Fig. 5). These are the predicted singly spliced messages capable of coding for Env using the splice acceptors for Nef and Tat, respectively (12, 13) (Fig. 1). In both cell lines, env cDNA could readily be detected before 12 hr after infection on a 24-hr exposure of the autoradiogram (Fig. 5); however fainter bands could be detected as early as regulatory cDNAs in both cell lines with longer exposure of the autoradiogram.

DISCUSSION

The method of cDNA synthesis and PCR amplification was highly sensitive and specific for the detection of the doubly spliced messages capable of coding for Tat, Rev, and Nef proteins. This method allowed for the distinction between these messages that are very similar in size and have extensive overlap of their reading frames and that cannot be distinguished by standard methods. This method should allow for the study of expression of these messages *in vivo*, where the relative abundance of the messages does not generally allow detection or distinction by other methods.

It has been demonstrated (7, 8) that a temporal pattern of RNA accumulation exists in early *in vitro* HIV-1 infection as well as in early reactivation of virus from chronically infected lines. We further investigated and identified the specific subgenomic messages expressed during the early phase of the productive infection when the 2-kilobase (kb) subgenomic mRNA species predominates. Our results indicate that the 2-kb spliced messages capable of coding for the specific regulatory proteins Tat, Rev, and Nef can be detected early after infection in both lymphocytes and monocytes and are expressed simultaneously rather than sequentially, although in different amounts relative to each other. The overall pattern of expression of the doubly spliced Tat, Rev, Nef, and

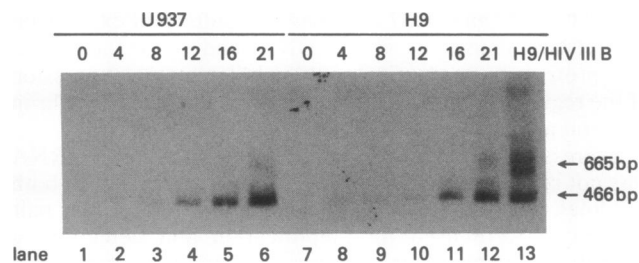


FIG. 5. Temporal expression of specific mRNAs for Env protein from HIV-infected U937 and H9 cell lines. cDNA was synthesized using the antisense primer ART2 and amplified with the primer pair ART2/US. The product was identified using probe S3. Lanes are as indicated in Fig. 3.

spliced Vif messages as well as their relative abundance were similar in both cell lines.

The simultaneous expression of doubly spliced regulatory messages is consistent with the presence of a single promoter; however, there are potentially important differences in the quantity of the specific messages. The Nef-specific mRNA was overproduced relative to Tat (and to a lesser extent Rev) at all time points examined in the acute infection as well as in the chronically infected cell line. The relative abundance of Nef mRNA agrees with recent work by others (14, 15). Robert-Guroff *et al.* (14), using similar methods, demonstrated a predominance of Nef-specific small mRNAs in H9 cells chronically infected with HIV-1_{IIIIB} as well as in peripheral blood monocyte/macrophages acutely infected with the HIV-1_{Ba-L} strain. However, they observed a predominance of a two-exon message for Nef in the peripheral blood monocyte/macrophages, which we did not observe in the acutely infected monocyte line U937 (data not shown). Guatelli *et al.* (16) also demonstrated the predominant use of the second-exon splice acceptor for the Nef-coding mRNA; however, because of the location of the primers used, they could not distinguish singly from doubly spliced messages because both Env-specific and Nef-specific messages use this acceptor (16).

Singly spliced Env-specific message could be detected by PCR in both cell lines simultaneously with Rev, consistent with our previous Northern hybridization analysis (7, 17). These results suggest that at least small quantities of singly spliced messages do not require prior accumulation of the Rev protein. This is not surprising in light of the extreme sensitivity of this detection method. Further quantitative comparisons cannot be made because these messages were detected with different primer pairs that can vary significantly in efficiency of amplification. The accumulation of Tat- and Rev-specific mRNA may be required before the appearance of unspliced genomic RNA, but it is difficult to distinguish amplified proviral DNA from genomic RNA by this method.

The differential expression of the doubly spliced messages represents another important posttranscriptional means of regulation of virus replication. It is known that the viral protein Rev acts posttranscriptionally *in trans* to enhance the expression of the singly spliced and unspliced structural mRNAs and genomic RNA (2-6). The mechanism(s) of control resulting in the balanced expression of the individual multiply spliced mRNAs coding for the regulatory genes is unknown. Our similar results in two different cell lines using the same virus do not exclude the possibility that cell factors acting *in trans* might influence this relative expression. Work done with other retroviruses supports the importance of the relative efficiencies of splice acceptors in the controlled expression of viral genes. The inefficient splice acceptor of the Env mRNA of avian sarcoma virus results in a balanced expression of spliced and unspliced mRNA. This inefficiency, at least in part, is mediated through the specific cis-acting sequences recognized by the cellular splicing machinery (18). The contribution of cis-acting sequences and trans-acting cellular factors to our observed differences in the efficiency of splice-acceptor use and mRNA expression is not known. The significance of the relative and persistent abundance of the Nef message remains unclear due to the uncertain role its gene product plays in the virus life cycle (19-25). In light of the unique activities of the proteins coded for by these multiply spliced mRNAs, understanding the control mechanism of such differential RNA expression will be vital to unraveling the pathogenetic mechanism of HIV-1.

We acknowledge the continued support of Dr. Robert C. Gallo. M.E.K. is supported through the career development program of the Veteran's Administration Hospital. S.K. was a fellow of the Jane

Coffin Childs Memorial Fund for Medical Research. This work is supported in part by Public Health Service Grants AI30897 (S.K.) and HL43510 (D.B.) and by the Clark Foundation for AIDS Research (S.K.).

1. Arya, S. K., Guo, C., Josephs, S. F. & Wong-Staal, F. (1985) *Science* **229**, 69–73.
2. Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C. & Wong-Staal, F. (1986) *Cell* **46**, 807–817.
3. Sadaie, M. R., Benter, T. & Wong-Staal, F. (1988) *Science* **239**, 910–913.
4. Sodroski, J., Goh, W. C., Rosen, C., Tartar, A., Portetelle, D., Burny, A. & Haseltine, W. (1986) *Science* **231**, 1549–1552.
5. Portetelle, D., Burney, A. & Haseltine, W. (1986) *Science* **231**, 1549–1552.
6. Malim, M. H., Hauber, J., Fenrick, R. C. & Cullen, B. R. (1988) *Nature (London)* **335**, 181–183.
7. Kim, S., Byrn, R., Groopman, J. & Baltimore, D. (1989) *J. Virol.* **63**, 3708–3713.
8. Pomerantz, R. J., Trono, D., Feinberg, M. B. & Baltimore, D. (1990) *Cell* **61**, 1271–1276.
9. Stinski, M. F., Thomsen, D. R., Stenberg, R. M. & Goldstein, L. C. (1983) *J. Virol.* **46**, 1–14.
10. Andreadis, A., Gallego, M. E. & Nadal-Ginard, B. (1987) *Annu. Rev. Cell Biol.* **3**, 207–242.
11. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **313**, 277–284.
12. Muesing, M., Smith, D., Cabradilla, C., Benton, C., Lasky, L. & Capon, D. (1985) *Nature (London)* **313**, 450–458.
13. Sadaie, M. R., Rappaport, J., Benter, T., Josephs, S. F., Willis, R. & Wong-Staal, F. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9224–9228.
14. Robert-Guroff, M., Popovic, M., Gartner, S., Markham, P., Gallo, R. C. & Reitz, M. S. (1990) *J. Virol.* **64**, 3391–3398.
15. Schwartz, S., Felber, B. K., Benko, D. M., Fenyo, E.-M. & Pavlakis, G. N. (1990) *J. Virol.* **64**, 2519–2529.
16. Guatelli, J. C., Gingeras, T. R. & Richman, D. D. (1990) *J. Virol.* **64**, 4093–4098.
17. Kim, S., Ikeuchi, K., Groopman, J. & Baltimore, D. (1990) *J. Virol.* **64**, 5600–5604.
18. Katz, R. A. & Skalka, A. M. (1990) *Mol. Cell. Biol.* **10**, 696–704.
19. Terwilliger, E., Sodrowski, J. G., Rosen, C. A. & Haseltine, W. A. (1986) *J. Virol.* **60**, 754–760.
20. Luciw, P. A., Cheng-Mayer, C. & Levy, J. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1434–1438.
21. Niederman, T. M. J., Thielman, B. S., Ratner, L. & Greene, W. C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9549–9553.
22. Salfeld, J., Gottlinger, H. G., Sia, R. E., Park, R. E., Sodroski, J. G. & Haseltine, W. A. (1990) *EMBO J.* **9**, 965–970.
23. Cheng-Mayer, C., Iannello, P., Shaw, K., Luciw, P. A. & Levy, J. (1989) *Science* **246**, 1629–1632.
24. Kim, S., Ikeuchi, K., Byrn, R., Groopman, J. & Baltimore, D. (1989) *Proc. Natl. Acad. Sci. USA* **87**, 9544–9548.
25. Hammes, S. R., Dixon, E., Malim, M. H., Cullen, B. R. & Greene, W. C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9549–9553.