proteins, this may imply that the function of p23 is subject to structural constraints exerted by the tryptophan residues. The observation that an analogous open reading frame of visna, designated orf Q (*29*), is also unusually rich in tryptophan residues, raises the possibility that some functional features common to these two cytopathic viruses are mediated by this region of the virus.

REFERENCES AND NOTES

1. M. Popovic, M. G. Sarngadharan, E. Read, R. C. Gallo, *Science* **224**, 497 (1984).
2. R. C. Gallo *et al.*, *ibid.*, p. 500.
3. J. Schüpbach *et al.*, *ibid.*, p. 503.
4. M. Sarngadharan *et al.*, *ibid.*, p. 506.
5. D. Klatzmann *et al.*, *ibid.* **225**, 59 (1984); F. Barré-Sinoussi *et al.*, *ibid.* **220**, 868 (1983).
6. J. A. Levy *et al.*, *ibid.* **225**, 840 (1984).
7. T. Folks *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4539 (1985).
8. L. Ratner *et al.*, *Nature (London)* **313**, 227 (1985).
9. S. Wain-Hobson, P. Sonigo, O. Danos, S. Cole, M. Alizon, *Cell* **40**, 9 (1985).
10. R. Sanchez-Pescador, *Science* **227**, 484 (1985).
11. M. A. Muesing *et al.*, *Nature (London)* **313**, 4590 (1985).
12. S. K. Arya, C. Guo, S. F. Josephs, F. Wong-Staal, *Science* **229**, 69 (1985).
13. T. Sodroski *et al.*, *ibid.*, p. 74.
14. J. S. Allan *et al.*, *ibid.* **230**, 810 (1985).
15. M. A. Gonda *et al.*, *ibid.* **227**, 173 (1985).
16. Kindly provided by R. C. Gallo and his co-workers.
17. T. H. Lee *et al.*, *Science* **226**, 57 (1984).
18. T. H. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3856 (1984).
19. L. W. Kitchen *et al.*, *Nature (London)* **312**, 367 (1984).
20. W. G. Robey *et al.*, *Science* **228**, 593 (1985).
21. F. Barin *et al.*, *ibid.*, p. 1094.
22. J. S. Allan *et al.*, *ibid.*, p. 1091.
23. T. H. Lee *et al.*, unpublished observation.
24. J. E. Coligan *et al.*, *Methods Enzymol.* **91**, 413 (1983).
25. N. T. Chang *et al.*, *Science* **228**, 93 (1985).
26. T. H. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7579 (1984).
27. A. B. Rabson *et al.*, *Science* **229**, 1388 (1985).
28. F. diMarzo Veronese *et al.*, *ibid.*, p. 1402.
29. P. Sonigo *et al.*, *Cell* **42**, 369 (1985).

30. We thank R. C. Gallo and his colleagues for reagents and B. Valas for excellent technical assistance. We also thank J. Kelley and J. Hopper for their help in preparing this manuscript. T. H. Lee is a Special Fellow of the Leukemia Society of America. J. Allan was supported by National Research Service Award ST 32 CA 9382. The research is supported by NIH grants CA 37466, CA 18216, and CA 39805.

20 November 1985; accepted 21 January 1986

Replicative and Cytopathic Potential of HTLV-III/LAV with *sor* Gene Deletions

JOSEPH SODROSKI, WEI CHUN GOH, CRAIG ROSEN, ANDRE TARTAR, DANIEL PORTETELLE, ARSENE BURNY, WILLIAM HASELTINE

The genome of the human T-lymphotropic virus type III (HTLV-III/LAV) has the potential to encode at least three polypeptides in addition to those encoded by the *gag*, *pol*, and *env* genes. In this study, the product of the *sor* (short open reading frame) region, which overlaps the 3′ end of the *pol* gene, was found to be a protein with a molecular weight of 23,000. An assay was developed for testing the ability of cloned HTLV-III proviruses to produce viruses cytopathic for T4$^+$ lymphocytes. In the cell line used, C8166, neither the HTLV-III *sor* gene product nor the complete 3′-*orf* gene product were necessary for the replication or cytopathic effects of the HTLV-III.

THE HUMAN T-LYMPHOTROPIC VIRUS type III (HTLV-III/LAV) is the primary etiological agent of the acquired immune deficiency syndrome (AIDS) and associated disorders (*1*). Infection with the virus can result in a depletion of the T4$^+$ subset of lymphocytes in patients, which results in the immune suppression characteristic of AIDS (*2*). HTLV-III can exert cytopathic effects on cultured peripheral blood lymphocytes as well as on established lymphocyte lines, mimicking the effects on the host cell in vivo (*1, 3*).

The complete nucleotide sequence of several independent isolates of the virus has been determined (*4*). In addition to the genes that encode structural components of the virus (the *gag* and *env* genes) and a gene that encodes functions required for replication (the *pol* gene) (*5*), the genome of all strains of this virus has the potential to encode at least three additional polypeptides. The tat_{III} gene, which includes two coding exons, one located just 5′ to *env* and a second located within an alternative reading frame of *env*, encodes the 14K transactivator protein (molecular weight 14,000) (*6, 7*). The product of tat_{III} induces high levels of expression of genes under the control of the HTLV-III long terminal repeat (LTR) via interaction with specific target sequences (*8*). Expression of tat_{III} in infected cells greatly stimulates the rate of virus protein expression and thereby contributes to efficient replication of the virus (*9*). A 27K protein of unknown function is encoded by the 3′-*orf* reading frame, located between the *env* gene and the 3′ LTR (*10*).

A third potential coding region in the HTLV-III genome, designated *sor* for short open reading frame, overlaps the 3′ end of the *pol* gene (*4*). Studies of virus-specific messenger RNA (mRNA) species present in infected cells revealed two spliced polyadenylated RNA species, of 5.5 and 5.0 kilobases

J. Sodroski, W. C. Goh, C. Rosen, Laboratory of Biochemical Pharmacology, Dana-Farber Cancer Institute, and Department of Pathology, Harvard Medical School, Boston, MA 02115.
A. Tartar, Pasteur Institute, Lille, France.
D. Portetelle and A. Burny, Department of Molecular Biology, University of Brussels, Brussels, Belgium.
W. Haseltine, Laboratory of Biochemical Pharmacology, Dana-Farber Cancer Institute; Department of Pathology, Harvard Medical School; and Department of Cancer Biology, Harvard School of Public Health, Boston, MA 02115.

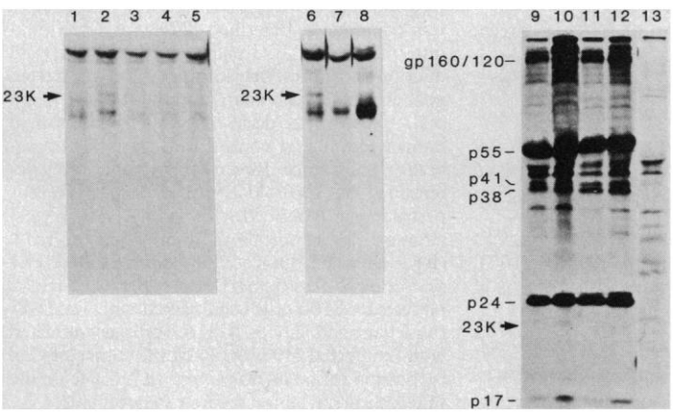

Fig. 1. Immunoprecipitation of extracts from HTLV-III–infected cells with antiserum to oligopeptide S. Cell lines were labeled overnight with [^{35}S]methionine at 100 μCi per milliliter and prepared for immunoprecipitation with rabbit antiserum to oligopeptide S (lanes 1 to 8) or sera from HTLV-III–infected individuals (lanes 9 to 13) as described (*17*). Lanes 1 and 2: uninfected H9 cells and H9 cells infected with HTLV-III (H9/IIIB), respectively, precipitated with the antiserum to oligopeptide S. Lanes 3, 4, and 5: immunocompetition of H9/IIIB cell lysates for the antiserum to oligopeptide S with increasing amounts (5, 10, and 20 μg, respectively) of oligopeptide S present in the reaction. Lanes 6, 10, and 12: Jurkat-tat_{III} cells transfected with pHXBc2. Lanes 7, 9, and 11: Jurkat-tat_{III} cells transfected with the pDsor1. Lanes 8 and 13: untransfected Jurkat-tat_{III} cells. Cells were labeled 5 to 6 days after transfection by the DEAE-dextran technique (*18*) with 15 μg of plasmid DNA. Antiserum to oligopeptide S was used to precipitate cell extracts in lanes 6 to 8, patient antiserum T.W. in lanes 9 and 10, and patient antiserum 4-3 in lanes 11 to 13.

(kb), from which the *sor* region could be expressed (*11*). Here we report that the HTLV-III *sor* region encodes a 23K protein and we discuss the role of this gene in viral replication and cytopathicity.

The nucleotide sequence of the *sor* open reading frame from several HTLV-III strains predicts that this region could encode a protein of 192 to 203 amino acids with a molecular weight of 22.5K to 23K. To determine whether a protein corresponding to the translation product of *sor* is made in HTLV-III–infected cells, a synthetic oligopeptide with 14 amino acid residues was prepared that corresponds to the predicted residues 166 to 179 (nucleotides 5083 to 5124 of the HTLV-III provirus) according to the numbering scheme of Ratner *et al.* (*4*). The oligopeptide (Thr-Pro-Lys-Lys-Ile-Lys-Pro-Pro-Leu-Pro-Ser-Val-Thr-Lys, oligopeptide S; >90 percent pure) was conjugated with tetanus toxoid and used to immunize rabbits.

The rabbit antiserum to oligopeptide S was used in an immunoprecipitation analysis of radiolabeled proteins in virus-infected cells. For this purpose, a cell line infected with HTLV-III (H9/IIIB) (*1*) was labeled for 12 hours with [^{35}S]methionine. The rabbit antiserum precipitated a protein with

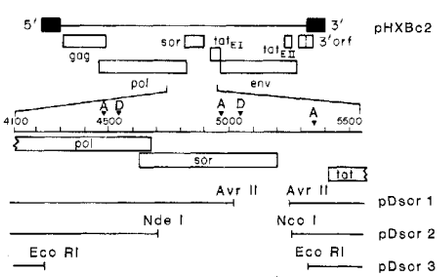

Fig. 2. Structure of the pHXBc2 proviral clone and deletion mutants. The structure of the HTLV-III provirus in plasmid pHXBc2 is shown (*12*). The dark boxes represent LTR sequences and the open boxes beneath the genome represent coding regions. The two *tat* exons are designated *tat* E1 and *tat* E2. The 3′-*orf* gene in the pHXBc2 provirus contains a stop codon (represented by a vertical broken line). The nucleotide sequence and predicted amino acid sequence near this stop codon, obtained by the method of Maxam and Gilbert (*19*), is as follows:

(8720)	CAC	ACA	CAA	GGC	TAC
	His	Thr	Gln	Gly	Tyr
TTC	CCT	GAT	TAG	CAG	AAC
Phe	Pro	Asp	*	Gln	Asn
	TAC	ACA	(8759)		
	Tyr	Thr			

The middle of the HTLV-III genome is shown in detail. The positions of splice donors (D) and acceptors (A) and nucleotide numbers are derived from the sequence of HTLV-III complementary DNA clones by Muesing *et al.* [in (*4*)]. Deletions made in pHXBc2 are shown beneath the genome. The locations of the deletion end points occur at restriction enzyme cleavage sites. All plasmid DNA's were purified by centrifugation in CsCl before being used in transfection assays.

an apparent molecular weight of 23K from extracts of the H9/IIIB cells (Fig. 1, lane 2), but not from extracts of uninfected H9 cells (lane 1). The protein was not precipitated by nonimmune rabbit serum. Oligopeptide S itself, when added in increasing amounts to the reaction, specifically competed for the ability of the antiserum to precipitate the 23K protein (Fig. 1, lanes 3 to 5).

If the 23K protein is the product of the *sor* region, viruses with the *sor* region deleted should not express it. To test this prediction, a human T-lymphocyte line, Jurkat-tat_{III}, was transfected with either plasmid pHXBc2, containing an infectious HTLV-III provirus (*12*), or plasmid pDsor1, which is identical to pHXBc2 except for a deletion within the *sor* region (Fig. 2). The Jurkat-tat_{III} line, established by infection of Jurkat cells with a recombinant retrovirus vector, constitutively expresses a functional tat_{III} protein (*13*). This cell line was chosen because when it is transfected with either pHXBc2 or pDsor1, the levels of viral replication and protein production are similar, in contrast to our observations in cell lines not expressing the tat_{III} product.

The rabbit antiserum to oligopeptide S precipitated a 23K protein from the pHXBc2 transfectants, but not from the pDsor1 transfectants or from untransfected Jurkat-tat_{III} cells (Fig. 1, lanes 6 to 8). Radiolabeled protein extracts from the transfected cells were also incubated with sera from HTLV-III–infected individuals to determine whether other viral proteins were produced after transfection with pDsor1. Serum samples from infected patients precipitated approximately equal amounts of the gp160, gp120 *env* gene products, the p55, p24-25, and p17 *gag* gene proteins, and the virus-specific proteins of 41K and 38K from cells transfected with either the pHXBc2 or the pDsor1 plasmids. Sera from some of the infected patients also apparently precipitated the 23K protein from the Jurkat-tat_{III} cells transfected with pHXBc2 but not from the cells transfected with pDsor1 (Fig. 1, lanes 9 to 13). These results demonstrate that a deletion in the *sor* region specifically curtails the expression of a 23K protein detected by rabbit antiserum directed against the predicted *sor* polypeptide. We conclude that the HTLV-III *sor* region encodes the 23K protein, which we suggest be called $p23^{sor}$.

To investigate the role of the *sor* gene in HTLV-III replication and cytopathicity, additional deletions were introduced into pHXBc2. The pHXBc2 provirus contains a termination codon at nucleotide 8744, 123 codons 3′ to the initiation site of the 3′-*orf* gene (see legend to Fig. 2). Upon transfection into peripheral blood lymphocytes, pHXBc2 produces replicating cytopathic virus (*12*). Wild-type proviruses and those containing deletions were introduced into appropriate recipient cell lines to monitor viral propagation and cytopathic effects.

The deletions introduced into pHXBc2 in the region of the *sor* gene are shown in Fig. 2. In plasmid pDsor1, the 3′ third of the *sor* gene is deleted, substituting a random open reading frame for the deleted segment. The deletion in a second plasmid, pDsor2, encompasses nearly the entire *sor* gene. Neither

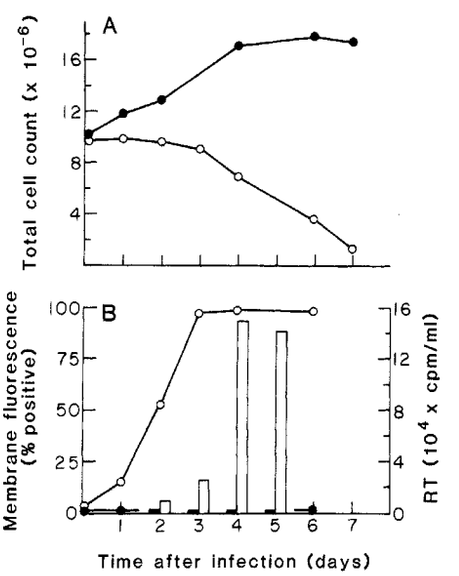

Fig. 3. (A) Infection of C8166 cells with HTLV-III. Approximately 3 × 10^6 donor H9 and H9/IIIB cells (*1*) were pelleted and resuspended in fresh RPMI 1640 medium supplemented with 20% fetal bovine serum and mitomycin C (140 μg/ml). After incubation at 37°C for 80 minutes, the cells were washed twice in phosphate-buffered saline and cocultivated for 24 hours with approximately 10^7 C8166 cells in 50% conditioned medium from the original donor cells and Polybrene (4 μg/ml). After cocultivation the cells were pelleted and resuspended in fresh medium. Medium was changed in this manner every 2 days. Only cells that excluded trypan blue were included in the total cell count. Pilot experiments indicated that less than 1% of the donor cells excluded trypan blue 48 hours after mitomycin C treatment, indicating that most of the counted cells were C8166 recipient cells. Membrane immunofluorescence was examined as described (*20*) with a 1:8 dilution of RV119 AIDS patient serum, and a 1:30 dilution of fluorescein isothiocyanate-conjugated goat antiserum to human immunoglobulin (Cappel). Experiments done in parallel with normal human serum as a control did not detect membrane fluorescence. Reverse transcriptase was determined on 100-fold concentrated cell-free supernatants as described with an oligo dT-poly(A) template-primer and magnesium cofactor (*21*). (B) Membrane fluorescence (curve) and reverse transcriptase (histogram). Open bars and circles represent C8166 cells cocultivated with H9/IIIB; solid bars and circles, C8166 cells cocultivated with uninfected H9 cells. The time course of viral replication varied by no more than 1 day from the values shown in three separate experiments.

of these deletions interfered with the coding regions of the adjacent *pol* or *tat* genes. The more extensive deletion in the pDsor3 provirus removes coding sequences at the 3′ end of the *pol* gene and extends beyond the 3′ end of the *sor* gene. The proviruses that carry these deletions express the tat_{III} gene (at about 40 percent of the level of an undeleted provirus) as judged by their ability to stimulate heterologous gene expression under the control of an HTLV-III LTR in transient transfection assays (*7*).

Cells of the C8166 $T4^+$ lymphocyte line were chosen as recipients for the assay because this line expresses markers typical of activated T cells and is exquisitely sensitive to HTLV-III infection and cytopathicity (*14*). Figure 3 shows the results of infection of C8166 cells with virus produced by the H9/IIIB prototype cell line (*1*). Just before the C8166 cells showed their maximum HTLV-III–specific positive membrane fluorescence and extracellular reverse transcriptase (RT) activity, cytopathic changes occurred that included syncytia formation, cellular enlargement, and extrusion of cell membranes (Fig. 3). The number of viable cells decreased rapidly thereafter. No such changes occurred in C8166 cultures exposed to uninfected H9 cells.

The ability of the full-length provirus and of the proviruses with *sor* gene deletions to produce infectious as well as cytopathic virus was determined. For this purpose, the pHXBc2 plasmid and plasmids with deletions were initially transfected into a human B-lymphocyte line, Raji-tat_{III}, constitutively expressing the HTLV-III *tat* protein (*13*). Viral *gag, tat*, and *env* products were readily detected at 72 hours in cells transfected with pHXBc2, pDsor1, pDsor2, or pDsor3 plasmid DNA. At this time, the transfected cells were treated with mitomycin C and cultivated with recipient C8166 cells at a ratio of 1 to 3. Within 4 days, most of the C8166 cells that had been cocultivated with cells that received pHXBc2 exhibited HTLV-III–specific membrane fluorescence (Fig. 4), and about half of them showed cytopathic changes indistinguishable from those observed after infection of C8166 cultures with the prototype virus derived from H9/IIIB cells. Six days after cocultivation with pHXBc2 transfectants, almost all of the C8166 cells showed cytopathic changes and there was a dramatic decrease in the viability of the culture. Extracellular RT activity was detected in cell-free supernatants of such cultures, albeit at lower levels than after infection of C8166 cultures (see Fig. 4 legend).

In parallel experiments, C8166 cells were cocultivated with Raji-tat_{III} cells transfected with the pDsor1, -2, and -3 plasmids. Although the levels of viral *gag* and *env* proteins transiently produced in the Raji-tat_{III} cells transfected with the deleted plasmids were the same as in the pHXBc2-transfected cells, the results of cocultivation with the C8166 recipients differed markedly. No evidence of the virus infection was detected in C8166 cells exposed to the pDsor3 transfectants. No cytopathic effect was noted and the number of viable cells was indistinguishable from that of control cultures cocultivated with mock-transfected Raji-tat_{III} cells, even after 3 weeks. A maximum of 4 percent of cells showed evidence of positive surface immune fluorescence, a value that did not increase with time. No viral DNA or proteins could be detected in these cells at 10 days after cocultivation. We conclude that the pDsor3 deletion renders the virus incapable of replication in this system.

In contrast, C8166 cells cocultivated with both the pDsor1- and pDsor2-transfected Raji-tat_{III} lymphocytes did produce virus (Fig. 4). However, the spread of virus infection in the culture was delayed. Greater than 95 percent of the cell population showed HTLV-III–specific membrane fluorescence at 10 days after exposure to the deleted proviruses; such fluorescence occurred at 4 days after exposure to pHXBc2. The C8166 cultures exposed to the deleted proviruses showed delayed cytopathic changes and longer cell viability, but nonetheless more than 95 percent of the cells died after 12 to 13 days.

The protein products present in the infected cell populations were analyzed by immunoprecipitation of cell lysates labeled at the time of maximum immunofluorescence, just prior to the onset of marked loss of cell viability (Fig. 5). When tested with serum from an AIDS patient (RV119), immunoprecipitates of H9 cells infected with HTLV-III reveal products of the *gag* gene (p55, p24-25, p17, and possibly p38) (*5*), *env* gene (gp160, gp120, and gp41) (*5*), and 3′-*orf* gene (p27) (*10*). When C8166 cells that had been exposed to pHXBc2 were tested with RV119 serum, the profile was similar except for the absence of the p27 3′-*orf* product. This was expected, because the provirus present on this plasmid contains a termination codon near the middle of the 3′-*orf* coding region (see Fig. 2). When

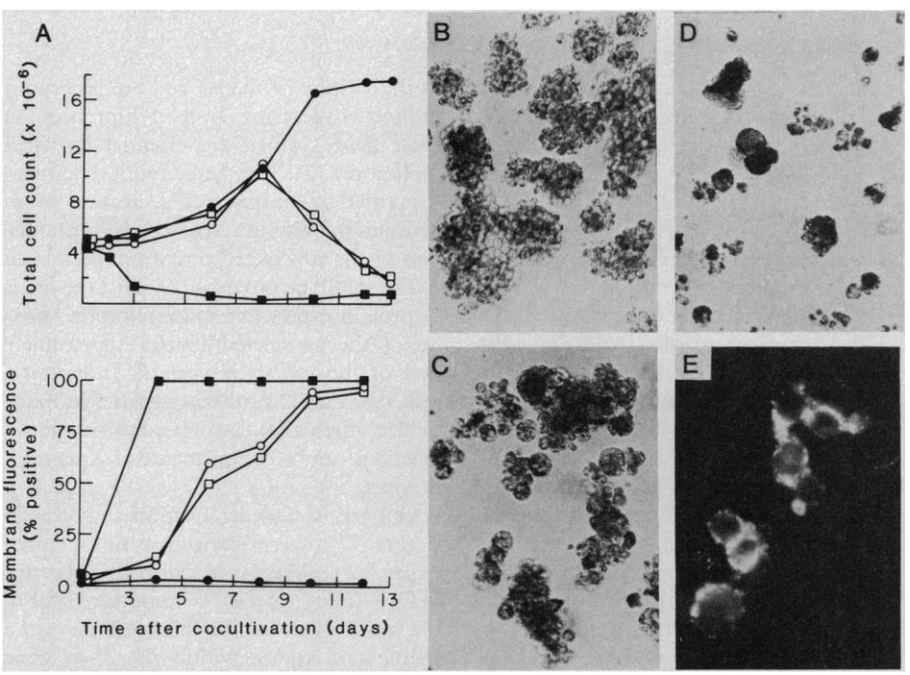

Fig. 4. (A) Infectivity and cytopathicity assays with C8166 cells. Approximately 2×10^6 Raji-tat_{III} cells constitutively expressing tat_{III} were transfected with 15 μg of plasmid DNA by the DEAE-dextran procedure (*18*). Approximately 72 hours after transfection, the cells were treated with mitomycin C and cocultivated with C8166 cells as described in Fig. 3. Plasmids used for the transfection included pHXBc2 (■), pDsor1 (□), pDsor2 (○), and pDsor3 (●). Total cell counts, reverse transcriptase assays, and membrane immunofluorescence were performed as described in Fig. 3. Reverse transcriptase levels (5×10^3 to 10×10^3 cpm/ml compared to background of 4×10^2 to 6×10^2 cpm/ml) peaked at day 5 after cocultivation with the pHXBc2 transfectants and at days 10 to 11 after cocultivation with either pDsor1 or pDsor2 transfectants. Reverse transcriptase was not elevated above background in C8166 cells cocultivated with the pDsor3 transfectants. The time course of viral replication varied by no more than 2 days from the values shown in three separate experiments. (B, C, and D) C8166 cells 9 days after cocultivation with Raji-tat_{III} cells transfected with pDsor3, pDsor1, and pHXBc2 plasmids, respectively (×160). Cytopathic effects are present in (C) and (D). (E) Typical pattern of fluorescence observed in C8166 cells 10 days after cocultivation with pDsor1 transfectants.

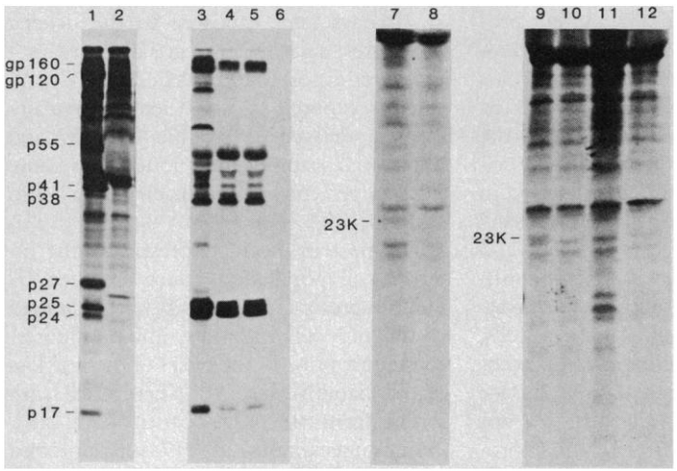

Fig. 5. Immunoprecipitation of C8166 proteins. The C8166 cells were labeled with [^{35}S]cysteine overnight at the time of maximal positive membrane fluorescence specific for HTLV-III (days 4 to 5 for C8166 cells cocultivated with pHXBc2 transfectants, days 9 to 10 for cells cocultivated with pDsor1, pDsor2, or pDsor3 transfectants). Labeled cell lysates were immunoprecipitated (*17*) with RV119, a positive AIDS patient serum (lanes 1 to 6), or rabbit antiserum to oligopeptide S (lanes 7 to 12). Sources of the lysates are H9 cells infected with HTLV-III (lanes 1 and 7), uninfected H9 cells (lanes 2 and 8), or C8166 cells cocultivated with transfected Raji-tat_{III} cells. The original transfected plasmids include pHXBc2 (lanes 3 and 9), pDsor1 (lanes 4 and 10), pDsor2 (lanes 5 and 11), or pDsor3 (lanes 6 and 12).

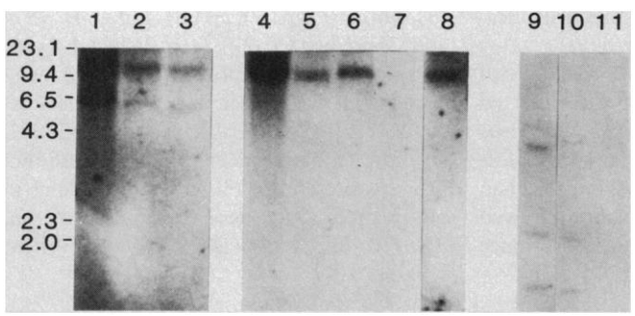

Fig. 6. Southern blot of DNA from C8166 cells. Total cellular DNA was isolated from C8166 cells at the time of maximum membrane immunofluorescence specific for HTLV-III (day 4 after cocultivation with pHXBc2-transfected Raji-tat_{III} cells; day 9 after cocultivation with Raji-tat_{III} cells transfected with pDsor1, pDsor2, or pDsor3). The DNA (10 μg) was digested with Xba I (lanes 1 to 3), Xho I (lanes 4 to 7), Bam HI (lane 8), or Hind III (lanes 9 to 11) and Southern blotted as described (*15*). The nitrocellulose filter was hybridized to a probe made from a pooled collection of gel-isolated Bgl II internal fragments derived from pHXBc2. The filter was washed in 0.2× SSC at 70°C for 2 hours before autoradiography. The original transfected plasmids used were pHXBc2 (lanes 1, 4, and 9), pDsor1 (lanes 2, 5, and 10), pDsor2 (lanes 3, 6, and 8), or pDsor3 (lanes 7 and 11). Xba I does not cut within the HTLV-III provirus, yielding apparent supercoiled and relaxed circular bands. Both Xho I and Bam HI cleave HTLV-III proviral DNA once, yielding an apparently linear form migrating slightly faster than the relaxed circular form observed in the Xba I digest. Hind III cuts the HTLV-III provirus at multiple internal sites, yielding fragments of approximately 4.2, 2.0, 1.6, and 0.6 kilobases (the latter fragment is not shown).

C8166 cells that had been exposed to the deleted proviruses pDsor1 and -2, were tested with RV119 serum, the profile was the same as that for cells exposed to pHXBc2 (Fig. 5). The 23K *sor* gene product was detected by the antiserum to oligopeptide S in C8166 cells exposed to the pHXBc2 transfectants, but not in C8166 cells exposed to pDsor1 or pDsor2.

The structure of proviral DNA present in cells infected with the intact virus or virus with *sor* region deletions was examined by Southern blot analyses of total undigested cellular DNA or DNA that had been treated with restriction enzymes (*15*). Although the same DNA forms, predominantly unintegrated closed circular and relaxed circular viral DNA forms, were present in cells infected with whole virus or virus with deletions, the amount of DNA was reduced in cells carrying the virus with *sor* deletions (Fig. 6).

These results show that the *sor* gene of HTLV-III is not required for replication of the virus or for its cytopathic effect on a T4$^+$ cell line. Viruses containing specific deletions in this gene replicate in and kill the C8166 cell line, but at a slower rate than the whole virus. Two explanations of this modest effect on virus replication can be offered. The *sor* gene itself may slightly accelerate the rate of virus replication. Alternatively, disturbance of the structure of the viral genome in the vicinity of the *pol* and *tat* genes may indirectly affect the level of expression of these genes, which are essential for virus replication (*9*). The latter interpretation is supported by the observed decrease in *trans*-activating ability in transient expression assays of the *sor*-deleted proviruses compared to the wild-type provirus (*7*). That the levels of protein expression and replication kinetics of the *sor*-deleted viruses approximate that of the wild-type virus in Jurkat-tat_{III} cells (see Fig. 1) further suggests that much of the attenuation associated with the *sor* deletions can be complemented by the tat_{III} product.

Our results also demonstrate that expression of the intact 3′-*orf* gene is not required either for replication or cytopathic effect of HTLV-III in the T4$^+$ cell line used. All of the proviruses used for this work contain a termination codon within the 3′-*orf* gene and do not make detectable levels of the 27K 3′-*orf* protein.

A function essential for virus replication is apparently deleted in the pDsor3 plasmid. This deletion removes 148 amino acids from the 3′ end of the *pol* gene. The deletion does not remove any coding sequences of tat_{III}, and the provirus that carries this deletion expresses tat_{III} function (*7*). The 3′ end of the *pol* gene of other retroviruses encodes an endonuclease or integrase function (*16*). Similarities in the predicted amino acid sequence of this region of the HTLV-III genome to analogous regions of other retroviruses suggest that the function of this region is conserved (*4*). For this reason, we speculate that inactivation of the endonuclease-integrase domain of the *pol* gene prevents virus replication.

The surprisingly mild effect of mutations in the *sor* and 3′-*orf* genes on the replication and cytopathic activity of the virus poses questions regarding the role of these genes in the natural life cycle of the virus. That the *sor* protein plays some role is suggested by the relatively high degree of conservation of this gene in independent virus isolates (*4*). One possibility is that this protein may be required for growth or regulation in cell types not examined here. Assays similar to those described here but with the target cells being derived from the central nervous or reticuloendothelial system may reveal novel functions for the *sor* and 3′-*orf* products.

REFERENCES AND NOTES

1. R. C. Gallo *et al.*, *Science* **224**, 500 (1984); M. Popovic, M. G. Sarngadharan, E. Read, R. C. Gallo, *ibid.*, p. 497; M. G. Sarngadharan, M. Popovic, L. Bruch, J. Schüpbach, R. C. Gallo, *ibid.*, p. 506; F. Barré-Sinoussi *et al.*, *ibid.* **220**, 868 (1983).
2. M. Seligman *et al.*, *N. Engl. J. Med.* **311**, 1286 (1984); R. W. Schroff, M. S. Gottlieb, H. E. Prince, L. L. Chai, J. L. Fahey, *Clin. Immunol. Immunopathol.* **27**, 300 (1983); H. C. Lane *et al.*, *N. Engl. J. Med.* **309**, 453 (1983); J. L. Fahey *et al.*, *Am. J. Med.* **76**, 95 (1984).
3. T. Folks *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4539 (1985); D. Klatzmann *et al.*, *Science* **225**, 59

(1984); E. Vilmer *et al.*, *Lancet* **1984-I**, 753 (1984); S. Harada, Y. Koyanagi, N. Yamamoto, *Science* **229**, 563 (1985).
4. L. Ratner *et al.*, *Nature (London)* **313**, 277 (1985); S. Wain-Hobson, P. Sonigo, O. Danos, S. Cole, M. Alizon, *Cell* **40**, 9 (1985); R. Sanchez-Pescador *et al.*, *Science* **227**, 484 (1985); M. A. Muesing *et al.*, *Nature (London)* **313**, 480 (1985).
5. W. G. Robey *et al.*, *Science* **228**, 593 (1985); F. diMarzo Veronese *et al.*, *ibid.* **229**, 1402 (1985); L. Kitchen *et al.*, *Nature (London)* **312**, 367 (1984); J. Schüpbach *et al.*, *Science* **224**, 503 (1984); J. S. Allan *et al.*, *ibid.* **228**, 1091 (1985).
6. W. C. Goh, C. Rosen, J. Sodroski, D. Ho, W. Haseltine, in preparation.
7. J. Sodroski *et al.*, *Science* **227**, 171 (1985); S. K. Arya, C. Guo, S. F. Josephs, F. Wong-Staal, *ibid.* **229**, 69 (1985); J. Sodroski, R. Patarca, C. Rosen, F. Wong-Staal, W. Haseltine, *ibid.*, p. 74; J. Sodroski, C. Rosen, W. Haseltine, unpublished observations.
8. C. A. Rosen, J. G. Sodroski, W. A. Haseltine, *Cell* **41**, 813 (1985).
9. A. Dayton, J. Sodroski, C. Rosen, W. C. Goh, W. Haseltine, *ibid.*, in press.
10. J. S. Allan *et al.*, *Science* **230**, 810 (1985).
11. A. B. Rabson *et al.*, *ibid.* **229**, 1388 (1985).
12. G. M. Shaw *et al.*, *ibid.* **226**, 1165 (1984); A. G. Fisher, E. Collalti, L. Ratner, R. C. Gallo, F. Wong-Staal, *Nature (London)* **316**, 262 (1985).
13. C. A. Rosen, J. G. Sodroski, K. Campbell, W. A. Haseltine, *J. Virol.* **57**, 379 (1986).
14. S. Z. Salahuddin *et al.*, *Virology* **129**, 51 (1983).
15. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
16. D. P. Grandgenett, A. C. Vora, R. D. Schiff, *Virology* **89**, 119 (1978); D. P. Grandgenett, R. J. Kraus, P. J. Hippenmeyer, *ibid.* **130**, 257 (1983); D. Schwartzberg *et al.*, *Cell* **37**, 1043 (1984); L. A. Donehower and H. E. Varmus, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6461 (1984); P. Hippenmeyer and D. Grandgenett, *Virology* **137**, 358 (1984); A. T. Panganiban and H. M. Temin, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7885 (1984).
17. T. H. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3856 (1984).
18. C. Queen and D. Baltimore, *Cell* **33**, 729 (1983).
19. A. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1977).
20. M. Essex *et al.*, *Science* **220**, 859 (1983); M. Essex *et al.*, *ibid.* **221**, 1061 (1983).
21. H. M. Rho, B. J. Poiesz, F. W. Ruscetti, R. C. Gallo, *Virology* **112**, 355 (1981).
22. We thank F. Wong-Staal, B. Hahn, G. Shaw, and R. Gallo for providing the pHXBc2 plasmid; D. Bolognesi, M. Hirsch, D. Ho, and J. Groopman for patient sera; Z. Salahuddin for providing the C8166 cell line; A. Dayton, M. F. McLane, J. Allen, T.-H. Lee, M. Cloyd, and D. Celander for discussions; K. Campbell and R. Zaya for technical assistance; and C. Owens and T. Parks for editorial assistance. This work was supported by grants to W.H. from the American Cancer Society and the State of Massachusetts. J.S. is a Special Fellow of the Leukemia Society of America and a recipient of NIH grant CA40658. C.R. is supported by NIH grant CA07580.

12 December 1985; accepted 27 January 1986

Identification of HTLV-III/LAV *sor* Gene Product and Detection of Antibodies in Human Sera

NANCY C. KAN, GENOVEFFA FRANCHINI, FLOSSIE WONG-STAAL, GARRETT C. DUBOIS, W. GERARD ROBEY, JAMES A. LAUTENBERGER, TAKIS S. PAPAS*

The nucleotide sequence of the genome of HTLV-III, the infectious agent etiologically associated with the acquired immune deficiency syndrome, predicts a small open reading frame, termed *sor*, located between the *pol* and *env* genes. A DNA segment containing 82 percent of the *sor* region was inserted into a prokaryotic expression vector, pJL6, to determine whether *sor* encodes a viral protein and to gain some insight into its possible function. The bacterially synthesized *sor* protein reacted with sera from individuals infected with HTLV-III, indicating that *sor* is expressed as a protein product or products that are immunogenic in vivo. Antibodies to the purified, bacterially synthesized *sor* protein were found to react specifically with the same protein and also with a protein of molecular weight 23,000 (23K) in HTLV-III–infected H9 cell extracts. The 23K protein comigrated with a protein immunoprecipitated by the serum of a hemophiliac patient with antibodies to HTLV-III, suggesting that this protein is probably the *sor* gene product.

HUMAN T-LYMPHOTROPIC VIRUS type III (HTLV-III/LAV) has been implicated as the causative agent of the acquired immune deficiency syndrome (AIDS) in humans (*1–5*). The published nucleotide sequences of three isolates of HTLV-III and related viruses predict the characteristic retroviral *gag*, *pol*, and *env* genes as well as two small open reading frames, *sor* and 3′-*orf* (*6–8*). The two small open reading frames in the HTLV-III genome are unusual in that they have no equivalents similar in size and location in the genomes of other exogenous human retroviruses such as HTLV-I and HTLV-II (*9, 10*). The *sor* region is located between the *pol* and *env* genes and can encode a protein of 192 amino acids with an estimated molecular weight of 22,500 (22.5K). Its initiator methionine overlaps with the end of *pol* and its terminator precedes the beginning of *env*. The 440 nucleotides between *sor* and *env* contain additional open reading frames that have the potential to encode small novel proteins. One of these has been shown to constitute the second exon of the *trans*-activating gene (tat_{III}) (*11, 12*). The putative initiator methionine of the 3′-*orf* is located 4 base pairs downstream from the stop codon of the *env* protein. Its open reading frame extends into the U3 element of the 3′ long terminal repeat (LTR). The native gene product of 3′-*orf* in HTLV-III–infected cells has been identified as a 27K protein and shown to be immunogenic in vivo (*13, 14*). In this report, we show that the HTLV-III *sor* gene product is a 23K protein that is also immunogenic in vivo.

HTLV-III is an exogenous retrovirus that infects mainly the $T4^+$ human T cells. Unlike most other nondefective retroviruses, infection with HTLV-III generally leads to cell death in vitro and to depletion of T cells with OKT4 phenotype in patients with AIDS (*2, 5, 15*). The small open reading frames may contribute to this cytopathic effect by encoding viral proteins that either directly or indirectly alter the cell metabolism. DNA probes derived from the *sor* region hybridize to RNA species in HTLV-III–infected cells (*12, 16*). However, whether the *sor*-related messenger RNA's (mRNA's) are translated into distinct protein species remains to be studied, and the functions of the putative *sor* gene products are unknown.

In one approach to these problems we have expressed the *sor* region in *Escherichia coli*. The Alu I–Alu I fragment of the BH10 clone (*17*) of HTLV-III corresponding to the segment from nucleotides 4724–5201 [numbered according to Ratner *et al.* (*6*)] was isolated from an Eco RI subclone. The 5′ Alu I site is located 34 amino acids downstream from the putative initiator methionine, while the 3′ Alu I site coincides with the termination codon of *sor*. Therefore, only 35 amino acids may have been deleted with respect to the native protein. The Alu I fragment was joined to the expression vector pJL6 (*18, 19*) at the Hind III site (Fig. 1). Thirteen amino acids at the NH_2-

N. C. Kan, J. A. Lautenberger, T. S. Papas, Laboratory of Molecular Oncology, National Cancer Institute, National Institutes of Health, Frederick, MD 21701-1013.
G. Franchini and F. Wong-Staal, Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892.
G. C. DuBois, Program Resources, Inc., Frederick Cancer Research Facility, Frederick, MD 21701-1013.
W. G. Robey, Virus Control Unit, National Cancer Institute, National Institutes of Health, Frederick, MD 21701-1013.

*To whom requests for reprints should be addressed.