Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 2209–2213, April 1986
Medical Sciences

Three novel genes of human T-lymphotropic virus type III: Immune reactivity of their products with sera from acquired immune deficiency syndrome patients

(*sor, tat* and *3' orf* genes/cDNA cloning/double splicing/*in vitro* translation/immunoprecipitation)

SURESH K. ARYA* AND ROBERT C. GALLO

Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205

Communicated by Peter C. Nowell, November 13, 1985

ABSTRACT Human T-lymphotropic virus type III or lymphoadenopathy associated virus (HTLV-III/LAV) is the cause of acquired immune deficiency syndrome (AIDS). In addition to the conventional retroviral genes involved in virus replication, namely, *gag, pol,* and *env* genes, DNA sequence analysis of HTLV-III genome predicted two additional open reading frames termed by us short open reading frame (*sor*) and 3' open reading frame (*3' orf*). Furthermore, functional analysis revealed another gene with transactivating function, termed *tat*. We have now structurally identified and functionally characterized these HTLV-III specific genes by way of cDNA cloning. DNA sequence analysis of the clones shows that the *tat* and *3' orf* genes contain three exons and their transcription into functional mRNA involves two splicing events and that the *sor* gene contains at least two exons. *In vitro* transcription and translation of the cloned spliced sequences show that the *sor, tat,* and *3' orf* genes code for polypeptides with apparent mobility of 24–25 kDa, 14–15 kDa, and 26–28 kDa, respectively. All three polypeptides are immune reactive and are immunogenic in the natural host. The results demonstrate that the three extra open reading frames of HTLV-III, two of which are unique to HTLV-III, are in fact genes that function *in vivo* and further allow the identification of three new and previously unrecognized HTLV-III antigens with differential immunogenicity in individuals with acquired immune deficiency syndrome and related disorders.

Human T-lymphotropic virus type III (HTLV-III) or the lymphoadenopathy associated virus (LAV) is etiologically linked to acquired immune deficiency syndrome (AIDS) and AIDS-related complex (ARC) (1–4). The overall genetic structure of HTLV-III/LAV is similar to that of other animal retroviruses. However, besides *gag, pol,* and *env* genes, DNA sequence analysis of HTLV-III/LAV genome predicted two additional open reading frames or potential genes (5–8), termed by us and others *sor* (short open reading frame) and *3' orf* (3' open reading frame). The presence of a third gene, termed *tat* (transactivation of transcription), was also suggested (9–12). Thus, HTLV-III/LAV contains coding potential for three genes that are specific to this virus. Two of these putative genes, *sor* and *3' orf*, are unique to HTLV-III but a functional analog of the third gene, *tat*, is also carried by other members of the HTLV-bovine leukemia virus (BLV) group of retroviruses (9–13). We and others have localized the *tat* gene of HTLV-III to a region between the putative *sor* and the *env* genes (10, 12), a region of the genome previously thought to be noncoding. This is distinct from the other members of the HTLV-BLV group where *tat* gene is located downstream from the *env* gene. Thus, even the *tat* gene is organized differently in HTLV-III. We report here that *sor, tat,* and *3' orf* genes all contain intron(s) and are respectively translated into polypeptides with apparent mobility of 24–25 kDa, 14–15 kDa, and 26–28 kDa on $NaDodSO_4$/PAGE. These gene products display differential immune reactivity for HTLV-III positive human sera, the *3' orf* gene product being the most immune reactive. The results demonstrate the existence of three new HTLV-III antigens.

MATERIALS AND METHODS

cDNA Cloning and DNA Sequencing. Poly(A)-selected RNA from HTLV-III-infected H4 cells, isolated as described (14, 15), was used to construct cDNA libraries as reported (10). The libraries were screened with subgenomic HTLV-III probes to obtain clones containing specific HTLV-III sequences (10, 11). The selected clones were characterized by restriction mapping and DNA sequencing by the method of Maxam and Gilbert (16).

***In Vitro* Transcription and Translation.** The inserts of selected cDNA clones were transferred to the vector pSP6 that transcribes inserted DNA under the influence of SP6 promoter (17). RNA was transcribed *in vitro* after linearization of the plasmid DNA with specific restriction enzymes. It was translated *in vitro* by using rabbit reticulocyte translation system and [^{35}S]methionine, and the products were analyzed by 12% $NaDodSO_4$/PAGE and radioautography by the standard procedures.

Immunoprecipitation with Human Sera. The *in vitro* translation products were incubated with normal human serum for 1–2 hr at 4°C. Suspension of *Staphylococcus aureus* (Staph A) cells was then added, and incubation was continued for an additional 1 hr. The sample was centrifuged, and the supernatant was divided into two equal parts, one of which was incubated with immune serum at 4°C for 18–24 hr. A suspension of Staph A was added to each sample and incubated at 4°C for 1 hr. The samples were centrifuged, and the pellets were repeatedly and sequentially washed with 50 mM Tris·HCl (pH 7.4)/50 mM EDTA/0.05% Nonidet P-40/1% aproteinin containing 0.5 M NaCl or 0.15 mM NaCl. The pellets were suspended in 75 mM Tris·HCl (pH 6.8)/0.7 mM 2-mercaptoethanol/2% (wt/vol) $NaDodSO_4$/10% (vol/vol) glycerol/0.001% bromophenol blue, boiled for 10 min, and centrifuged. The supernatants were subjected to 12% $NaDodSO_4$/PAGE analysis.

RESULTS

cDNA Clones of HTLV-III Specific Genes. To identify HTLV-III specific genes, we took the direct approach of

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.

Abbreviations: HTLV-III/LAV, human T-lymphotropic virus type III/lymphoadenopathy associated virus; AIDS, acquired immune deficiency syndrome; bp, base pair(s); kb, kilobase(s).
*To whom correspondence should be addressed.

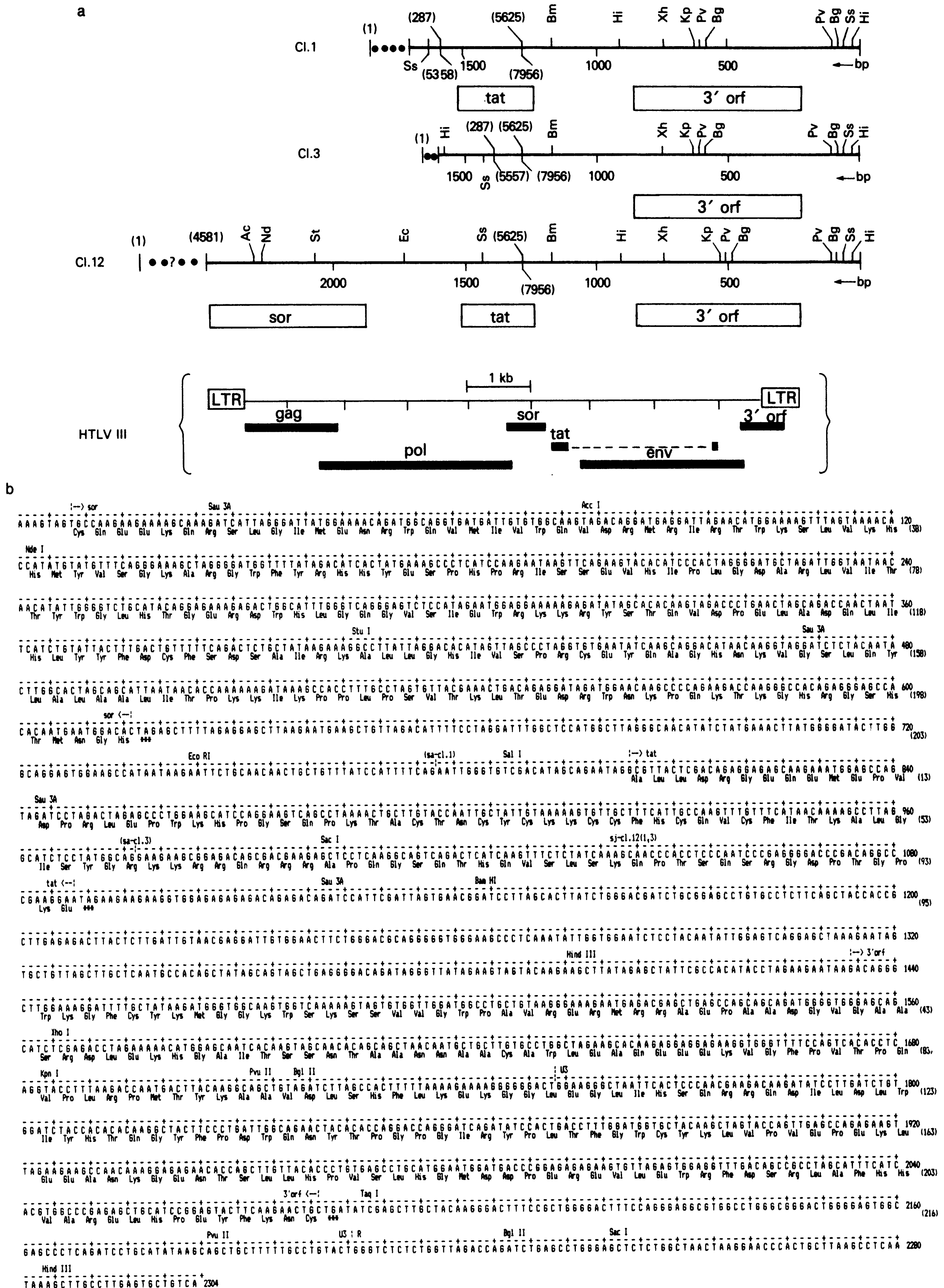
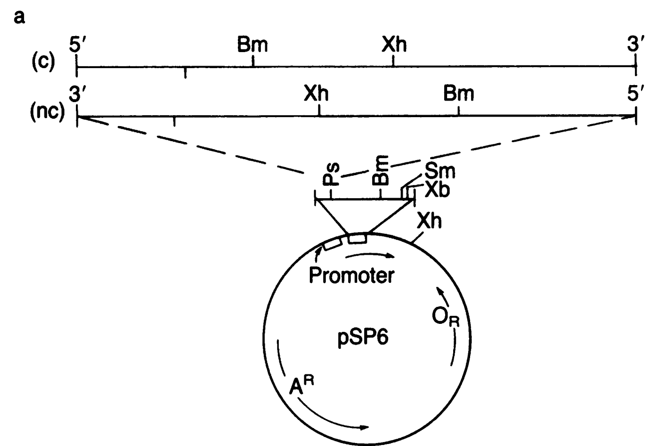


FIG. 1. *(Legend appears at the bottom of the opposite page.)*

obtaining functional cDNA clones by screening cDNA libraries with specific subgenomic HTLV-III probes to obtain the desired clones. We have previously described a functional cDNA clone (clone 1) corresponding to the mRNA of the *tat* gene (10). We also described in this previous report a second cDNA clone (clone 3) that we speculated may correspond to the mRNA of *3' orf* gene. Both of these clones were copies of the mRNAs that were generated by double splicing events. Thus, the *tat* gene and the putative *3' orf* gene consisted of three exons and two introns (Fig. 1). We have now obtained another cDNA clone (clone 12) that contains the complete open reading frame of the putative *sor* gene, in addition to the open reading frames of the *tat* gene and the putative *3' orf* gene (Fig. 1). DNA sequence analysis of clone 12 [2304 base pairs (bp)] showed it to be an incomplete cDNA clone as it lacked the mRNA cap site and possibly other sequences on the 5' side of the *sor* open reading frame. However, it contained the 3'-splice junction that was identical to the 3'-splice junction of clones 1 and 3 (Fig. 1).

Translation Products of HTLV-III Specific Genes. To characterize the gene products of the putative *sor*, *tat*, and *3' orf* genes, the cDNA were transferred to the transcription vector pSP6, and the plasmids containing clones 1, 3, and 12 cDNA inserts were designated pSP-1, pSP-3, and pSP-12, respectively (Fig. 2*a*). RNA was transcribed after linearization of the plasmid DNAs with specific restriction enzymes that were chosen because they will either retain a given open reading frame as a part of pSP6 transcriptional unit or delete it. The transcription of the plasmid DNAs cleaved with specific restriction enzymes gave RNA transcripts of the appropriate sizes (data not shown). These transcripts were translated and products analyzed. Representative results are shown in Fig. 2. The transcripts of pSP-1 DNA linearized with *Xba* I or *Sma I* gave two polypeptides with apparent mobility of 25–26 kDa and 14–15 kDa, the 14–15 kDa polypeptide being in smaller relative amounts. Digestion of this plasmid DNA with *Bam*HI or *Xho* I, which deletes *3' orf* open reading frame from the transcriptional unit, gave only the polypeptide with 14–15 kDa apparent mobility. These results suggest that 25–26 kDa and 14–15 kDa polypeptides were products of the *3' orf* and *tat* open reading frames, respectively. While pSP-12 DNA linearized with *Xba* I displayed three polypeptides of 25–26 kDa, 23–24 kDa, and 14–15 kDa, this DNA linearized with *Bam*HI or *Xho* I gave only two polypeptides of 23–24 kDa and 14–15 kDa apparent mobility (Fig. 2). These results again suggest that 25–26 kDa and 14–15 kDa polypeptides are the product of the *3' orf* and *tat* open reading frames, respectively, and further suggest that 23–24 kDa polypeptide is the product of the *sor* open reading frame.

The transcripts of *Xba* I linearized pSP-3 DNA, which contains only the *3' orf* open reading frames, though not always translated efficiently, displayed a distinct polypeptide with apparent mobility of 27–28 kDa. This polypeptide was not detected when pSP-1 DNA was linearized with *Bam*HI or *Xho* I, which removes the *3' orf* open reading frame from the transcriptional unit. These results suggest that the *3' orf* open reading frame contained in pSP-3 DNA was being translated into a 27–28 kDa polypeptide. The plasmid DNAs containing cDNA inserts in the incorrect orientation with respect to the SP6 promoter gave transcripts of the appropriate sizes but none of these transcripts were translated into distinct polypeptides (Fig. 2).



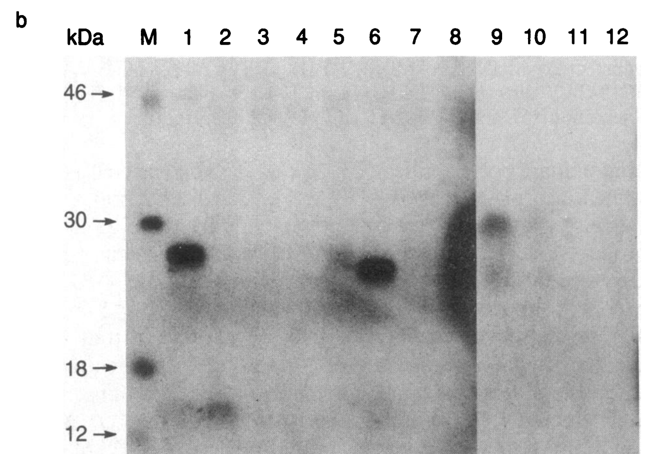


FIG. 2. (*a*) Physical map of pSP-1 containing HTLV-III cDNA clone 1. pSP-3 and pSP-12 were similarly constructed. (c) and (nc) refer to the correct and noncorrect orientation of the cDNA insert with respect to the SP6 promoter. (*b*) $NaDodSO_4$/PAGE analysis of the translation products of the transcripts from pSP-1, pSP-12, and pSP-3 plasmid DNAs. Lanes 1 and 2, *Xba* I- and *Bam*HI-digested pSP-1(c) DNA; lanes 3 and 4, *Xba* I- and *Bam*HI-digested pSP-1(nc) DNA; lanes 5 and 6, *Xba* I- and *Bam*HI-digested pSP-12(c) DNA; lanes 7 and 8, *Xba* I- and *Bam*HI-digested pSP-12(nc) DNA; lanes 9 and 10, *Xba* I- and *Bam*HI-digested pSP-3(c) DNA; lanes 11 and 12, *Xba* I- and *Bam*HI-digested pSP-3(nc) DNA; lane M, molecular size standards.

Since clone 12 contained the *tat* open reading frame in addition to the *sor* and *3' orf* open reading frames, we tested its transactivating capacity in a transfection system that measures transactivation of the chloramphenicol acetyl transferase (CAT) gene (see ref. 10). Representative results for human lymphoid JM cells are shown in Fig. 3. Clearly, clone 12 DNA transactivated the *CAT* gene activity. Thus, the *tat* open reading frame contained in clone 12 was transcribed and translated into a functionally active polypeptide.

Immune Reactivity of HTLV-III Specific Gene Products. To evaluate the immune reactivity of the polypeptides directed by the *sor*, *tat*, and *3' orf* open reading frames, translation products were immune precipitated with HTLV-III-positive human sera from several individuals. Representative results are shown in Figs. 4 and 5, and data are compiled in Table 1. HTLV-III-positive serum specifically immune precipitated a

FIG. 1 (*on opposite page*). (*a*) Physical maps of HTLV-III cDNA clones 1, 3, and 12. The two splice junctions for clones 1 and 3 and one splice junction for clone 12 are indicated. The nucleotide numbering in parentheses is according to Ratner *et al.* (6). (*b*) DNA sequence of HTLV-III cDNA clone 12. The three open reading frames contained in this clone along with the predicted amino acid sequences are shown. DNA sequences of clones 1 and 3 have been reported before (10). The open reading frame for the *tat* gene is in a different frame than those of the *sor* and *3' orf* genes.

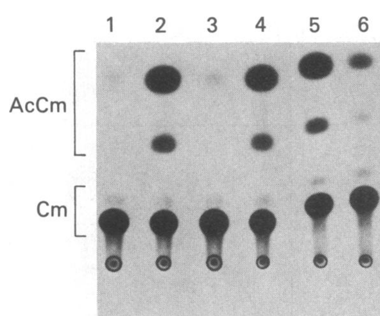


FIG. 3. Enhancement of HTLV-III LTR-promoted *CAT* gene expression by HTLV-III cDNA clone 12. Human lymphoid JM cells were cotransfected with clone 12 DNA in expression vector pCV (pCV-12) and HTLV-III LTR-CAT (pC15-CAT) plasmid DNA by the DEAE-dextran protocol (10). *CAT* gene product activity in the extract of transfected cells was measured by analyzing the conversion of [^{14}C]chloramphenicol (Cm) into its acetylated forms (AcCm) by thin layer paper chromatography. Lanes 1 to 6 are respectively for cells transfected with DNAs of pSVoCAT (1), pRSVCAT (2), pC15-CAT (3), pC15-CAT plus pCV-HXb3 (4), pC15-CAT plus pCV-12 (correct orientation) (5), and pC15-CAT plus pCV-12 (incorrect orientation) (6).

predominant polypeptide of 25–26 kDa for *Xba* I as well as *Sma* I linearized plasmid pSP-1 DNA. The 25–26 kDa polypeptide was also specifically immune precipitated from the translation products of *Xba* I as well as *Sma* I linearized plasmid pSP-12 DNA. Similar results were obtained with plasmid pSP-3 DNA, except the apparent size of this polypeptide was 27–28 kDa (Fig. 4). The marginal detection of this polypeptide in translation products of *Bam*HI-digested pSP-1 and pSP-3 plasmid DNAs was probably the result of incomplete enzyme digestions; it was not detected for *Bam*HI-digested pSP-12 plasmid DNA. Instead, translation products of *Bam*HI-digested pSP-12 plasmid DNA displayed a band at 23–24 kDa that was immune precipitated with HTLV-III-positive serum but also to a lesser extent with some normal human sera (see Table 1). Consistent with our interpretation of the translation products noted above, we infer that the 26–28 kDa and 23–24 kDa polypeptides are the immune reactive products of *3' orf* and *sor* open reading frames, respectively. Immunoprecipitation of the 14–15 kDa *tat* gene product was not obvious with this particular HTLV-III-positive serum but could be detected to varying extent by some of the other HTLV-III-positive sera as shown in Fig. 5 and listed in Table 1.

DISCUSSION

The HTLV-III open reading frames termed *sor*, *tat*, and *3' orf* are specific to this virus and two of these, *sor* and *3' orf*, are

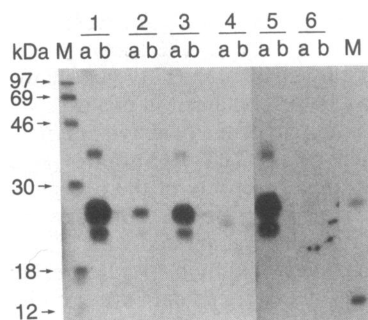


FIG. 4. NaDodSO$_4$/PAGE analysis of immune precipitates of translation products of pSP-1, pSP-12, and pSP-3 DNA transcripts. Lanes 1 and 2, *Xba* I- and *Bam*HI-digested pSP-1 DNA; lanes 3 and 4, *Xba* I- and *Bam*HI-digested pSP-12 DNA; lanes 5 and 6, *Xba* I- and *Bam*HI-digested pSP-3 DNA. Sublanes (a) and (b) are for HTLV-III-positive and normal human serum, respectively. Lane M, molecular size standards.

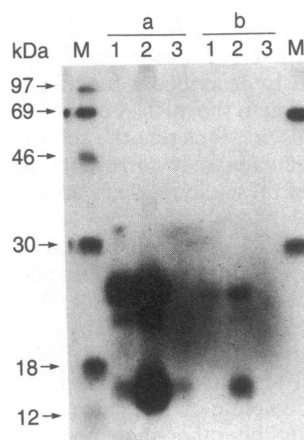


FIG. 5. NaDodSO$_4$/PAGE analysis of immune precipitates of translation products of *Xho* I-digested pSP-1 (*a*) and pSP-12 DNA (*b*) transcripts. Lanes 1 and 2 are for two different immune sera and lane 3 is for normal serum. Lane M, molecular size standards. Analysis was performed as described in Fig. 3. (The 25–26-kDa band is the *3' orf* gene product, presumably the result of incomplete enzyme digestion of the plasmid DNA.)

unique to it. The products of the three HTLV-III specific open reading frames immune react with antibodies in sera of individuals with AIDS and ARC. Therefore, these open reading frames are in fact genes that are expressed *in vivo*.

Our results allow the structural definition of the three HTLV-III specific genes. We have previously characterized the functional domain of the *tat* gene (10). Like the *tat* gene, the *3' orf* gene (clone 3) also consists of three exons (287 bp, 69 bp, and 1258 bp) and two introns (5268 bp and 2330 bp), and its transcription into a functional mRNA involves double splicing. The *3' orf* gene differs from the *tat* gene in having a truncated second exon involving splicing out of the putative initiation codon of the *tat* gene product (10).

The *sor* gene contains at least two and probably three exons. The 3' exon (1258 bp) of this gene (clone 12) is identical to the third exon of the *tat* and *3' orf* genes. The sequences on the 5' side (1114 bp) of this exon in clone 12 are shared with the second exon of the *tat* gene and extend upstream to include *sor* open reading frame. We suspect that the generation of the *sor* mRNA also involves two splicing events. It is possible that the synthesis of this mRNA involves the same first donor site (at nucleotide 287) as other mRNAs and one of the many potential acceptor sites located to the 5' side of the *sor* open reading frame. If the consensus acceptor site nearest to the 5' side of the *sor* open reading frame located at nucleotide 4494 is utilized, the functional *sor* gene will generate a message of about 2.7 kilobases (kb). However, if the *sor* message involves only one splicing event demonstrated in clone 12, the mRNA would be about 7.0 kb.

Table 1. Immune reactivity of the *sor*, *tat*, and *3' orf* gene products

Serum		Gene product		
Number	Diagnosis	*sor*	*tat*	*3' orf*
1	AIDS	+	±	++
2	AIDS	+	+	++
3	AIDS	+	±	++
4	AIDS	±	—	±
5	ARC	+	++	++
6	ARC	+	—	++
7	Healthy homosexual	—	—	—
8	Healthy homosexual	—	—	++
9	Healthy homosexual	—	—	—
10	Healthy homosexual	—	—	—
11	Healthy heterosexual	±	—	—
12	Healthy heterosexual	+	—	—
13	Healthy heterosexual	—	—	—
14	Healthy heterosexual	—	—	—

+, Reactive; ±, detectable; ++, strongly reactive; —, not detected.

Muesing *et al.* (8) have suggested that the *sor* gene consists of two exons generating a message of about 5.0 kb. Their suggestion is inconsistent with clone 12 that contains an intron located within their suggested second exon. It is possible to postulate other combinations of potential 5′ donor and acceptor splice sites to generate a 5-kb message involving double splicing. It is, of course, possible that more than one species of the *sor* mRNA is synthesized utilizing alternative splicing events. We have previously reported four abundant mRNAs of 9.4 kb, 4.2 kb, 2.0 kb, and 1.8 kb in HTLV-III-infected cells (10, 11). We also observed other less abundant RNA species of about 7 kb, 5 kb, 3.2 kb, and 2.8 kb in these cells. One or more of these species could correspond to the *sor* message.

The *sor*, *tat*, and *3′ orf* genes synthesize polypeptides with apparent mobilities of 23–24 kDa, 14–15 kDa, and 26–28 kDa, respectively. The *3′ orf* open reading in clone 3 and in clones 1 and 12 was translated into a polypeptide of 27–28 kDa and 25–26 kDa, respectively. This open reading frame contains two initiation codons (ATG) 57 bp apart in phase in its 5′ portion (Fig. 1). We suggest that the first and second ATGs are used for translation in pSP-3 DNA and pSP-1 and pSP-12 DNAs, respectively. Both of these ATG triplets are flanked by the appropriate consensus sequence requisite for efficient translation initiation by the eukaryotic ribosomes (18). Furthermore, the coding potential of the open reading frames for the *sor*, *tat*, and *3′ orf* genes, starting from the first in phase initiation codon is respectively 192, 86, and 206 amino acid residues, predicting the respective polypeptides of about 20 kDa, 9 kDa, and 21 kDa. The observed mobility of the products of these genes in $NaDodSO_4$/PAGE was uniformly higher than predicted. This may suggest anomalous conformation and/or posttranslational modifications of the proteins.

The products of the *sor*, *tat*, and *3′ orf* genes are immunogenic *in vivo*, thus identifying three new antigens for HTLV-III, in addition to the previously described *gag* and *env* gene products (19–22). The three gene products appear to be differentially immunoreactive and immunogenic, the *3′ orf* gene product apparently being the most potent and the *sor* gene product being the least potent in this regard. The lesser immunogenicity of the *sor* gene product may be due to its diminished expression *in vivo* and its particular intracellular localization, or it may be related to its structure (Fig. 6). The predicted amino acid sequence of the *sor* gene product does not contain a cluster of amino acid residues that will impart to this protein hydrophilic structure with β-turns-two parameters generally thought to be responsible for strong immunogenicity (23, 24). Notably, the predicted amino acid sequence of both the *sor* and *tat* gene products lacks typical sequence (-aspargine-Xaa-threonine or serine-) that generally serves as a glycosylation site and such a sequence is present twice in the predicted sequence of *3′ orf* gene product.

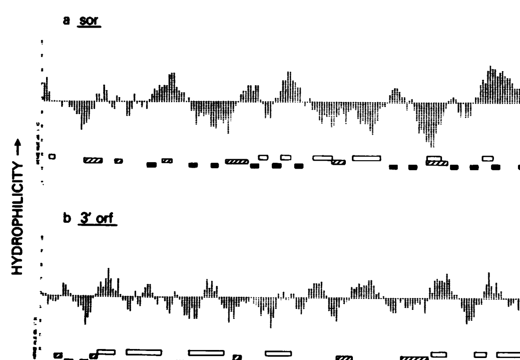


FIG. 6. Hydrophobicity profile and predicted secondary structure of *sor* (*a*) and *3′ orf* (*b*) polypeptides analyzed according to Kyte and Doolittle (23) and Chou and Fasman (24). Secondary structure is depicted by boxes and vertical lines represent amino acid residues. Open box, α-helix; hatched box, β-sheets; closed box, β-turns.

With regard to any correlation between the progression of the disease and expression of the HTLV-III specific genes, the survey reported here is too small to detect meaningful trends. We think it is premature to draw conclusions from the observation that antibodies to the *3′ orf* gene product were detected in all but one of the six sera from patients with AIDS and ARC but in only one out of four sera from HTLV-III-positive healthy homosexual individuals in this study. Further, some of the normal human sera reacted, though poorly, with the *sor* gene product. Although we cannot presently rule out artifactual interactions, this may suggest that a normal cellular gene with some homology to the *sor* gene exists, and its product is synthesized in some instances. The differential expression of the *sor*, *tat*, and *3′ orf* genes *in vivo* may reflect mutual modulatory role(s) of the products of these genes.

We thank M. Sarngadharan, M. Guroff, and their colleagues and collaborators for providing serum samples used in this study. Additionally, thanks are due to L. Jagodzinski and R. Liou for assistance with DNA sequencing, and M. B. Eiden, C. Guo, and S. F. Josephs for useful discussions.

1. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rosenbaum, W. & Montagnier, L. (1983) *Science* **220**, 868–870.
2. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. (1984) *Science* **224**, 497–500.
3. Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Okeske, J., Safai, B., White, C., Foster, P. & Markham, P. D. (1984) *Science* **224**, 500–503.
4. Hahn, B. H., Shaw, G. M., Arya, S. K., Popovic, M., Gallo, R. C. & Wong-Staal, F. (1984) *Nature (London)* **313**, 166–168.
5. Sanchez, R., Power, M. D., Barr, P. J., Steimer, K. S., Stempien, M. M., Brown-Shimer, S. L., Gee, W. W., Renard, A., Randolph, A., Levy, J. A., Dina, D. & Luciw, P. A. (1985) *Science* **227**, 484–492.
6. 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.
7. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. (1985) *Cell* **40**, 9–17.
8. Muesing, M. A., Smith, D. H., Cabradilla, C. D., Benton, C. V., Lasky, L. A. & Capon, D. J. (1985) *Nature (London)* **313**, 450–458.
9. Sodroski, J., Rosen, C., Wong-Staal, F., Salahuddin, S. Z., Popovic, M., Arya, S. K., Gallo, R. C. & Haseltine, W. A. (1985) *Science* **227**, 171–173.
10. Arya, S. K., Guo, C., Josephs, S. F. & Wong-Staal, F. (1985) *Science* **229**, 69–73.
11. Arya, S. K., Gallo, R. C., Hahn, B. H., Shaw, G. M., Popovic, M., Salahuddin, S. Z. & Wong-Staal, F. (1984) *Science* **225**, 927–930.
12. Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F. & Haseltine, W. (1985) *Science* **229**, 74–77.
13. Sodroski, J. G., Rosen, C. A. & Haseltine, W. A. (1984) *Science* **225**, 381–385.
14. Arya, S. K. (1982) *Int. J. Biochem.* **14**, 19–24.
15. Arya, S. K. & Gallo, R. C. (1984) *Biochemistry* **23**, 6690–6696.
16. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
17. Melton, D., Krieg, P., Rebagliati, M., Maniatis, T., Zinn, K. & Green, M. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
18. Kozak, M. (1984) *Nature (London)* **308**, 241–245.
19. Sarngadharan, M. G., Bruch, L., Popovic, M. & Gallo, R. C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3481–3484.
20. Allan, J. S., Colligan, J. E., Barin, F., McLane, M. F., Sodroski, J. G., Rosen, C. A., Haseltine, W. A., Lee, T. H. & Essex, M. (1985) *Science* **228**, 1091–1094.
21. Barin, F., McLane, M. F., Allan, J. S. & Lee, T. H. (1985) *Science* **228**, 1094–1096.
22. Crowl, R., Ganguly, K., Gordon, M., Conroy, R., Schaber, M., Kramer, R., Shaw, G., Wong-Staal, F. & Reddy, E. P. (1985) *Cell* **41**, 979–986.
23. Kyte, J. & Doolittle, R. F. (1981) *J. Mol. Biol.* **157**, 105–132.
24. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13**, 222–228.