Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 5282–5285, July 1986
Medical Sciences

Expression of the protein encoded by the 3′ open reading frame of human T-cell lymphotropic virus type III in bacteria: Demonstration of its immunoreactivity with human sera

G. FRANCHINI*, M. ROBERT-GUROFF*, F. WONG-STAAL*, J. GHRAYEB†, I. KATO†, T. W. CHANG†‡, AND N. T. CHANG†‡

*Laboratory of Tumor Cell Biology, Development Therapeutic Program, Division of Cancer Treatment, National Cancer Institute, Bethesda, MD 20892; and †Centocor, 244 Great Valley Parkway, Malvern, PA 19355

Communicated by Janet D. Rowley, March 10, 1986

ABSTRACT **The genome of human T-cell lymphotropic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV), the infectious agent etiologically associated with the acquired immunodeficiency syndrome, contains, in addition to the genes for the polymerase, core, and envelope proteins, several open reading frames. To investigate whether the 3′ open reading frame (3′*orf*) located between the envelope gene and the 3′ long terminal repeat is a gene expressed *in vivo* in infected individuals, we inserted a fragment of 3′*orf* in a prokaryotic expression vector. The protein product synthesized in bacteria was purified and allowed to react with sera from individuals infected with human T-cell lymphotropic virus type III as indicated by seropositivity for other viral proteins. Two-thirds of the sera, regardless of the clinical status of the individuals, reacted with the purified protein indicating that 3′*orf* is a viral gene the product of which is immunogenic *in vivo*. A polyclonal rabbit antiserum reacting against the 3′*orf* gene product was obtained by serial injection of rabbits with the purified bacterial protein. The antiserum recognized a 27-kDa protein in the human T-cell lymphotropic virus type III-infected lymphocytes.**

The acquired immunodeficiency syndrome (AIDS) and associated diseases (1–3) have been epidemiologically correlated with a retrovirus, human T-cell lymphotropic virus type III (HTLV-III) or lymphadenopathy-associated virus (LAV) (4, 5). The T-cell tropism of HTLV-III (4, 5) and the cytopathic effect of the virus or cloned viral DNA on the OK T4+ lymphocytes *in vitro* (5, 6) suggest that one or more viral gene products may be responsible for the depletion of T cells of the OK T4 phenotype observed in patients with AIDS (1, 2). Computer analysis of the DNA sequence of the HTLV-III genome (7) led to the identification of the *gag*, *pol*, and *env* genes as well as the prediction of at least two additional open reading frames, *sor* and 3′*orf*, located between the *pol* and *env* genes and between the *env* gene and the 3′ long terminal repeat, respectively. To date, many of the proteins encoded by the HTLV-III genome have been identified: the core proteins, p24, p17, p15, and p9, the exterior portion of the envelope protein gp120, and the transmembrane protein gp41 (8, 9).

Another functional gene, *tat* III, has been localized in the HTLV-III genome (10, 11) and has been shown to mediate transcriptional activation of genes linked to the HTLV-III long terminal repeats (10, 11). However, *sor* and 3′*orf* have not been demonstrated to be functional genes. We have approached this problem by expressing a portion of 3′*orf* in bacteria. Using the partially purified protein, we were able to demonstrate that sera of some people infected with HTLV-III are immunoreactive against it, indicating that the 3′*orf* gene product is indeed expressed *in vivo* and is immunogenic.

MATERIALS AND METHODS

Preparation of the 3′orf Recombinant Expression Vector. The HTLV-III DNA clone BH-8 (12) was cleaved with *Xho* I and *Sst* I, and the resulting 679-base-pair *Xho* I–*Sst* I fragment was isolated. The fragment was treated with DNA polymerase (Klenow fragment) in the presence of all four deoxyribonucleotides and ligated in the repligen expression vector (a kind gift from Scott Putney, Repligen Corp., 101 Binney St., Cambridge, MA 02142), which had been previously cleaved with *Sal* I and *Sst* I and treated with DNA polymerase (Klenow fragment). The ligated DNA was transfected in *Escherichia coli*, and a colony, designated E-10 containing the HTLV-III 3′*orf* recombinant plasmid with the HTLV-III sequences in the correct orientation, was selected.

Detection of the Hybrid 3′orf Protein by the Immunoblot Technique. The E10 containing the 3′orf plasmid and the control untransfected bacterial strain were cultured overnight in conditions described (13). The cells were lysed in RIPA buffer (14) and sonicated for 1 min. One one-hundredth of the cell lysate was applied to a 15% acrylamide/bisacrylamide gel containing $NaDodSO_4$ (14). The electrophoresed proteins were electrophoretically transferred to nitrocellulose sheets at 60 V overnight and reacted with human and mouse antibodies according to the described procedure (15). We used a 1:100 dilution of the human sera and a 1:50 dilution of the mouse monoclonal antibodies. The antigen–antibody reaction was detected using ^{125}I-labeled protein A at a concentration of 5×10^5 cpm/ml (14).

The size of the 3′orf protein expressed in bacteria is calculated with respect to the prestained low molecular weight standards obtained from Bethesda Research Laboratories.

Purification of the 3′orf Protein from the E-10 Strain. Several liters of E-10 cells were grown in L broth containing ampicillin at 100 μg/ml for 8 hr (13). Cells were harvested, resuspended in one-hundredth the volume of the original culture in 50 mM Tris·HCl, pH 8, and disrupted by sonication. The E-10 protein, present in the insoluble material obtained after sonication, was solubilized in 50 mM Tris·HCl, pH 8, containing 1 M sodium chloride, 10 mM dithiothreitol, and 8 M urea, was applied to a Sephacryl S-300 column and eluted with the same buffer. Fractions containing the E-10 protein were pooled, dialyzed extensively, and applied to a

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, human T-cell lymphotropic virus type III; AIDS, acquired immunodeficiency syndrome; ARC, AIDS-related complex; HH, healthy homosexuals.

‡Present address: Center for Biotechnology, Baylor College of Medicine, 6501 Fannin, Room C-506, Houston, TX 77030.

column of monoclonal antibody 3H-8-coupled Sepharose 4B. The 3H-8 monoclonal antibody recognizes the 36 amino acids encoded by the repligen expression vector. The specifically bound material was eluted with 2% (vol/vol) formic acid. The E-10 protein still contained 10–20% impurities as judged by $NaDodSO_4$/PAGE.

Detection of the Protein by Immunoprecipitation. Rabbit immune sera reactive with the 3'orf protein were obtained injecting New Zealand White male rabbits with the purified recombinant E-10 protein. E-10 protein (250 μg) in complete Freund's adjuvant was injected subcutaneously in the rabbit's back. Four boosters of E-10 protein (250 μg) in incomplete Freund's adjuvant were given 1 week apart. The animals were boosted again weekly with E-10 protein at 500 μg in incomplete Freund's adjuvant from the 11th to the 14th week. Immune sera were obtained by serial bleedings of the immunized animals.

HTLV-III-infected lymphocytes (1×10^7 cells) were metabolically labeled with 100 μCi of [^{35}S]methionine in methionine-free medium for 8 hr at 37°C (1 Ci = 37 GBq). The cells were harvested, lysed with RIPA buffer (14), and centrifuged 20 min at 10,000 rpm in an Eppendorf centrifuge. Portions of the lysate supernatant were mixed with 50 μl of serum obtained from a rabbit immunized against the purified E-10 protein synthesized in *E. coli* for 1 hr at 4°C. The immunocomplexes, bound to protein-A-Sepharose beads, were analyzed on a 10–20% polyacrylamide/$NaDodSO_4$ gradient gel.

RESULTS

Generation of the Recombinant Expression Vector Containing the 3'*orf* of HTLV-III. We introduced the DNA fragment obtained by an *Xho* I–*Sst* I cleavage of the HTLV-III DNA clone BH-8 (12) into the prokaryotic repligen expression vector (13). *Xho* I cleaves in position 8474 of the HTLV-III nucleotide sequence (7). Therefore, we eliminated 44 amino acids at the amino terminus of the predicted open reading frame. The methionine codon presumably required for initiation of translation of this protein is located 1 amino acid downstream from the termination of the envelope gene open reading frame. Therefore, only 34 amino acids may have been deleted with respect to the native protein. The *Sst* I site is 158 nucleotides downstream from the TGA termination codon, which is located in position 8991. We expected to have a protein product of 172 amino acids derived from the 3'*orf* region. The prokaryotic vector that we used provides 36 amino acids at the amino terminus of the protein, which are derived from the polylinker sequences. Thus, the hybrid protein should contain 208 amino acids.

Detection of the Bacterially Synthesized 3'orf Hybrid Protein. The total lysates of the bacteria containing the recombinant 3'orf, designated E-10, and the parental bacteria control were analyzed by immunoblot (15) using human sera of patients infected by HTLV-III and a mouse monoclonal antibody (3H-8), which recognizes an epitope located in the 36 amino acids encoded by the linkers of the vector (J. Ghrayeb, unpublished data). Several human sera and the 3H-8 monoclonal antibody detected a 27-kDa protein on immunoblots (Fig. 1). The protein was detected in the E-10 lysate but not in the control untransfected bacteria using sera of patients with AIDS, AIDS-related complex (ARC), and clinically healthy homosexuals (HH) (Fig. 1, lanes 1–6). No reaction with the 27-kDa protein was observed when we used a serum from a healthy individual not in a risk group for AIDS (Fig. 1, lanes 7 and 8). The 27-kDa protein was also detected in the E-10 lysate by the mouse monoclonal 3H-8 antibody directed against the vector epitopes (Fig. 1, lanes 9 and 10). All the sera showed reactivity with other *E. coli* proteins as well. The sera of infected patients displayed higher reactivity with *E. coli* protein perhaps indicating a greater exposure to bacterial antigens. Similar results were also obtained using a radioimmunoprecipitation assay (14) of metabolically labeled bacterial proteins (not shown). The apparent size of the protein is again larger than expected and this may reflect an abnormal conformation of the protein in $NaDodSO_4$.

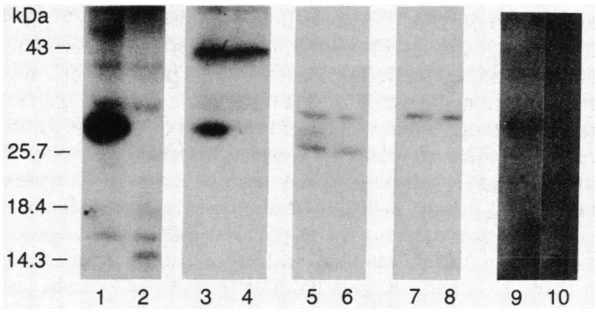

FIG. 1. The autoradiograms shown above were obtained by reacting the *E. coli* lysates with human sera in an immunoblot assay. *E. coli* strain E-10 containing the recombinant 3'orf plasmid (lanes 1, 3, 5, 7, and 9). Untransfected bacteria (lanes 2, 4, 6, 8, and 10). HH sera (lanes 1, 2, 5, and 6). ARC sera (lanes 3 and 4). Sera from healthy individuals not at risk (lanes 7 and 8). The 3H-8 monoclonal antibody (lanes 9 and 10) recognizes the amino terminus of the hybrid protein.

Detection of Natural Antibodies Against the 3'orf Protein in Sera from Patients. The hybrid protein was purified from E-10 bacterial lysate by immunoaffinity chromatography using the mouse monoclonal antibody irreversibly bound to Sepharose beads. The purified protein still reacted with the human sera previously tested as well as the mouse monoclonal antibody proving that it is the hybrid protein that is recognized by the sera rather than two proteins comigrating on $NaDodSO_4$/PAGE (Fig. 2). The purified protein was subsequently used in an immunoblot assay to study the distribution of natural antibodies in sera of patients at different stages of the disease. We analyzed the sera of 82 individuals, including AIDS and ARC patients, HH, intravenous drug users, relatives of AIDS patients (healthy individuals at risk), and

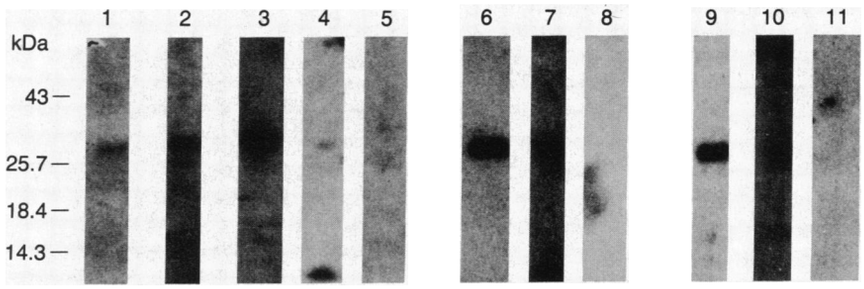

FIG. 2. Three micrograms of purified protein obtained from the E-10 strain were loaded in each lane and incubated with human sera in an immunoblot assay. HH sera (lanes 1, 9–11). 3H-8 monoclonal antibody (lane 2). AIDS sera (lanes 3–5). ARC sera (lanes 6–8).

sera of people not included in populations at risk (2). An example is shown in Fig. 2. The sera of some patients recognized the purified protein whereas healthy persons not at risk were negative. The serum of some patients was more strongly reactive and some contained no detectable antibodies. Of the 82 sera tested 79 were seropositive for other viral proteins (gp41 and p24); the 3 seronegative cases represent the normal donors without known risk factors. All the sera were then tested against the purified 3'orf protein, and the presence of antibodies in the different groups is summarized in Table 1. The three normal donors not at risk were negative, whereas approximately two-thirds of the sera from AIDS and ARC patients and healthy people at risk reacted with the purified protein. The data indicate that the 3'*orf* of HTLV-III is indeed a gene, the protein product of which is immunogenic in humans.

Identification of the Native 3'orf Protein in HTLV-III-Infected T Cells. Rabbits were immunized with the purified protein in an attempt to obtain polyvalent antibodies reactive with the native 3'orf protein. The animals were bled, and the antisera were used to analyze the native protein in the HTLV-III-infected lymphocytes. A radioimmunoprecipitation assay was performed on the metabolically labeled HTLV-III infected cell lines H9 (16). The rabbit antiserum recognized a protein of ≈27 kDa in the infected lymphocytes, whereas the serum obtained from the animal prior to immunization with the purified E-10 protein failed to detect the 27-kDa protein (Fig 3, lanes A and B). Background bands, probably due to a specific binding of cellular proteins to the protein-A-Sepharose were also detected in lanes A and B. A 27-kDa protein, in addition to a high molecular weight band probably representing an envelope precursor protein (lane C), was also immunoprecipitated by serum from a HH infected by HTLV-III. A human serum obtained from a healthy uninfected donor did not recognize the 27-kDa 3'orf protein in the H9 HTLV-III-infected cells (lane D). Thus, the 3'orf protein product is expressed in the infected T cells.

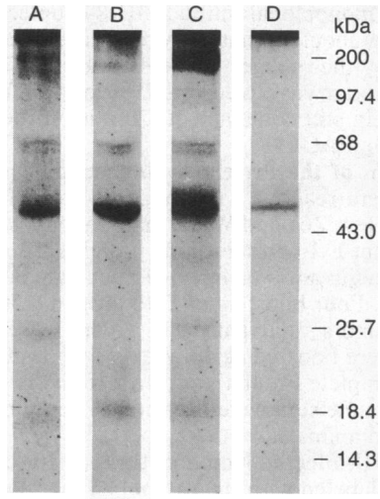

FIG. 3. Autoradiogram of the immunoprecipitation of the 3'orf protein from the cell lysates of H9 HTLV-III-infected cells. Aliquots of 100 μl of labeled cellular proteins were immunoprecipitated with 50 μl of rabbit antiserum (lane A), 50 μl of preimmune rabbit serum (lane B), 10 μl of human serum from a healthy HTLV-III-infected patient (lane C), and 10 μl of serum from an uninfected donor (lane D). The arrow indicates the position of the protein recognized by the antiserum (lane A) as well as the human serum (lane C) in the HTLV-III-infected cells. The size corresponds to approximately 27 kDa (arrow).

DISCUSSION

Our data indicate that the 3'*orf* of HTLV-III is indeed a gene, the protein product of which is immunogenic in humans. Allan *et al.* (17) reported a lower incidence of natural antibodies directed against the p27 3'*orf* product in patients infected by HTLV-III. We believe that the difference may be explained by the fact that the use of purified protein in our assay allowed us to score as positive sera containing low titres of antibodies. The fact that the protein elicits a detectable immunoresponse in only two-thirds of the infected people suggests the following possible interpretations: (*i*) the 3'orf protein may be less immunogenic than the gp41 envelope protein in humans; (*ii*) the rate of cell death may vary in different individuals with consequent variable exposure of viral encoded antigens; (*iii*) the 3'orf protein may be relegated to the cellular compartment rather than to the membrane; (*iv*) it may reflect a different degree of failure of the immune system in different patients; and (*v*) the rate of viral replication in different individuals in combination with one or more of the above mentioned factors may play an important role. The availability of the antiserum against the 3'orf protein will allow the identification of the subcellular localization of the protein and give some insights on its function in the infected cells. The 3'*orf* gene transcript in the infected cells is a doubly spliced mRNA of 2 kilobases that contains only one open reading frame (10). The cDNA of this transcript does not exhibit a transcriptional activation function like the cDNA of the *tat* III gene (10). The fact that the 3'*orf* is expressed *in vivo* suggests that it may be important either for viral replicative functions or for a biological effect exerted by the virus on the infected T cells. Of interest is the finding that the p27 3'orf protein of HTLV-III is also recognized by sera of monkeys affected by a disease comparable to human AIDS (17). These monkeys have been found to be infected by a retrovirus, simian T-cell lymphotropic virus type III (18), serologically related to HTLV-III. The molecular characterization of the equivalent gene in the simian T-cell lymphotropic virus type III genome and the possibility of having an animal model to study the function of this gene will facilitate the study of the role of the 3'*orf* gene product.

Table 1. Natural antibodies against the 3'*orf* gene product in infected patients

Clinical status	Tested, no.	Positive, no.	% positive
Healthy donors (not at risk)	3	0	0
Healthy individuals at risk (homosexuals, i.v. drug users, relatives of AIDS patients)	29*	19	65
ARC patients	22*	15	68
AIDS patients	28*	17	60

*All these sera were also positive for antibodies to other viral antigens (either p24 and/or gp41).

We thank Dr. R. C. Gallo for support; Drs. M. Eiden and M. Feinberg for critical review of the manuscript; and Drs. T. H. Lee, J. S. Allan, and M. Essex for providing us with an HTLV-III-infected cell line expressing high level of the 3'orf protein and for sharing results prior to publication.

1. Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D. & Farr, P. T. (1981) *N. Engl. J. Med.* **305**, 1425–1430.
2. Goedert, J. J. & Blattner, W. A. (1985) in *AIDS, Etiology, Diagnosis, Treatment and Prevention*, eds. De Vita, V., Hellman, S. & Rosemberg, S. A. (Lippincott, Philadelphia), Vol. 1, pp. 1–30.
3. Killen, J. (1984) *AIDS Memo.* **1**, 4.
4. 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., Kogenbaum, J. V. & Montagnier, L. (1983) *Science* **220**, 868–870.

5. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. (1984) *Science* **224,** 497–500.
6. Fisher, A. G., Collalti, E., Ratner, L., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **316,** 262–265.
7. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Joseph, S. F., Doran, E. R., Rafalski, J. A., Whitehozn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautemberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C. & Wong-Staal, F. (1985) *Nature (London)* **313,** 277–282.
8. Sarngadharan, M. G., Popovic, M., Bruch, L., Schupach, J. & Gallo, R. C. (1984) *Science* **224,** 506–509.
9. diMarzo Veronese, F., DeVico, A. L., Copeland, T. D., Droszlan, R., Gallo, R. C. & Sarngadharan, M. G. (1985) *Science* **229,** 1402–1405.
10. Arya, S. K., Chan, G., Josephs, S. & Wong-Staal, F. (1985) *Science* **229,** 69–74.
11. Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F. & Haseltine, W. A. (1985) *Science* **229,** 74–77.
12. Hahn, B. H., Show, G. M., Arya, S. K., Popovic, M., Gallo, R. C. & Wong-Staal, F. (1984) *Nature (London)* **312,** 166–169.
13. Ghrayeb, J., Kato, I., McKinney, S., Huang, J. J., Chanda, P. K., Ho, D. D., Sarangadharan, M. G., Chang, T. W. & Chang, N. T. (1986) *DNA* **5,** 93–99.
14. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
15. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4350–4353.
16. Popovic, M., Read-Connole, E. & Gallo, R. C. (1984) *Lancet* **ii,** 1472–1473.
17. Allan, J. S., Coligan, J. E., Lee, T. H., McLane, M. F., Kanki, P. J., Groopman, J. E. & Essex, M. (1985) *Science* **230,** 810–812.
18. Kanki, P. J., Alroy, J. & Essex, M. (1985) *Science* **230,** 951–954.