JOURNAL OF VIROLOGY, June 1990, p. 3074–3077
0022-538X/90/063074-04$02.00/0

Vol. 64, No. 6

Chemical Synthesis of Biologically Active *tat trans*-Activating Protein of Human Immunodeficiency Virus Type 1

RENE CHUN, CHARLES G. GLABE, AND HUNG FAN*

Department of Molecular Biology and Biochemistry and Cancer Research Institute, University of California, Irvine, California 92717

Received 31 October 1989/Accepted 5 February 1990

Full-length (86-residue) polypeptide corresponding to the human immunodeficiency virus type 1 *tat trans*-activating protein was chemically synthesized on a semiautomated apparatus, using an Fmoc amino acid continuous-flow strategy. The bulk material was relatively homogeneous, as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and isoelectric focusing, and it showed *trans*-activating activity when scrape loaded into cells containing a human immunodeficiency virus long terminal repeat-chloramphenicol acetyltransferase reporter plasmid. Reverse-phase high-pressure liquid chromatography yielded a rather broad elution profile, and assays across the column for biological activity indicated a sharper peak. Thus, high-pressure liquid chromatography provided for enrichment of biological activity. Fast atom bombardment-mass spectrometry of tryptic digests of synthetic *tat* identified several of the predicted tryptic peptides, consistent with accurate chemical synthesis.

Human immunodeficiency virus types 1 and 2 (HIV-1 and -2), the etiological agents of acquired immunodeficiency syndrome, are retroviruses of the lentivirus subclass (5). In contrast to most other replication-competent retroviruses, which have only three genes (*gag*, *pol*, and *env*), HIV-1 and -2 contain a number of additional genes, which specify several regulatory proteins (8). Prominent among these regulatory proteins is the *tat* protein, a positive regulator of viral expression (1, 24, 25). Functional *tat* protein is absolutely required for viral infectivity (1, 25), and it is probably required for activation of quiescent virus from latently infected cells as well.

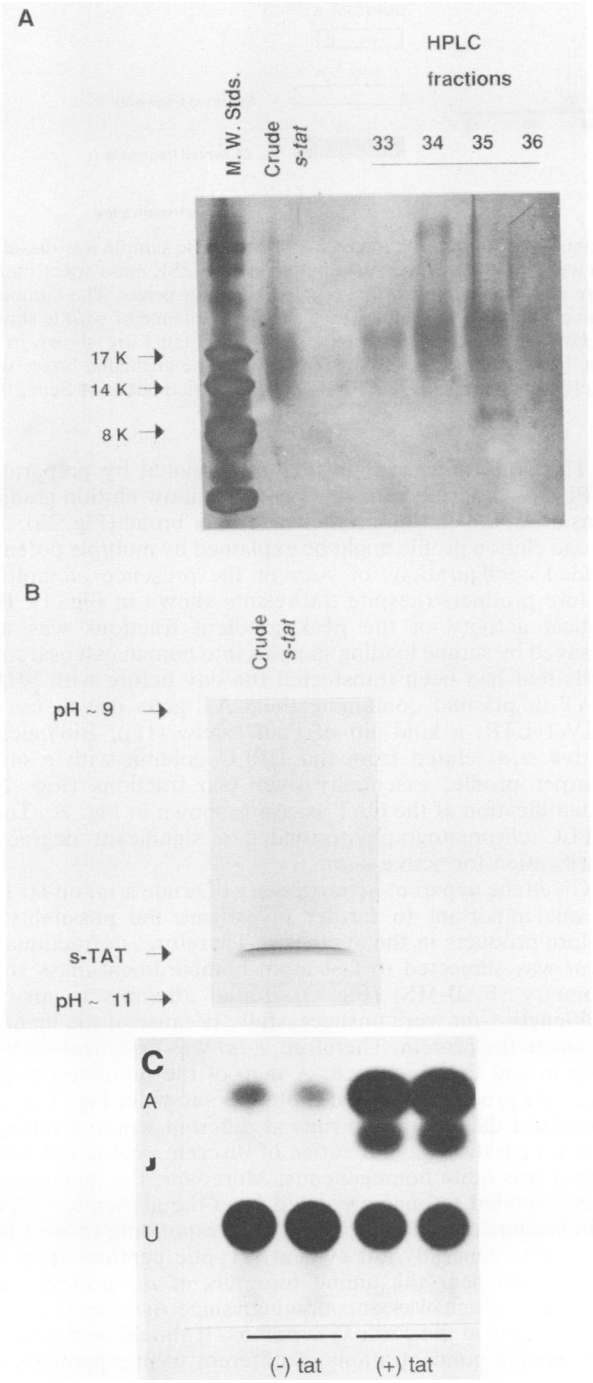
In HIV-1-infected cells, the *tat* protein is expressed from a doubly spliced mRNA (1, 23). It is an 86-amino-acid protein with an apparent size of 14 kilodaltons in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (1, 6, 25). Sequences within the R region of the HIV-1 long terminal repeat (LTR)—the TAR sequences (+1 to +45 base pairs [16, 19, 21])—are the targets for *tat* action; these sequences must be downstream and in the positive orientation to function as TAR sequences (19, 21). The mechanism of *tat* action is still unclear, with various experiments supporting transcription initiation (17, 19, 20), antiattenuation (14), translational efficiency (22), or a combination thereof (3, 25). *tat* protein does not show specific affinity for TAR DNA (6), although recent experiments indicate an interaction with TAR RNA (4).

Further investigations of *tat* action would be facilitated by the availability of large amounts of pure, biologically active *tat* protein. Given its small size, direct chemical synthesis of the protein seemed feasible. Therefore, we synthesized the entire 86-amino-acid *tat* protein from the sequence published for the HXB2 isolate (1, 25; see Fig. 3a). This was carried out on a semiautomated apparatus, using the relatively recent 9-fluorenylmethoxycarbonyl (Fmoc) amino acid continuous-flow strategy (8; C. G. Glabe, submitted for publication). Briefly, 1.1 g of polyamide-Kieselguhr resin containing approximately 0.1 μmol of the C-terminal amino acid (glutamic acid) was washed with dimethyl formamide and packed into a glass column (6.6 by 100 mm). The protein was extended by the sequentional coupling of Fmoc-protected amino acids as BOP-activated Hobt esters (13) with a 30-min acylation time. This methodology offered several advantages. (i) Synthesis on site was considerably more rapid than having peptides made commercially. The entire *tat* protein was synthesized in 27 days (14 working days). (ii) Synthesis and deprotection were conducted under milder conditions than for the traditional *tert*-butyloxycarbonyl (tBOC) methodology. The Fmoc alpha-amino protecting group is cleaved by mild base treatment, avoiding repeated acidolysis of the peptide at each coupling. This was especially important for a large peptide such as the *tat* protein. (iii) It was relatively inexpensive; our cost was about $10 per residue in 0.1-mmol

* Corresponding author.

FIG. 1. Synthesis and activity of s-*tat*. (A) SDS-PAGE of s-*tat*. Crude s-*tat* and pooled HPLC fractions were run on a 10 to 20% acrylamide gel under reducing conditions and then silver stained. Lanes (from left to right): molecular weight standards; crude s-*tat*; HPLC fractions from the column shown in Fig. 2 (the numbers correspond to the fraction numbers in Fig. 2). K, Kilodaltons. (B) Analytical isoelectric focusing of s-*tat* carried out under denaturing conditions. HPLC-purified s-*tat* (10 μg) was loaded on a pH 9 to 11 analytical isoelectric focusing acrylamide gel in 6 M urea–10 mM dithiothreitol and run at constant current for 6 h. The gel was fixed and stained with Coomassie blue. (C) A portion of the s-*tat* preparation that was deblocked and prepared as described in the text. HeLa-U3R-III cells growing in monolayer were seeded at 10^6 cells per 10-cm-diameter dish. The next day, cells were scrape loaded. Briefly, they were washed with phosphate-buffered saline, 1 ml of phosphate-buffered saline containing 20 μg of crude s-*tat* was added, and the cells were scraped from the dish with a rubber policeman. They were then replated in 5 ml of medium containing 10% calf serum and allowed to reattach. The next day, the dead cells were removed by aspiration of the medium, and the monolayer was washed and harvested for CAT enzyme assays as described previously (15). The figure shows duplicate assays for cells scrape loaded with s-*tat* or mock scrape-loaded cells. CAT assays on equal amounts of cell extract were carried out, and acetylated chloramphenicol (A) was separated from unacetylated chloramphenicol (U) by ascending chromatography on silica gel plates. An autoradiogram is shown. Marked *trans* activation of CAT activity resulted from scrape loading of s-*tat*.

quantities (compared with $50 from commercial sources). To maximize coupling, the extent of reaction at each cycle was monitored by testing the synthesis column colorimetrically for the presence of unreacted alpha-amino groups, using trinitrobenzene sulfonic acid (except for proline [12]), and the coupling process was repeated if necessary. The following residue was added only after reaction at the previous cycle was judged to be ≥99% complete. The monitoring was probably important for successful synthesis of a protein as long as the *tat* protein. Repeated coupling for several of the residues (His-22, Arg-30, Tyr-40, Ile-42, Val-51, Thr-64, and Trp-76) was necessary to obtain complete amino acylation.



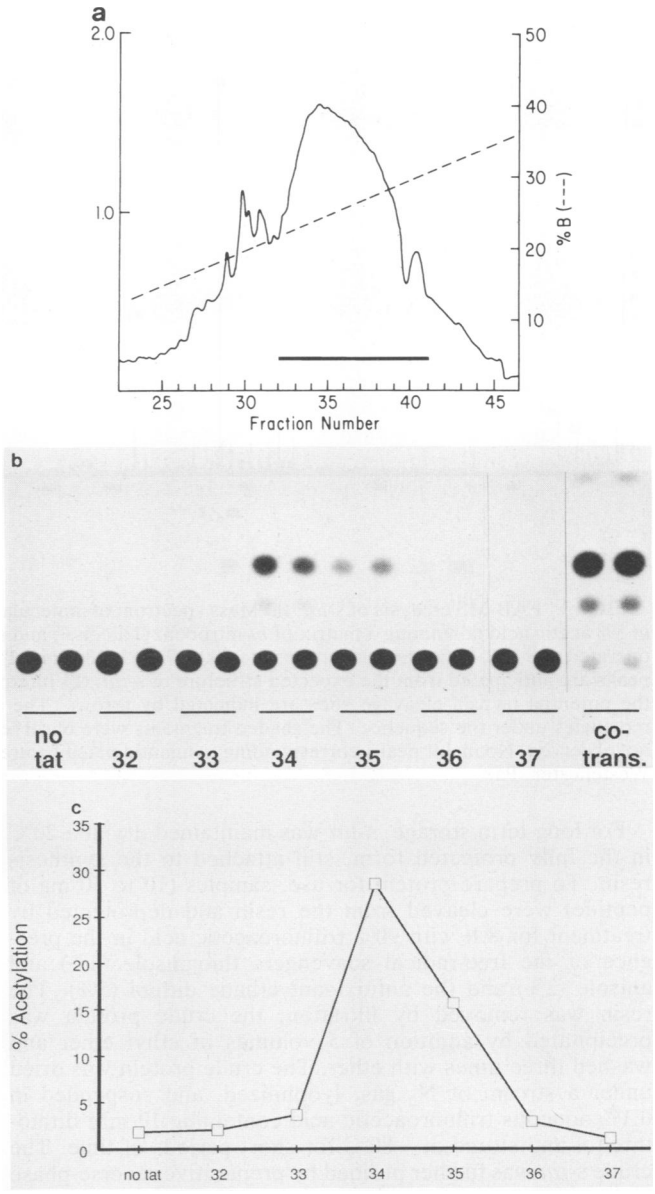


FIG. 2. HPLC of s-*tat*. (a) HPLC of 5 mg of s-*tat*. Chromatography was on a 10 to 50% acetonitrile gradient over 50 min; 1-min fractions were collected. An A_{214} profile of the peptide-containing region is shown. (b) In panel b, fractions 32 to 41 from the HPLC column were recovered by lyophilization and resuspension in 1 ml of 10 mM Tris (pH 7.4). Samples (20 μl) were scrape loaded as for Fig. 1 into human osteosarcoma cells that had been transfected (2) the previous day with 3 μg of plasmid pHIV-CAT. Incubation was continued overnight, and CAT enzyme activity was then measured. All samples were tested in duplicate. Assays of fractions 32 to 37 (corresponding to the bar in panel a) are shown. Controls included no addition of s-*tat* and cotransfection of pHIV-CAT with a *tat* expression plasmid (SV-tat; each plasmid at 3 μg). Total amounts of s-*tat* scrape loaded ranged from ~5 μg (fraction 32) to ~10 to 13 μg (fractions 33 to 36). There was also no activity in fractions 31 or 38 to 41. (c) Quantification of the CAT assays of panel b. Values represent averages for the duplicate samples.

Automated synthesis without monitoring coupling efficiency would certainly have given lower yields of authentic product. The total amount of synthetic *tat* protein (s-*tat*) prepared was approximately 800 mg.

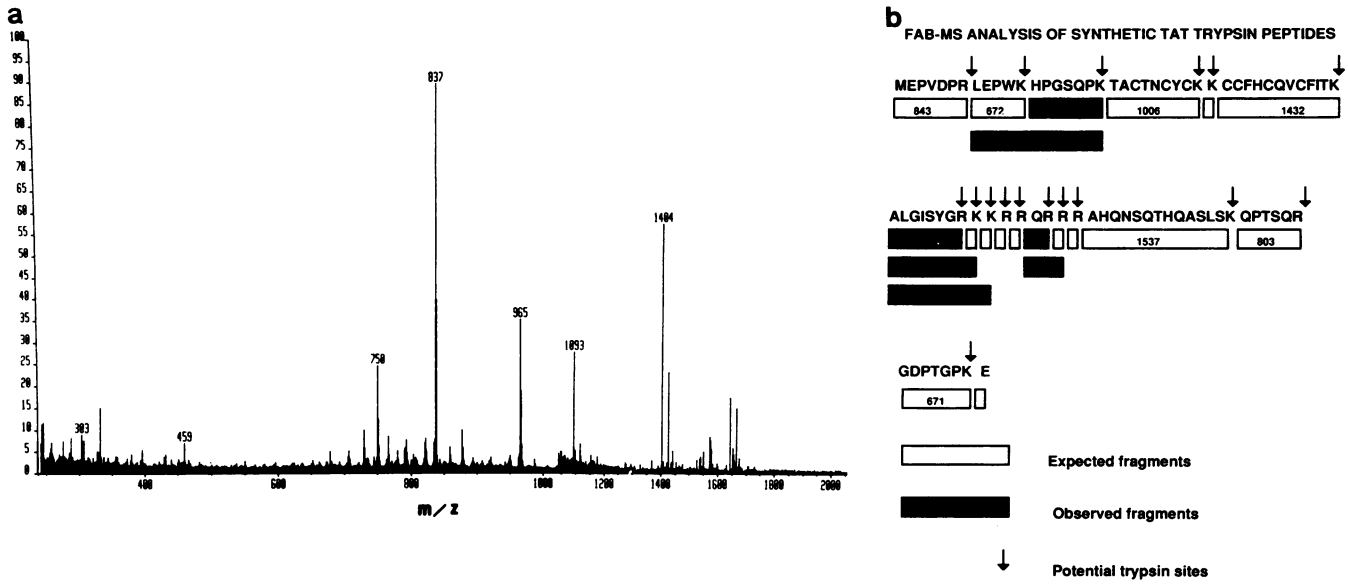


FIG. 3. FAB-MS analysis of s-*tat*. (a) Mass spectrum of molecular ions obtained from tryptic fragments of s-*tat*. The sample was dissolved in 5% acetic acid containing a matrix of *m*-nitrobenzyl alcohol, and spectra were obtained with a VG analytical 2AB-2SE mass spectrometer operated at 8 kV. Major molecular ions at 750, 837, 965, 1093 and 1404 were observed among with a number of minor peaks. The numbered peaks are anticipated from the expected structure of s-*tat*. (b) Interpretation of the FAB-MS data. The expected sequence of s-*tat* is shown; the potential trypsin cleavage sites are indicated by arrows. The molecular masses of the expected trypsin fragments are shown in the rectangles under the sequence. The shaded fragments were observed in the FAB-MS spectra. Molecular ions for the unshaded boxes were not observed. No major peaks corresponding to unanticipated failure products were observed. The FAB-MS was carried out at M-Scan, Inc., Westchester, Pa.

For long-term storage, s-*tat* was maintained dry at −20°C in the fully protected form, still attached to the synthesis resin. To prepare protein for use, samples (10 to 50 mg of peptide) were cleaved from the resin and deprotected by treatment for 8 h with 90% trifluoroacetic acid in the presence of the free-radical scavengers thioanisole (3%) and anisole (2%) and the antioxidant ethane dithiol (5%). The resin was removed by filtration; the crude protein was precipitated by addition of 5 volumes of ethyl ether and washed three times with ether. The crude protein was dried under a stream of N_2 gas, lyophilized, and suspended in 0.1% aqueous trifluoroacetic acid containing 10 mM dithiothreitol and stored at −80°C for short periods of time. The crude s-*tat* was further purified by preparative reverse-phase high-pressure liquid chromatography (HPLC) on a Vydac C4 column (10 by 250 mm), using an acetonitrile gradient (e.g., 10→100%) for elution (see below). The s-*tat* migrated as a relatively homogeneous band on SDS-PAGE, with a molecular mass of approximately 17 kilodaltons (Fig. 1A), consistent with published reports for native *tat* protein (6) and our own immunoprecipitations of native *tat* protein (not shown). Isoelectric focusing gels under denaturing conditions (Fig. 1B) also indicated a relatively homogeneous product, with the predicted isoelectric point of ~pH 10.5.

The biological activity of the s-*tat* was assessed by scrape-loading (7, 9) into HeLa-U3R-III cells. These cells (kindly provided by Craig Rosen) harbor an integrated plasmid containing the bacterial chloramphenicol acetyltransferase (CAT) gene driven by the HIV LTR and respond to biologically active *tat* protein by expression of CAT enzyme. Bulk s-*tat* showed significant biological activity when scrape loaded into these cells at 20 μg/ml (Fig. 1C). Gentz et al. (9) reported somewhat higher *trans* activations when a bacterially synthesized *tat* was scrape loaded at 20 μg/ml into the same cells.

The crude s-*tat* was further fractionated by preparative HPLC as described above. When a shallow elution gradient was used, the elution profile was rather broad (Fig. 2a). The broad elution profile might be explained by multiple potential folded configurations of s-*tat* or the presence of multiple failure products (despite the results shown in Fig. 1). Biological activity of the peak gradient fractions was also assayed by scrape loading samples into human osteosarcoma cells that had been transfected the day before with pHIV-CAT (a plasmid containing the CAT gene driven by the HIV-1 LTR; a kind gift of Paul Luciw [19]). Biologically active s-*tat* eluted from the HPLC column with a much sharper profile, essentially over two fractions (Fig. 2b). Quantification of the CAT assays is shown in Fig. 2c. Thus, HPLC chromatography provided a significant degree of purification for active s-*tat*.

Given the apparent heterogeneity of crude s-*tat* on HPLC, it was important to further investigate the possibility of failure products in the synthesis. Therefore, unfractionated s-*tat* was subjected to fast atom bombardment-mass spectrometry (FAB-MS) (Fig. 3). Initial attempts to analyze full-length s-*tat* were unsuccessfully because of the inability to ionize the protein. Therefore, s-*tat* was first digested with trypsin and then analyzed. A map of the predicted tryptic cleavage products from *tat* protein is shown in Fig. 3, as are combined data from two runs at different window settings. The preparation gave a series of discrete peaks, indicating that it was quite homogeneous. Moreover, the major peaks corresponded to ones predicted from the authentic *tat* protein sequence, and major peaks corresponding to $n-1$ failures were generally not evident. Tryptic peptides from the center and near the amino terminus of *tat* protein were detected, which was encouraging, since the chemical synthesis began at the carboxy terminus. It should be noted that the relative concentrations of different tryptic peptides de-

tected will be influenced by the efficiency of trypsin cleavage at each site and also be the relative ease with which the different peptides are ionized during the mass spectrometry. Thus, the failure to identify all predicted tryptic peptides was not surprising.

In this report, we describe preparation of biologically active synthetic *tat* protein for HIV-1. As indicated above, new synthetic chemistries facilitated this project. The availability of large (milligram) quantities of s-*tat* offer several potential advantages for studies of *tat* action. In particular, even if s-*tat* contains some synthetic failure products, it is not contaminated with other proteins of eucaryotic or bacterial origin. Thus, it should be very useful in generation of affinity columns, raising of antibodies, and cross-linking experiments. Ultimately, s-*tat* enriched for biological activity may be useful in structure or structure-function studies.

While this work was in progress, Green and co-workers (10, 11) reported synthesis of biologically active *tat* protein, using tBOC synthetic chemistries and an automated synthesizer. Thus, two different chemical synthesis approaches have yielded biologically active *tat* protein.

This work was supported by a grant from the University of California Universitywide Taskforce on AIDS. R.C. was supported by a predoctoral fellowship from the UC Irvine Corporate-University Partners Program.

LITERATURE CITED

1. **Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal.** 1985. *Trans*-activator gene of human T-lymphotropic virus type III (HTLV-III). Science **229**:69–73.
2. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**:2745–2752.
3. **Cullen, B. R.** 1986. *Trans*-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell **46**:973–982.
4. **Dingwall, C., I. Ernberg, M. J. Gait, S. M. Green, S. Heaphy, J. Karn, A. D. Lowe, M. Singh, M. A. Skinner, and R. Valerio.** 1989. Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA *in vitro*. Proc. Natl. Acad. Sci. USA **86**:6925–6929.
5. **Fauci, A. S.** 1988. The human immunodeficiency virus: infectivity and mechanisms of pathogenesis. Science **239**:617–622.
6. **Frankel, A. D., D. S. Bredt, and C. O. Pabo.** 1988. Tat protein from human immunodeficiency virus forms a metal-linked dimer. Science **240**:70–73.
7. **Frankel, A. D., and C. O. Pabo.** 1988. Cellular uptake of the Tat protein from human immunodeficiency virus. Cell **55**:1189–1193.
8. **Gallo, R., F. Wong-Staal, L. Montagnier, W. A. Haseltine, and M. Yoshida.** 1988. HIV/HTLV gene nomenclature. Nature (London) **333**:504.
9. **Gentz, R., C.-H. Chen, and C. A. Rosen.** 1989. Bioassay for trans-activation using purified human immunodeficiency virus *tat*-encoded protein: transactivation requires mRNA synthesis. Proc. Natl. Acad. Sci. USA **86**:821–824.
10. **Green, M., M. Ishino, and P. M. Loewenstein.** 1989. Mutational analysis of HIV-1 tat minimal domain peptides: identification of *trans*-dominant mutants that suppress HIV-LTR-driven gene expression. Cell **58**:215–223.
11. **Green, M., and P. M. Loewenstein.** 1988. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat transactivator protein. Cell **55**:1179–1188.
12. **Hancock, W. S., and J. E. Battersby.** 1976. A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid. Anal. Biochem. **71**:260–264.
13. **Hudson, D.** 1988. Methodological implications of simultaneous solid-phase peptide synthesis. 1. Comparison of different coupling procedures. J. Org. Chem. **53**:617–624.
14. **Kao, S.-Y., A. F. Colman, P. A. Luciw, and B. M. Peterlin.** 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by Tat gene product. Nature (London) **330**:489–493.
15. **Kitado, H., I. S. Y. Chen, N. P. Shah, A. J. Cann, K. Shimotohno, and H. Fan.** 1987. U3 sequences from HTLV-I and -II LTRs confer p*x* protein response to a murine leukemia virus LTR. Science **235**:901–904.
16. **Muesing, M. A., D. H. Smith, and D. J. Capon.** 1987. Regulation of mRNA accumulation by a human immunodeficiency virus *trans*-activator protein. Cell **48**:691–701.
17. **Okomoto, T., and F. Wong-Staal.** 1986. Demonstration of virus-specific transcriptional activator(s) in cells infected with HTLV-III by an in vitro cell-free system. Cell **47**:29–35.
18. **Paivinen, A., P. Coassin, P. Reiss, L. Udell, and R. Karol.** 1987. Solid phase peptide synthesis by the Fmoc-polyamide continuous flow method. J. Anal. Purification **2**:36–46.
19. **Peterlin, B. M., P. A. Luciw, P. J. Barr, and M. D. Walker.** 1986. Elevated levels of mRNA can account for the *trans*-activation of human immunodeficiency virus (HIV). Proc. Natl. Acad. Sci. USA **83**:9734–9738.
20. **Rice, A. R., and M. B. Mathews.** 1988. Transcriptional but not translation regulation of HIV-1 by the Tat gene product. Nature (London) **332**:551–553.
21. **Rosen, C. A., J. G. Sodroski, F. Campbell, and W. A. Haseltine.** 1985. The location of *cis*-acting regulatory sequences in the human T cell lymphotrophic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41**:813–823.
22. **Rosen, C. A., J. G. Sodroski, W. C. Goh, A. I. Dayton, J. Lippke, and W. A. Haseltine.** 1986. Post-transcriptional regulation accounts for the *trans*-activation of the human T-lymphotrophic virus type III. Nature (London) **319**:555–559.
23. **Sodroski, J., R. Patarca, C. Rosen, F. Wong-Staal, and W. Haseltine.** 1985. Location of the *trans*-activating region on the genome of human T-cell lymphotrophic virus type III. Science **229**:74–77.
24. **Sodroski, J. G., C. A. Rosen, and W. A. Haseltine.** 1985. *Trans*-acting transcriptional regulation of human T-cell leukemia virus type III long terminal repeat. Science **227**:171–173.
25. **Wright, C. M., B. K. Felber, H. Paskalis, and G. N. Pavlakis.** 1986. Expression and characterization of the *trans*-activator of HTLV-III/LAV virus. Science **234**:988–992.