Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 6543–6546, September 1988
Microbiology

95- and 25-kDa fragments of the human immunodeficiency virus envelope glycoprotein gp120 bind to the CD4 receptor

(identification of cleavage sites/possible domain arrangements/sequence analysis with ^{125}I-labeled tyrosine)

ANDERS NYGREN*†, TOMAS BERGMAN‡, THOMAS MATTHEWS§, HANS JÖRNVALL‡, AND HANS WIGZELL*

Departments of *Immunology and ‡Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden; and §Department of Surgery, Duke University, Medical Center, Durham, NC 27710

Communicated by Sune Bergström, March 30, 1988

ABSTRACT **^{125}I-labeled gp120 (120-kDa envelope glycoprotein) from the BH10 isolate of human immunodeficiency virus is cleaved to a limited extent with the glutamate-specific protease from *Staphylococcus aureus*. After disulfide bond reduction, fragments with approximate molecular masses of 95, 60, 50, and 25 kDa are produced. Tests for binding to CD4-positive cells show that only two fragments, the 95- and 25-kDa peptides, are observed in cleavage products that retain the selective binding capacity of gp120. Radiosequence analysis of the fragments after sodium dodecyl sulfate/polyacrylamide gel electrophoresis and electroblotting demonstrates that the 95-kDa fragment lacks the N-terminal region of gp120 and starts at position 143 of the mature envelope protein. The 50-kDa fragment starts at the same position. The 25-kDa binding fragment was similarly deduced to be generated as a small fragment from a cleavage site in the C-terminal part of gp120. The identifications of these fragments demonstrate that radiosequence analysis utilizing ^{125}I-labeled tyrosine residues can function as a useful and reliable method for small-scale determination of cleavage sites in proteins. Combined, the data suggest domain-like subdivisions of gp120, define at least two intervening segments especially sensitive to proteolytic cleavage, and demonstrate the presence of a functional region for receptor binding in the C-terminal part of the molecule.**

The human immunodeficiency virus (HIV), the etiological agent for acquired immunodeficiency syndrome (AIDS), shows specific tropism for the CD4 molecules present on the surface of certain human T lymphocytes and monocytes (1–5). Extensive genomic variation (6–8) is characteristic for HIV, but CD4 tropism is a constant feature of the virus, suggesting that the responsible region of the HIV envelope protein, gp120 (120-kDa glycoprotein), may be of interest as an epitope for a vaccine candidate. Different studies have tried to map this region by recombinant DNA techniques utilizing site-directed mutagenesis and monoclonal antibodies for blocking the interaction between gp120 of HIV and cellular CD4 (9, 10). The available data indicate that the CD4-binding region extends over a sizeable area in the C-terminal part of gp120. However, they only demonstrate that this region is important in creating a CD4-binding capacity of gp120 and do not give proof for the localization of the CD4-binding region. Difficulties in demonstrating direct binding, using synthetic peptides or fragments of gp120 from prokaryotic recombinant DNA systems, are not surprising, since there is a requirement for N-linked carbohydrates on gp120 to maintain the CD4-binding property (11).

In the present study we report that it is possible, starting from native gp120 molecules, to generate large proteolytic fragments with retained CD4-binding ability and to identify them by utilizing ^{125}I-labeled tyrosine as a marker in radiosequence analysis. The fragments isolated by NaDodSO$_4$/polyacrylamide gel electrophoresis were electroblotted onto glass-fiber filter discs for analysis (12). The CD4-binding region of gp120 could then be localized to a 25-kDa segment. Our study also provides strong evidence for the usefulness of ^{125}I-labeled tyrosine in small-scale structural analysis of proteins.

MATERIALS AND METHODS

gp120 was affinity purified from HIV (BH10 isolate)-infected (13) H9 cells (14) by using a detergent extraction previously described (15). Radio-iodination was performed by using the Enzymobead reagent (Bio-Rad) according to the manufacturer's protocol; gp120 (2 μg) dissolved in Dulbecco's phosphate-buffered saline (PBS; 25 μl) was mixed with Enzymobeads (50 μl), PBS, pH 7.2 (50 μl), and 10 μl (1 mCi; 1 Ci = 37 GBq) of Na^{125}I (Amersham). Finally, 25 μl of a 2% aqueous solution of α-D-glucose (allowed to mutarotate to 1% β-D-glucose overnight) was added. The mixture was incubated at room temperature for 20 min, and then the labeled protein was separated from free ^{125}I on Sephadex G-25 (Pharmacia).

^{125}I-labeled gp120 (^{125}I-gp120; 5 × 10^6 cpm) was mixed with 2 μl of glutamate-specific endoproteinase from *Staphylococcus aureus* V8 (5 μg/μl) and incubated for at least 15 min at 37°C in 0.1 M ammonium bicarbonate/0.25% bovine serum albumin (50 μl), pH 7.8.

To test cell binding, 10^5 cpm of proteolytically cleaved gp120 was mixed with 2 × 10^6 Jurkat cells and incubated for 1 hr as described (16). The cells were then washed three times with PBS/0.5% bovine serum albumin and once in PBS only. The cells were harvested in water and sample buffer was added (10% glycerol, 5% mercaptoethanol, 2.3% NaDodSO$_4$ by weight, in 0.0625 M Tris·HCl, pH 6.8), and the lysate was subjected to NaDodSO$_4$/polyacrylamide gradient (10–20%) gel electrophoresis (17). The gels were dried and autoradiographed for 48 hr with Kodak XAR-5 film.

To identify CD4-binding fragments, 5 × 10^6 cpm of proteolytically cleaved gp120 was fractionated by NaDodSO$_4$/polyacrylamide gel electrophoresis. The gel was cut corresponding to the electrophoretic bands, and the accuracy of excision was checked by autoradiography. The gel pieces were electroblotted in an ISCO model 1750 electrophoretic concentrator, transferring fragments onto glass-fiber filter discs (12 mm diameter) as described (12). After electroblotting, apomyoglobin (Beckman) was added to each filter disc, as carrier and standard, and amino acid sequence analysis was performed in an Applied Biosystems 470A gas-phase sequencer. Myoglobin phenylthiohydantoin derivatives were

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.

Abbreviation: HIV, human immunodeficiency virus.

†To whom reprint requests should be addressed at: Department of Immunology, Karolinska Institutet, Box 60400, S-104 01 Stockholm, Sweden.

identified by reverse-phase high-performance liquid chromatography (18) of aliquots; remaining parts were analyzed for ^{125}I radioactivity in a γ counter (Compu Gamma 1282, LKB).

The electrophoretically separated 95-kDa fragment was also electroeluted into a small volume of transfer buffer in the same blotting system but excluding the filter disc. This preparation was then digested with trypsin (10 μg/22,000 cpm) for 4 hr at 37°C in 0.1 M ammonium bicarbonate/0.25% bovine serum albumin, pH 7.8. The tryptic digest was degraded directly in a Beckman 890D liquid-phase radiosequence analysis, after application into glycine-precycled Polybrene (19).

RESULTS

Generation of gp120 Fragments with CD4-Binding Capacity. The ability to generate gp120 fragments of reasonable size and stability was first studied by using different proteolytic enzymes and periods of digestion. Trypsin and several other commonly used proteases failed to give suitable fragments. However, the glutamate-specific extracellular protease from *S. aureus* was found to yield fragments of adequate size, semistable to further degradation. Fig. 1 displays the fragments generated from ^{125}I-gp120 by using this enzyme for different times under conditions of limited proteolysis. Three fragments with sharply defined borders and approximate molecular sizes of 95, 60, and 50 kDa were generated after disulfide reduction, as well as a broader band centering at 25 kDa in size.

We next studied whether any of these fragments had retained the CD4-binding properties of gp120. With CD4-positive cells and ^{125}I-labeled fragments obtained after staphylococcal protease digestion (Fig. 2), two of the four fragments (the 95- and 25-kDa peptides) were observed in products displaying CD4 binding like the original gp120 protein. The specificity of binding was assessed by several means, one of which, the specific blockage of binding by the T4.2 anti-CD4 monoclonal antibody (16), is included in Fig. 2. Monoclonal antibodies directed against other cell-surface molecules failed to block the binding of the fragments and none of the fragments did bind to CD4-negative cells (data not shown).

We can conclude that it is possible to produce from gp120 proteolytic fragments (95 and 25 kDa after reduction) with retention of select binding to CD4.

Identification of Cleavage Sites and Position of Fragments. Having established that the CD4 binding ability can be expressed by discrete proteolytic fragments of gp120, we explored the possibility of identifying them by direct radiosequence analysis. The fragments obtained by limited proteolysis of ^{125}I-gp120 with the staphylococcal protease were separated by NaDodSO$_4$/polyacrylamide gel electrophoresis, electroblotted onto glass-fiber filter discs, and subjected to radiosequence analysis. A comparison of potential cleavage sites for the protease with the positions of ^{125}I-labeled tyrosine residues in the fragments analyzed, utilizing the known genomic sequence of the HIV BH10 isolate (13), identified the cleavage sites. Positions given below indicate the amino acid residues in native gp120 lacking signal peptide (20).

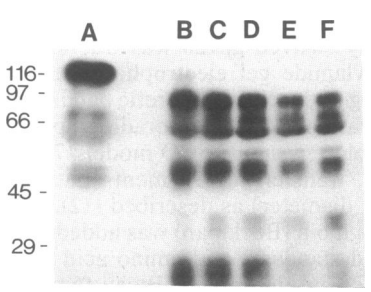

FIG. 1. NaDodSO$_4$/polyacrylamide gel electrophoresis of gp120 fragments generated by glutamate-specific staphylococcal protease digestion. Autoradiography of ^{125}I-labeled molecules. Lane A, gp120, starting material. Lanes B–F, protease-generated fragments obtained after digestion for 15, 25, 35, 45, or 55 min. Molecular masses in kDa of marker proteins are given to the left.

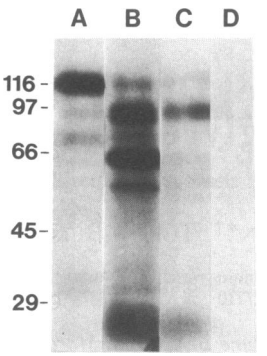

FIG. 2. Generation of CD4-binding fragments from ^{125}I-gp120 by using glutamate-specific staphylococcal protease digestion, as analyzed by NaDodSO$_4$/polyacrylamide gel electrophoresis and autoradiography. Lane A, gp120, starting material. Lane B, protease-generated fragments after 15 min of digestion. Lane C, fragments from the lane B sample allowed to react with CD4-positive (Jurkat) cells, followed by washing, cell lysis, and application of cell-bound material to NaDodSO$_4$/polyacrylamide gel electrophoresis. Lane D, same as for lane C but with the Jurkat cells previously incubated with an anti-CD4 monoclonal antibody, T4.2, at 10 μg/ml. Masses are given as in Fig. 1.

The CD4-binding 95-kDa fragment could be identified as generated from a cleavage after position 142 of gp120. Radiosequence analysis proves positions 1 and 5 to be ^{125}I-labeled tyrosines (Fig. 3 *Left*), and this tyrosine pattern uniquely places the start of the 95-kDa fragment at residue 143. Digestion of the 95-kDa fragment with trypsin followed by direct radiosequence analysis shows position 3 to contain ^{125}I-labeled tyrosine (Fig. 3 *Right*). This proves the presence of a tryptic fragment starting at residue 403 in the 95-kDa digest, further confirming the designated position of the 95-kDa fragment within gp120.

Although neither the 60-kDa nor the 50-kDa fragment is displayed in products with detectable CD4-binding capacity, sequence analysis of the 50-kDa peptide yielded helpful information. The 60-kDa fragment failed to display ^{125}I-labeled tyrosines in the N-terminal region, excluding the possibility of positioning by radiosequence analysis. However, the analysis of the 50-kDa fragment proves that positions 1 and 5 contain ^{125}I-labeled tyrosine residues. It is therefore clear that the 50-kDa fragment is generated by the same N-terminal cleavage site as the 95-kDa fragment. The fact that proteolytic products with the 50-kDa peptide, in contrast to those with the 95-kDa fragment, lack CD4-binding capacity strongly suggests that the CD4-binding region must reside in the C-terminal part of the 95-kDa peptide.

Sequence analysis of the electroblotted 25-kDa fragment failed to produce any ^{125}I-labeled tyrosine residue; 21 cycles analyzed gave no significant liberation of radioactivity. The lack of tyrosine residues localizes the 25-kDa fragment with certain precision. Previous studies using mutational analysis and specific monoclonal antibodies suggest that, at a minimum, the CD4-binding region covers positions 332 to 444 (9, 10). An analysis of a cleavage site on the gp120 that should create a glycosylated fragment of 25 kDa containing this minimal region, but lacking tyrosine within the first 21 residues and fitting with the 95- and 50-kDa fragment data, leaves position 322 as the likely N-terminal starting point of

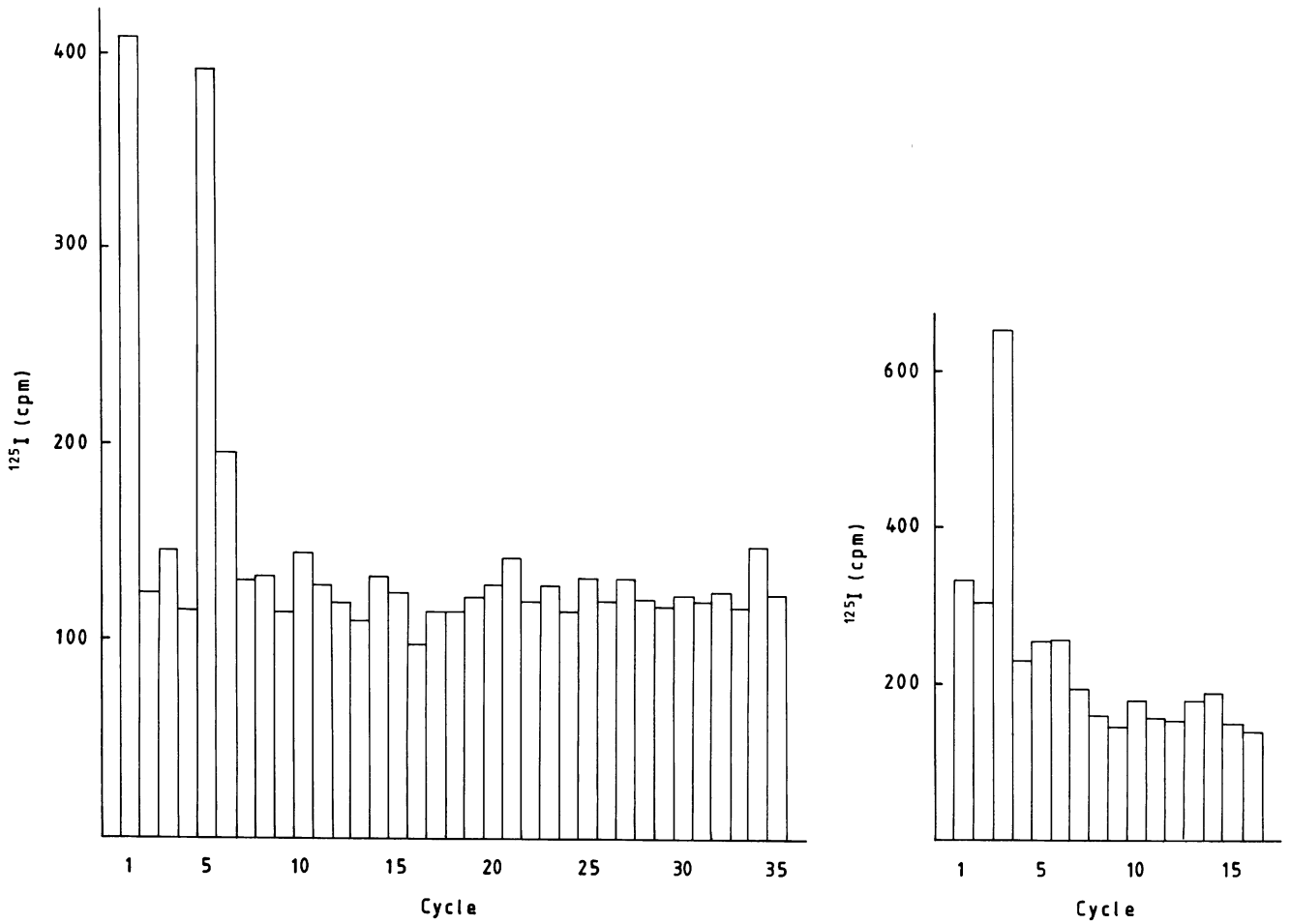

FIG. 3. N-terminal radiosequence analyses. (*Left*) ^{125}I-labeled 95-kDa CD4-binding fragment after NaDodSO$_4$/polyacrylamide gel electrophoresis and electroblotting (the 50-kDa CD4-nonbinding fragment gave identical results). (*Right*) Tryptic digest of the ^{125}I-labeled 95-kDa CD4-binding fragment.

the 25-kDa fragment. In an analysis including the total broad gel band from the 25-kDa region, ^{125}I-labeled tyrosine was recovered at cycles 9 and 10. This pattern fits exactly with the N-terminal segment of gp120, showing that this additional fragment also exists in the 25-kDa fragment area as expected from the cleavage generating the 95-kDa fragment. A provisional map of the 95-, 50-, and 25-kDa fragments within gp120 is given in Fig. 4.

DISCUSSION

The ability of HIV-1 isolates to display select tropism for CD4-positive cells is a stable feature found in all virus strains

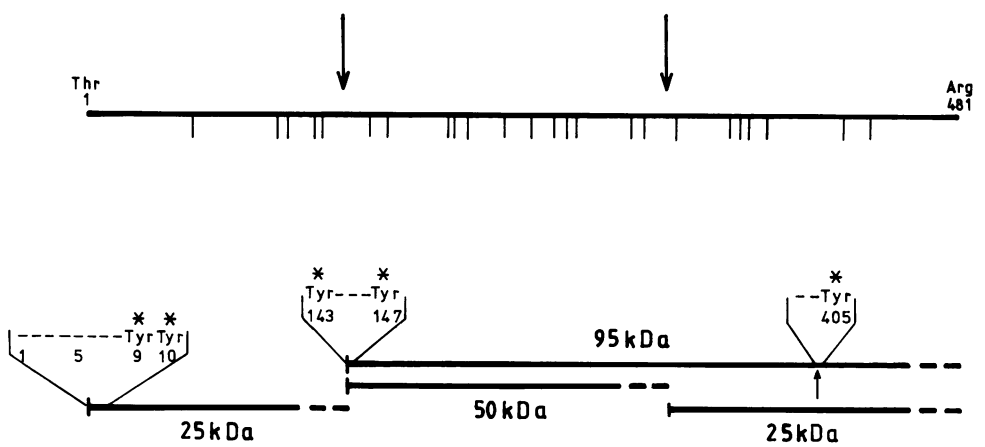

FIG. 4. Schematic representation of the native gp120 molecule (*Upper*), the subdomain fragments (*Lower*), and the relative positions of cleavage sites deduced. The heavy arrows show the glutamate-specific cleavage sites separating subdomains (cf. Fig. 1) corresponding to the approximate extension of the fragments discussed (heavy lines). The thin arrow shows the tryptic cleavage site defined in the 95-kDa fragment. Relevant portions of the gp120 amino acid sequence are indicated with ^{125}I-labeled positions (Fig. 3) identified by asterisks. Also shown are potential N-glycosylation sites (given by vertical bars).

so far analyzed. This requirement for CD4 further extends to the variants of HIV called HIV-2 and to the related simian immunodeficiency virus (SIV) (21). It is also possible to efficiently interfere with HIV infections *in vitro* by blocking the gp120/CD4 interaction (1–5). In all, these data indicate that the CD4 receptor interaction with gp120 plays an essential role for the infectivity of HIV. Attempts with vaccines to produce efficient blocking of gp120 binding to its cellular CD4 receptor have high priority. Human sera from HIV-seropositive individuals can be shown to contain antibodies capable of blocking this interaction in a group-specific manner, but neutralizing titers are normally low (22–24). Generation of isolated regions of gp120 retaining CD4-binding properties may allow a more focused generation of HIV-neutralizing antibodies.

In the present study, we have generated from gp120 proteolytic fragments that retain CD4-binding. In the search for CD4-binding capacity, recombinant peptides produced in prokaryotes as well as synthetic peptides have thus far failed to retain binding (25). In part this may be due to the fact that N-linked carbohydrate chains seem necessary for the capacity of gp120 molecules to bind to CD4 (11). Alternatively, correct disulfide bridges may be necessary. Previous attempts to delineate the CD4-binding region in gp120 by "negative" approaches, that is by inhibition induced by mutations or monoclonal antibodies, indicate that the region involved may extend over a sizeable part of the gp120 molecule in its C-terminal half (9, 10). Our present data provide strong positive evidence for the localization of the CD4-binding region to the C-terminal part of gp120 as well, since the proteolytic products have been identified (Fig. 4) and those with the 50-kDa fragment of the 95-kDa part do not bind. The data suggest domain-like subdivisions of gp120, allowing the generation of fragments with CD4-binding properties. Presumably, the 25-kDa segment from the C- terminal part corresponds to a functionally intact subdomain of the complex glycoprotein. Two further subdomains, although nonfunctional in CD4 binding, are also defined by the second cleavage site, at position 142 (Fig. 4). In all, these studies demonstrate the usefulness of ^{125}I-labeling coupled with $NaDodSO_4$/polyacrylamide gel electrophoresis, electroblotting, and radiosequence analysis for small-scale mapping of functional domains in proteins.

This work was supported by the Swedish Medical Research Council (Grants 03X-3532, B88-16H-07738-03A, and K87-16P-07785-02B), the Swedish Cancer Society (Grant 1806), the U.S. Army Medical Research Acquisition Activity (Contract DAMD17-86-G-6039), the Swedish National Board for Technical Development (Project 87-02744P), and the U.S. National Cancer Institute (Contract FOD-0758).

1. Dalgleish, A. G., Beverley, P. C. L., Clapham, P. R., Crawford, D. H., Greaves, M. F. & Weiss, R. A. (1984) *Nature (London)* **312**, 763–767.
2. Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C. & Montaigner, L. (1984) *Nature (London)* **312**, 767–768.
3. McDougal, J. S., Mawle, A., Cort, S. P., Nicholson, J. K. A., Cross, G. D., Scheppler-Campbell, J. A., Hicks, D. & Sligh, J. (1985) *J. Immunol.* **135**, 3151–3162.
4. McDougal, J. S., Nicholson, J. K. A., Cross, D., Cort, S. P., Kennedy, M. S. & Mawle, A. C. (1986) *J. Immunol.* **137**, 2937–2944.
5. McDougal, J. S., Kennedy, M. S., Sligh, J. M., Cort, S. P., Mawle, A. & Nicholson, J. K. A. (1986) *Science* **231**, 382–385.
6. Wong-Staal, F., Shaw, G. M., Hahn, B. H., Salahuddin, S. Z., Popovic, M., Markham, P., Redfield, R. & Gallo, R. C. (1985) *Nature (London)* **316**, 759–762.
7. Hahn, B. H., Gonda, M. A., Shaw, G. M., Popovic, M., Hoxie, J. A., Gallo, R. C. & Wong-Staal, F. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4813–4817.
8. Starcich, B. R., Hahn, B. H., Shaw, G. M., McNeely, P. D., Modrow, S., Wolf, H., Parks, E. S., Parks, W. P., Josephs, S. F., Gallo, R. C. & Wong-Staal, F. (1986) *Cell* **45**, 637–648.
9. Kowalski, M., Potz, J., Basiripour, L., Dorfman, T., Goh, W. C., Terwilliger, E., Dayton, A., Rosen, C., Haseltine, W. & Sodroski, J. (1987) *Science* **237**, 1351–1355.
10. Lasky, L. A., Nakamura, G., Smith, D. H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory, T. & Capon, D. J. (1987) *Cell* **50**, 975–985.
11. Matthews, T. J., Weinhold, K. J., Lyerly, H. K., Langlois, A. J., Wigzell, H. & Bolognesi, D. P. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5424–5428.
12. Bergman, T. & Jörnvall, H. (1987) *Eur. J. Biochem.* **169**, 9–12.
13. 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., Jr., 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.
14. Popovic, M., Sarngadharan, M. G., Read, E. & Gallo, R. C. (1984) *Science* **224**, 500–503.
15. Robey, W. G., Arthur, L. O., Matthews, T. J., Langlois, A., Copeland, T. D., Lerche, N. W., Oroszlan, S., Bolognesi, D. P., Grilden, R. V. & Fischinger, P. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7023–7027.
16. Lundin, K., Nygren, A., Arthur, L. O., Robey, W. G., Morein, B., Ramstedt, U., Gidlund, M. & Wigzell, H. (1987) *J. Immunol. Methods* **97**, 93–100.
17. O'Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 4007–4021.
18. Kaiser, R., Holmquist, B., Hempel, J., Vallee, B. L. & Jörnvall, H. (1988) *Biochemistry* **27**, 1132–1140.
19. Jörnvall, H. & Philipson, L. (1980) *Eur. J. Biochem.* **104**, 237–247.
20. Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-Staal, F. & Wolf, H. (1987) *J. Virol.* **61**, 570–578.
21. Kanki, P. J., Barin, F., M'Boup, S., Allan, J. S., Romet-Lemonne, J. L., Marlink, R., Mclane, M. F., Lee, T.-H., Arbeille, B., Denis, F. & Essex, M. (1986) *Science* **232**, 238–243.
22. Weiss, R. A., Clapham, P. R., Cheingsong-Popov, R., Dalgleish, A. G., Carne, C. A., Weller, I. V. D. & Tedder, R. S. (1985) *Nature (London)* **316**, 69–72.
23. Robert-Guroff, M., Brown, M. & Gallo, R. C. (1985) *Nature (London)* **316**, 72–74.
24. Lasky, L. A., Groopman, J. E., Fennie, C. W., Benz, P. M., Capon, D. J., Dowbenko, D. J., Nakamura, G. R., Nunes, W. M., Renz, M. E. & Berman, P. W. (1986) *Science* **233**, 209–212.
25. Putney, S. D., Matthews, T. J., Robey, W. G., Lynn, D. L., Robert-Guroff, M., Mueller, W. T., Langlois, A. J., Ghrayeb, J., Petteway, S. R., Jr., Weinhold, K. J., Fischinger, P. J., Wong-Staal, F., Gallo, R. C. & Bolognesi, D. P. (1986) *Science* **234**, 1392–1395.