*Journal of Chromatography*, 476 (1989) 99–112
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 1596

# ISOLATION OF RECOMBINANT PARTIAL *gag* GENE PRODUCT p18 (HIV-$1_{Bru}$) FROM *ESCHERICHIA COLI*

HANNO V. J. KOLBE*, FRANCINE JAEGER, PIERRE LEPAGE, CAROLYN ROITSCH, GEORGES LACAUD, MARIE-PAULE KIENY, JEAN SABATIE, STEPHEN W. BROWN and JEAN-PIERRE LECOCQ

*Transgene SA, 11 Rue de Molsheim, 67000 Strasbourg (France)*

and

MARC GIRARD

*Pasteur Vaccins, 3 Avenue Pasteur, 92430 Marnes-la-Coquette (France)*

## SUMMARY

The membrane-associated structural protein, p18, of the human immunodeficiency virus (HIV-1), has been expressed in *Escherichia coli*. The recombinant protein was purified by cation-exchange chromatography on S Sepharose followed by cation-exchange high-performance liquid chromatography (HPLC) on Sulfoethyl Aspartamide. The isolation of 28.7 mg of recombinant p18 from 16.7 l of cell culture represents an overall yield of *ca.* 20%.

Recombinant p18 was characterized by sodium dodecyl sulphate polyacrylamide gel electrophoresis, reversed-phase HPLC, amino acid composition and amino acid sequence analysis of the N-terminus. Edman degradation of peptides generated by trypsin or *Staphylococcus aureus* V8 proteolytic digestion, including the C-terminus, confirmed the amino acid sequence to be that predicted from the cDNA. A C-terminally cleaved form of recombinant p18, p18LM, was separated in the cation-exchange HPLC step and was partially characterized in parallel with the intact molecule. By Western blotting it was shown that recombinant p18 in addition to the cleaved form p18LM is recognized by a monoclonal antibody which was generated against the natural protein from HIV-1.

## INTRODUCTION

The *gag* gene of the human immodeficiency virus (HIV-1) codes for a precursor protein of 500 amino acids[1]. This precursor is processed by a protease which is coded by a region on the *pol* gene of HIV-1 to generate the structural proteins p18 ($Met_1$–$Tyr_{132}$, calculated molecular mass $M_c$ = 14883 Da; calculated isoelectric point, $pI_c$ = 9.23), p25 ($Pro_{133}$–$Leu_{363}$, $M_c$ = 25580 Da, $pI_c$ = 6.51), and p15 ($Ala_{364}$–$Leu_{500}$, $M_c$ = 15497 Da, $pI_c$ = 9.54). However, during processing of p18 in the mammalian host cell, $Met_1$ is cleaved off and $Gly_2$ is myristylated. Natural p18 thus shows a blocked N-terminus[2]. It is assumed that the myristate moiety anchors p18

0021-9673/89/$03.50 

to the lipid membrane of the virus. This assumption is supported by the results of Gelderblom *et al.*[3], who localized p18 by electron microscopy after immunolabelling on the inner side of the lipid membrane of human immunodeficiency virus particles, showing that p18 forms a spherical shell around the inner core.

Quantitative Western blot analyses of sera from HIV-1-seropositive individuals showed that antibodies against the core proteins p25 and p18 are discovered predominantly before and during the early stages of the disease. The recession of anti-p25 and anti-p18 antibodies during the development of the disease can consequently be used to diagnose the clinical status of the patient[4,5].

Antigens for Western blots are usually prepared by growing HIV-1 in culture, inactivating the virus and separating the viral proteins by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Considering the potentially dangerous manipulation of live virus and the limited availability of these antigens, the advantages of using recombinant protein antigens in Western blot analysis become apparent[6].

Naylor *et al.*[7] have shown that p18 displays a 44% similarity in a stretch of 18 amino acids ($Ile_{92}$–$Asn_{109}$) when compared with the hormone peptide thymosin alpha-1 ($Ile_{11}$–$Asn_{28}$). Sera raised against thymosin alpha-1 recognized p18, isolated from HIV-1, and partially neutralized viral activity, as measured by inhibition of expression of p15, p25 and reversed transcriptase activity in the cell culture medium[8]. These findings make recombinant p18, or peptides thereof, interesting as potential components of an AIDS vaccine.

In this paper, we describe the purification of recombinant p18 from *E. coli*. The protein was characterized biochemically and appears suitable for tests in preliminary experiments regarding its diagnostic and immunological use.

## EXPERIMENTAL

### *Chemicals*

All buffers were prepared with Milli-Q water (Millipore-Waters, Milford, MA, U.S.A.). Cation-exchange high-performance liquid chromatographic (HPLC) buffers were filtered in a 0.45-μm Nalgene unit (Sybron/Nalge, Rochester, NY, U.S.A.). Acetonitrile and 1-propanol were from Carlo Erba (Milan, Italy) and trifluoroacetic acid (TFA) and heptafluorobutyric acid (HFBA) from Pierce (Rockford, IL, U.S.A.). All other chemicals were of analytical-reagent grade.

### *Plasmid construction and fermentation of E. coli*

The complete *gag* gene of HIV-$1_{Bru}$[1] was constructed by juxtaposing the Hind III restriction fragments of plasmids pJ19-1, pJ19-13 and pJ19-17[9]. Using site-directed mutagenesis, a stop codon was introduced at the position of the C-terminal tyrosine of p18. This synthetic gene, coding for $Met_1$ to $Asn_{131}$ (numerical order of amino acids taken from mature *gag* gene product precursor protein[1]) was inserted into a procaryotic expression vector under control of the bacteriophage lambda PL promoter. *E. coli*, strain TGE901, contains a temperature-sensitive cI repressor. *E. coli*, transformed with this construction (pTG-2153-HIV-1), was grown overnight in a 500-ml preculture in Luria–Bertani medium (+ ampicillin) at 30°C. A 20-l LSL-Biolafitte (St. Germain-en-Laye, France) bioreactor was inoculated to an absorbance of 0.1 (650 nm, 1-cm

cuvette). Cells were grown for 2.5 h at 30°C until the cell density reached an absorbance of 0.3. Expression of p18 was then induced by raising the temperature and maintaining it at 42°C. After 7 h, the absorbance had reached 2.4. From the final volume of 16.7 l, cells were harvested by a 10-min centrifugation at 5000 rpm (Sorvall RC-3B, H-6000A rotor; DuPont, Wilmington, DE, U.S.A.).

*S Sepharose chromatography*

The cell pellet (7.87 g of protein) was homogenized (hand-operated glass/glass homogenizer) in 300 ml of phosphate-buffered saline (PBS). Recombinant p18 was released from *E. coli* cells by first freezing the suspension at −80°C and then thawing it at 0–2°C. The supernatant was collected by a 20-min centrifugation at 10 000 rpm and 0–5°C (Sorvall RC-5B, GSA rotor). Broken cells were carried through another two cycles, as described above, and a second and third supernatant were recovered. The pool of the first and second supernatants (540 ml) was diluted with 540 ml of Milli-Q water, adjusted to pH 6.0 with hydrochloric acid and left overnight at 2–4°C. The resulting precipitate was removed by a 15-min centrifugation at 7000 rpm and 0–5°C (Sorvall RC-5B, GSA rotor), At room temperature, protein continued to precipitate. To avoid application of precipitate to the column, the sample was loaded at 200 ml/h from a reservoir, containing a magnetic stirrer, via tubing equipped with a 10-μm solvent filter (Gilson, Middleton, WI, U.S.A.) on to a column (10 cm × 4.4 cm I.D.) of S Sepharose Fast Flow (Pharmacia, Uppsala, Sweden). equilibrated with half-concentrated PBS (pH 6.0). Precipitate concentrated in the sample reservoir and was discarded. The column was washed successively with 50 ml of half-concentrated PBS (pH 6.0) and 200 ml of PBS and protein was eluted with 200 ml each of 40 mM sodium phosphate (pH 7.0) containing 0.2, 0.4, 0.6, 0.8 and 1 *M* sodium chloride. Absorbance was measured at 275 nm with a Model 2158 Uvicord SD instrument (2.5-mm flow cell, 70 μl, LKB, Bromma, Sweden). Eluted fractions were pooled according to the purity of recombinant p18 and concentrated at room temperature in an ultrafiltration cell on a YM5 membrane (Amicon, Danvers, MA, U.S.A.). The concentrated sample was thoroughly washed with CatEx-A buffer (see below).

*Preparative high-performance liquid chromatography*

Preparative HPLC was performed on a Beckman Model 421A gradient liquid chromatograph, equipped with two Model 110A solvent pumps, a Model 340 organizer/sample injector (Beckman Instruments, Palo Alto, CA, U.S.A.), a Uvicord 2158 SD absorbance detector (275 nm, 2.5-mm HPLC flow cell, 8 μl), a Model 2210 recorder and a Model 2112 Redirac fraction collector (LKB). The chromatographic unit consisted of a Sulfoethyl Aspartamide HPLC column (300 Å, 5 μm, 200 mm × 9.4 mm I.D.; Nest Group, Southborough, MA, U.S.A.) and a guard column (20 mm × 2 mm I.D.), dry-filled with Zorbax Diol WR955 (DuPont). Separation of proteins was performed at room temperature. Prior to sample loading, the column was equilibrated with a blank gradient (flow-rate, 1.3 ml/min); 100% CatEx-A buffer [20 m*M* sodium phosphate (pH 7.0)–40 m*M* sodium chloride] from 0 to 5 min, 0–70% CatEx-B buffer [40 m*M* sodium phosphate (pH 7.0)–1 *M* sodium chloride] from 5 to 20 min, 70–100% CatEx-B from 20 to 40 min, 100% CatEx-B from 40 to 45 min, 100% CatEx-B to 100% CatEx-A from 45 to 50 min, 100% CatEx-A from 50 to 70 min. The sample was loaded via pump A at 1.5 ml/min and, after rinsing the system with 20 ml of

CatEx-A, the above-described gradient was started. Fractions were collected automatically every 0.7 min. The salt concentration in the fraction was determined by conductimetry (CDM3 conductivity meter, Radiometer, Copenhagen, Denmark).

*Proteolytic digestion*

*Trypsin digestion.* Recombinant p18 and p18LM (100 μg each) were separately desalted by reversed-phase HPLC in system 1 (see below). The protein peak was collected and solvent was removed in a Speedvac concentrator (Savant Instruments, Framingdale, NY, U.S.A.). Trypsin (5 μg) (sequence grade, Boehringer, Mannheim, F.R.G.) was added in 100 m*M* N-ethylmorpholine (Janssen Chimica, Beerese, Belgium)–0.1 m*M* calcium chloride (pH 8.3) (final volume 400 μl) and digestion was performed for 6 h at 37°C. The reaction was stopped by addition of 2% (final concentration) TFA. Samples were kept at −20°C until analysed by reversed-phase HPLC.

*V8 protease digestion.* To recombinant p18 and p18LM (100 μg each), 20 μg of *S. aureus* V8 protease (Boehringer) were added in 50 m*M* sodium phosphate (pH 7.8) (final volume 500 μl) and digestion was performed for 24 h at 37°C. The reaction was stopped by addition of 4% (final concentration) acetic acid. Samples were kept at −20°C until analysed by reversed-phase HPLC.

*Analytical reversed-phase high-performance liquid chromatography*

Analytical reversed-phase HPLC was performed at room temperature on a Hewlett-Packard ChemStation 1090 M liquid chromatograph (500-μl sample loop) with a photodiode-array detector, connected to a Model 9000/300 computer with monitor, a Model 9122 disk drive unit, a ThinkJet and a ColorPro plotter. For editing of chromatograms Operating Software, Revision 4.05, was used (Hewlett-Packard, Avondale, PA, U.S.A.).

*System 1.* The chromatographic unit consisted of a Nucleosil ($C_4$) HPLC column (4000 Å, 7 μm, 150 mm × 4.6 mm I.D.; Macherey, Nagel & Co., Düren, F.R.G.). Separation of purified proteins was performed using the following gradient (flow-rate, 0.6 ml/min): 90% eluent RP1-A (0.1% TFA in Milli-Q water) at 0 min, 10–90% eluent RP1-B [0.1% TFA in acetonitrile–Milli-Q water (70:30, v/v)] from 0 to 14 min, 90% RP1-B from 14 to 21 min, 90% RP1-B to 90% RP1-A from 21 to 24 min, 90% RP1-A from 24 to 30 min. The absorbance was measured at 205 and 280 nm. Fractions were collected manually. Samples for amino acid sequence analysis were dried in a Speedvac concentrator.

*System 2.* The chromatographic unit consisted of a Vydac 218TP54 ($C_{18}$) HPLC column (300 Å, 5 μm, 4.6 mm × 250 mm I.D.; Separation Group, Hesperia, CA, U.S.A.). Separation of peptides after proteolytic digestion was performed using the following gradient (flow-rate, 0.6 ml/min): 99% eluent RP1-A (see above) from 0 to 2 min, 1–100% eluent RP1-B (see above) from 2 to 22 min, 100% RP1-B from 22 to 26 min, 100% RP1-B to 99% RP1-A from 26 to 29 min, 99% RP1-A from 29 to 35 min. The absorbance was measured at 205 and 280 nm. Fractions were collected manually. Samples for amino acid sequence were prepared as described above.

*System 3.* The chromatographic unit was the same as in System 2. Separation of protein and peptides was achieved with the following gradient (flow-rate, 0.6 ml/min): 99% eluent RP3-A (0.1% HFBA in Milli-Q water) from 0 to 2 min, 1–80% eluent

RP3-B [0.1% HFBA in acetonitrile–1-propanol–Milli-Q water (60:30:10, v/v/v)] from 2 to 62 min, 80% RP3-B from 62 to 66 min, 80% RP3-B to 99% RP3-A from 66 to 69 min, 99% RP3-A from 69 to 75 min. The absorbance was measured at 205 and 280 nm. Fractions were collected manually. Samples for amino acid sequence analysis were prepared as described above.

*Protein quantitation*

Depending on the amount and purity of the protein, the following quantitation methods were used: a modified biuret method[10], the Bio-Rad kit (Bio-Rad Labs., Richmond, CA, U.S.A.), based on the method of Bradford[11], and integration of absorbance at 205 nm in reversed-phase HPLC (system 1; see above) with bovine RNAse and BSA as standards.

*Sodium dodecyl sulphate polyacrylamide gel electrophoresis and Western blotting*

SDS-PAGE was performed in 13% gels [acrylamide–bisacrylamide (37:1); gel dimensions, 0.75 mm thick × 13 cm wide × 7 cm long] according to the method of Laemmli with modifications as described in ref. 12. Molecular mass markers were from Amersham (Amersham), U.K.): lysozyme, 14.3 kDa; trypsin inhibitor, 21.5 kDa; carbonic anhydrase, 30 kDa; ovalbumin, 46 kDa; bovine serum albumin, 69 kDa, phosphorylase *b*, 92.5 kDa and myosin 200 kDa. Coomassie Brilliant Blue-stained gels were dried between Cellophane sheets on a Model 1125 slab gel dryer (Bio-Rad Labs.). The dried gels were scanned with a Model CS-930 dual-wavelength TLC scanner, connected to a Model DR-2 data recorder (Shimadzu, Kyoto, Japan).

Electrotransfer of proteins to nitrocellulose (0.45 μm, Schleicher & Schüll, Dassel, F.R.G.) was achieved according to ref. 13. After saturation with 1% BSA–0.01% Tween 20, the membrane was incubated with the monoclonal anti-p18 antibody described in ref. 14, excess of antibody was washed off and direct immunostaining was performed with anti-mouse Ig, horseradish peroxidase-linked $F(ab')_2$ fragment from sheep (Amersham) and 4-chloro-1-naphthol (Bio-Rad Labs.) as substrate.

*Amino acid sequence and composition*

Amino acid sequence analysis was performed on an Applied Biosystems (Foster City, CA, U.S.A.) Model 470A or 477A protein sequencer, connected to a Model 120A PTH-amino acid analyser.

After hydrolysis of protein with 6 *M* hydrochloric acid under nitrogen for 24 h at 106°C in a PicoTag Workstation (Waters), the amino acid composition analysis was performed on an Applied Biosystems Model 420A derivatizer connected to a Model 130A separation system. To obtain a best fit for the composition data of p18LM, the sums of squared errors ($AA_{\text{theor.}} - AA_{\text{determ.}}$) were calculated for p18LM between 95 and 130 amino acids. A minimum was found for p18LM, corresponding to p18, missing 13 amino acids from the C-terminus.

*Calculation of protein parameters by computer*

Protein parameters, *i.e.*, molecular mass ($M_c$), isoelectric point ($pI_c$) and Hopp index of antigenicity, were calculated on an IBM AT computer, using the DNA-star program (Computer Systems for Molecular Biology, Madison, WI, U.S.A.).

## RESULTS AND DISCUSSION

Extraction of *E. coli* cells (TGE901 pTG-2153-HIV-1) showed that recombinant p18 is a soluble protein and that it can be separated from cell fragments and aggregated proteins by collecting the supernatants of repetitive freezing–thawing cycles. Fig. 1 shows the SDS-PAGE analysis of successive homogenates (H1, H2 and H3) and the corresponding supernatants (S1, S2 and S3). Recombinant p18 is found in the first and second supernatants, which were pooled and adjusted to pH 6.0. On leaving the pool in the cold overnight, we observed a protein precipitate, which did not contain recombinant p18 (Fig. 1, lane PP). The clarified pool was applied to a column of S Sepharose. A soft gel cation-exchange sorbent was chosen because protein continued to precipitate during sample application and the regeneration of Sepharose was considered to be easier than the regeneration of a closed HPLC column.

SDS-PAGE analysis of the non-adsorbed fraction (Fig. 1, lane 1) shows a protein band with about the same mobility as recombinant p18 ($M_r$ 17.4 kDa). This protein is not identical with recombinant p18, as will be shown below (Fig. 8, lane C). Fractions which were step-eluted from S Sepharose at different salt concentrations are shown in Fig. 1, lanes 2–9. Recombinant p18 appears in the first and second protein peaks at 400 m*M* sodium chloride (lanes 6 and 7) and is completely eluted with 600 m*M* sodium chloride (lane 8). The fractions in lanes 7 and 8 were pooled (pool A); the fraction in lane 6 was kept separately (pool B).

After ultrafiltration of pools A and B, we observed a contaminant derived from

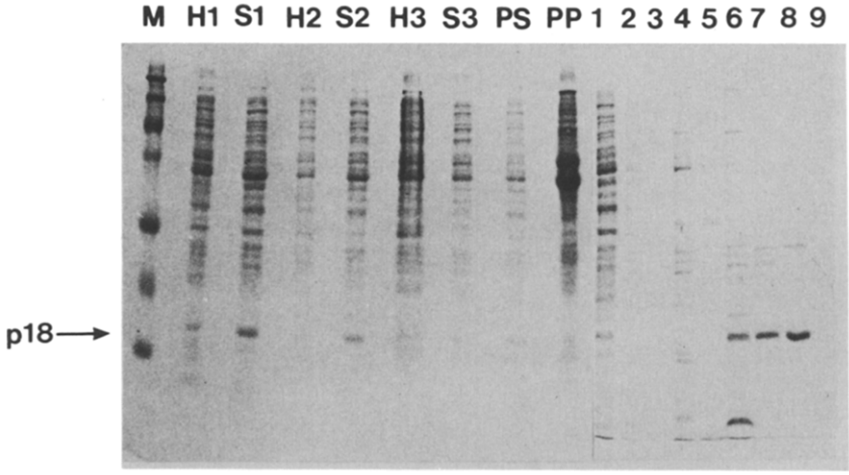


Fig. 1. Coomassie Brilliant Blue-stained SDS-polyacrylamide gel of recombinant p18 after extraction of protein from *E. coli* and after S Sepharose chromatography. M = molecular mass markers (see Experimental); H1 = *E. coli* cell suspension (first homogenate); S1 = first supernatant after freezing of *E. coli* cells; H2 = *E. coli* cell suspension (second homogenate); S2 = second supernatant after freezing of *E. coli* cells; H3 = *E. coli* cell suspension (third homogenate); S3 = third supernatant after freezing of *E. coli* cells; PS = pool of supernatants S1 and S2; PP = protein precipitate after adjusting PS to pH 6.0. Fractions 1–9 = S Sepharose chromatography: 1 = non-absorbed material; 2 and 3 = elution with half-concentrated PBS; 4 = PBS elution; 5 = 200 m*M* NaCl; 6 = 400 m*M* NaCl (first peak); 7 = 400 m*M* NaCl (second peak); 8 = 600 m*M* NaCl; 9 = 800 m*M* NaCl.

recombinant p18, namely p18LM, appearing with a lower molecular mass (15.1 kDA), as determined by SDS-PAGE. In different preparations of recombinant p18 we always found p18LM in the ultrafiltration step. For the following reasons we decided not to add protease inhibitors during this step; (1) purified recombinant p18 was to be used in animal experiments, and no potential health hazard to the animals could be tolerated; (2) as will be shown below, during cation-exchange HPLC recombinant p18 and p18LM could be separated; (3) the loss of *ca.* 13% C-terminally cleaved recombinant p18 (4.8 mg p18LM, Table I) was still considered tolerable; and (4) purified recombinant p18, stored in phosphate–sodium chloride buffer at −80°C, proved to be stable.

The concentrated pools A and B were separately chromatographed on a Sulfoethyl Aspartamide HPLC column. Fig. 2 shows the chromatographic separation of pool A. Elution of recombinant p18 was achieved with 32 m*M* sodium phosphate–635 m*M* sodium chloride and p18LM was eluted with 33 m*M* sodium phosphate–665 m*M* sodium chloride. The inset in Fig. 2 shows the SDS-PAGE analysis of the starting material (lane S) and of fractions A–L, as indicated in the chromatogram. Most of recombinant p18 is found in fractions E, F and G, whereas the majority of p18LM is eluted in fractions H and I. Pool B was analogously chromatographed. Fractions from both runs, containing either predominantly recombinant p18 or p18LM, were pooled separately and rechromatographed on the same column under identical conditions.

Table I summarizes the protein balance of recombinant p18 purification from *E. coli*. Scanning of Coomassie Brilliant Blue-stained SDS-polyacrylamide gels of the starting material (homogenate, *i.e.*, cell suspension before lysis of cells; see Experimental) and of purified recombinant p18 (Fig. 3, lane A) gives an estimate of about a 20% final yield. However, this value may be underestimated owing to the 17.4 kDa protein which co-migrates with recombinant p18 and which becomes visible in the S Sepharose flow-through (Fig. 1, lane 1), thus contributing to the scanned peak of recombinant p18 in the homogenate.

Purified recombinant p18 and p18LM were analysed by SDS-PAGE (Fig. 3) and

TABLE I

PROTEIN BALANCE OF RECOMBINANT p18 PURIFICATION FROM *E. COLI*

| *Analysed fraction*[a] | *Total protein (mg)* | *Recovery of protein (%)* | *Purity (%)*[b] |
|---|---|---|---|
| Homogenate (*E. coli*, suspended in PBS) | 7870 | 100.0 | 1.6 |
| Pool of supernatants $S_1$ + $S_2$ after freezing–thawing | 1440 | 18.3 | 6.2 |
| Pool of supernatants before S Sepharose chromatography | 367 | 4.7 | N.D. |
| 400 m*M* elution (1st peak) | 16.2 | 0.21 | 23.0 |
| 400 m*M* elution (2nd peak) | 13.9 | 0.18 | 67.3 |
| 600 m*M* elution | 15.5 | 0.20 | 80.8 |
| Purified recombinant p18 | 28.7 | 0.3 | 90.0 |
| Purified p18LM | 4.8 | 0.06 | 93.3 |

[a] For a detailed description, see Experimental.

[b] Coomassie Brilliant Blue-stained SDS-polyacrylamide gels were scanned, and the area corresponding to recombinant p18 (or p18LM) was expressed as a percentage of the total area scanned.

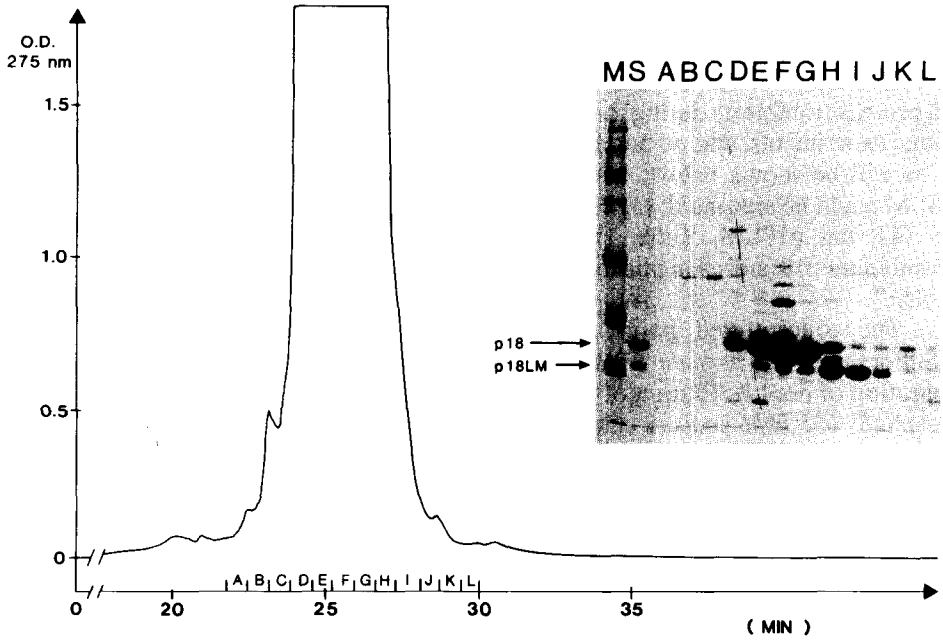


Fig. 2. Cation-exchange high-performance liquid chromatography of recombinant p18 and analysis of fractions by SDS-PAGE. Cation-exchange HPLC on Sulfoethyl Aspartamide was performed as described under Experimental. Protein (20 mg) was loaded in 40 ml of CatEx-A buffer. Inset: Coomassie Brilliant Blue-stained SDS-polyacrylamide gel, showing the analysis of fractions A–L. M = molecular mass markers (see Experimental). The positions of recombinant p18 and p18LM are indicated by arrows.

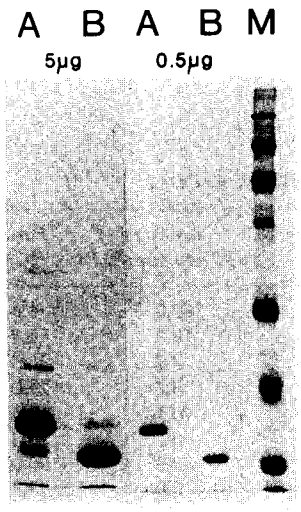


Fig. 3. Coomassie Brilliant Blue-stained SDS-polyacrylamide gel of purified recombinant p18 and p18LM at loadings of 5 and 0.5 μg per lane. A = Purified recombinant p18; B = purified p18LM; M = molecular mass markers (see Experimental).

showed purities of 90% and 93.3%, respectively. When the SDS-PAGE lanes were overloaded with 5 $\mu$g of protein per slot (width 2 mm), in addition to contamination of recombinant p18 with p18LM, and *vice versa*, faint protein bands of $M_r$ 22.4 and 32.4 kDa could be detected (Fig. 3).

Analysis of purified recombinant p18 and p18LM by reversed-phase HPLC (Fig. 4) showed identical retention times for both proteins, despite the long gradient (system 3). Chromatography of a mixture of recombinant p18 and p18LM in system 3 showed no peak broadening but an increased peak height (not shown). The same result was obtained with reversed-phase system 1 (not shown), demonstrating the potential limitations of reversed-phase HPLC analysis of closely related proteins.

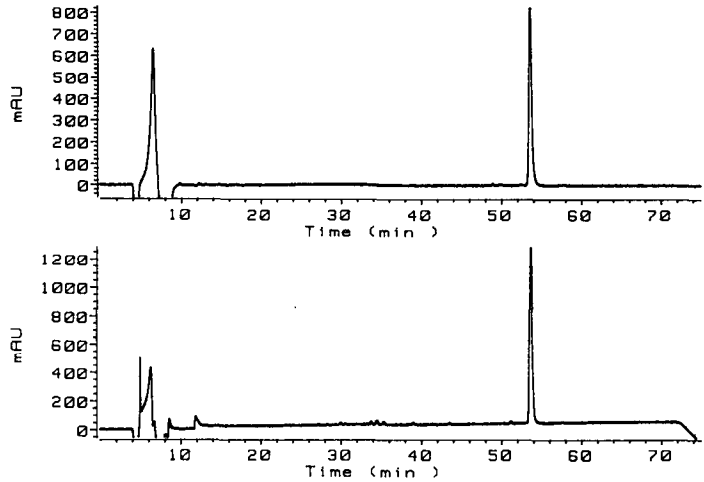


Fig. 4. Comparison of purified recombinant p18 and p18LM by reversed-phase HPLC in system 3, recorded at 205 nm. Upper panel, purified recombinant p18; lower panel, purified p18LM. Chromatograms are shown after baseline subtraction of a blank run. Retention times: recombinant p18, 53.644 min, p18LM, 53.687 min.

The same sequence (Gly–Ala–Arg) was obtained on N-terminal amino acid sequence determination of purified recombinant p18 and p18LM. Fig. 5 summarizes the amino acid sequence results obtained from tryptic and V8 proteolytic peptides of purified recombinant p18 and p18LM. The tryptic peptide of purified recombinant p18, eluted at a retention time of 19.4 min (Fig. 6, upper panel), represents the C-terminus of the protein (Table II) and is absent from the tryptic digest of p18LM (Fig. 6, lower panel). The same evidence of a missing C-terminus is obtained after V8 proteolysis: the peptide peak at retention time 11.8–12.2 min (Fig. 7, upper panel) of the recombinant p18 digest contains two successive C-terminal peptides (Table II), which are absent from the p18LM digest (Fig. 7, lower panel). The peptides eluted at retention time 16.1–16.6 min, displaying switched intensities of recombinant p18 and p18LM (Fig. 7, upper and lower panel), show an identical amino acid sequence. The different relative mobilities in reversed-phase HPLC may be due to differences in the oxidation state of $Cys_{56}$ in recombinant p18 and p18LM. Table II identifies all sequenced peptides by their retention times in reversed-phase system 2 and/or system 3.

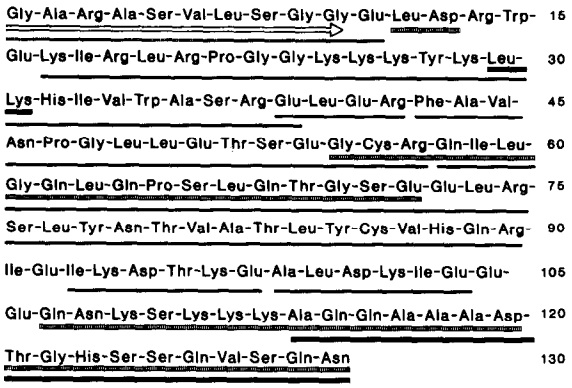


Fig. 5. Amino acid sequence of recombinant p18 and alignment of sequenced tryptic and V8 proteolytic peptides. Solid bars = tryptic peptides; hatched bars = V8 proteolytic peptides; thick bars = peptides obtained from recombinant p18 digests; thin bars = peptides obtained from p18LM; arrow = N-terminal amino acid sequence obtained from recombinant p18.

The amino acid composition of purified recombinant p18 and p18LM is shown in Table III. As identical N-termini were found for p18 and p18LM, we assumed that p18LM is a C-terminally cleaved form or a mixture of very similar C-terminally cleaved forms of p18, differing only in a few amino acids. We calculated for p18LM the sums of squared errors, comparing the theoretical amino acid composition of successive C-terminally cleaved molecules with the experimentally determined amino acid composition. A best fit was obtained for p18LM consisting of the first 117 amino acids ($Gly_1$–$Ala_{117}$) of recombinant p18. Calculation of the isoelectric point of this

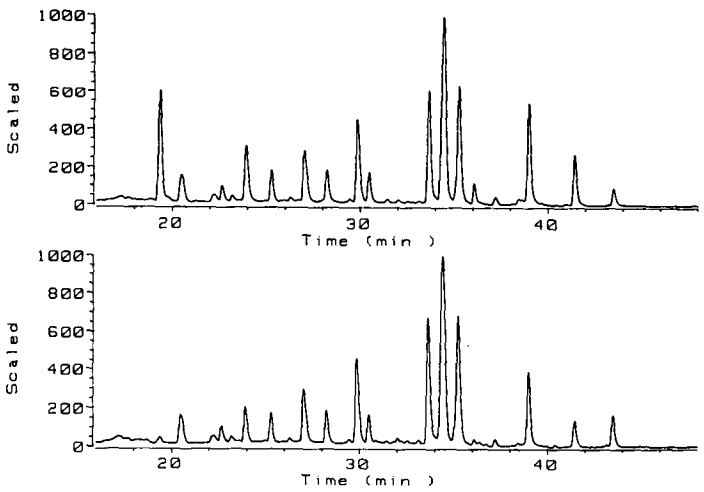


Fig. 6. Comparison of recombinant p18 and p18LM tryptic digests by reversed-phase HPLC in system 3, recorded at 205 nm. Upper panel, peptide map of recombinant p18; lower panel, peptide map of p18LM. Only the region containing peptide peaks is shown. To facilitate comparison, both chromatograms were normalized to the highest peak (value of 1000) at 34.4 min.

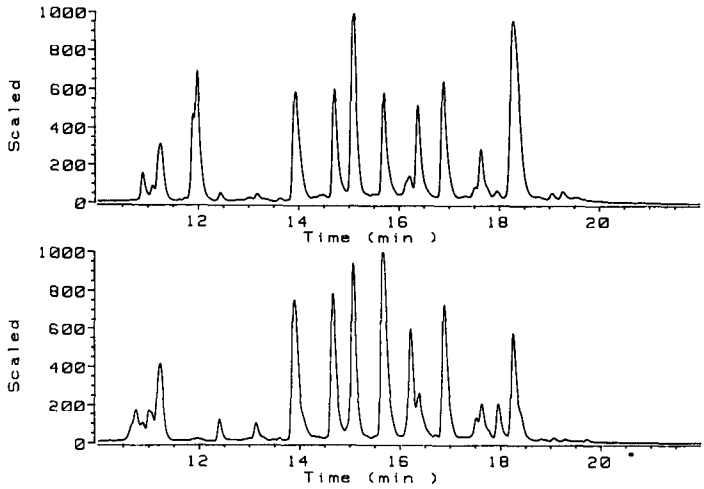


Fig. 7. Comparison of recombinant p18 and p18LM V8 protease digests by reversed-phase HPLC in systems 2, recorded at 205 nm. Upper panel, peptide map of recombinant p18; lower panel, peptide map of p18LM. Only the region containing peptide peaks is shown. To facilitate comparison, both chromatograms were normalized to the highest peak.

TABLE II

REVERSED-PHASE HPLC RETENTION TIMES OF TRYPTIC AND V8 PROTEOLYTIC PEPTIDES DERIVED FROM PURIFIED RECOMBINANT p18 AND p18LM

| *Peptide* | *Retention time (min)* | |
|---|---|---|
| | *System 2*[a] | *System 3*[a] |
| *Recombinant p18* | | |
| V8 protease: $Leu_{12}$–$Asp_{13}$ | 11.8–12.2 | 30.5 |
| V8 protease: $Gly_{55}$–$Glu_{72}$ | 16.35 | 38.7 |
| V8 protease: $Gln_{107}$–$Asp_{120}$ | 11.8–12.2 | 33.1 |
| V8 protease: $Thr_{121}$–$Asn_{130}$ | 11.8–12.2 | 30.5 |
| Trypsin: $Leu_{30}$–$Lys_{31}$ | N.D. | 36.2 |
| Trypsin: $Ala_{114}$–$Asn_{130}$ | N.D. | 19.3 |
| *p18LM* | | |
| V8 protease: $Gly_{1}$–$Glu_{11}$ | N.D. | 28.7 |
| V8 protease: $Lys_{17}$–$Glu_{39}$ | N.D. | 43.8 |
| V8 protease: $Gly_{55}$–$Glu_{72}$ | 16.5 | 38.7 |
| V8 protease: $Ile_{93}$–$Glu_{98}$ | 11.2 | 26.8 |
| V8 protease: $Ala_{99}$–$Glu_{104}$ | N.D. | 28.7 |
| Trypsin: $Glu_{39}$–$Arg_{43}$ | N.D. | 23.9 |
| Trypsin: $Phe_{43}$–$Arg_{57}$ | N.D. | 34.4 |
| Trypsin: $Gln_{58}$–$Arg_{75}$ | N.D. | 34.4 |
| Trypsin: $Ser_{76}$–$Arg_{90}$ | N.D. | 43.4 |

[a] V8 proteolytic peptide separation of purified recombinant p18 and p18LM in system 2 is shown in Fig. 7. Tryptic peptide separation of purified recombinant p18 and p18LM in system 3 is shown in Fig. 6.

TABLE III

AMINO ACID COMPOSITION OF RECOMBINANT p18 AND p18LM

| *Amino acid* | *Recombinant p18* | *p18LM* |
|---|---|---|
| Asp + Asn | 8.0 ( 8)[a] | 6.0 ( 6)[b] |
| Glu + Gln | 24.8 (23) | 21.8 (21) |
| Ser | 10,4 (11) | 8.1 ( 8) |
| Gly | 9.9 (10) | 9.7 ( 9) |
| His | 3.0 ( 3) | 2.0 ( 2) |
| Arg | 10.4 ( 9) | 10.1 ( 9) |
| Thr | 6.0 ( 6) | 5.2 ( 5) |
| Ala | 10.1 (10) | 8.8 ( 8) |
| Pro | 3.0 ( 3) | 3.1 ( 3) |
| Tyr | 2.9 ( 3) | 2.9 ( 3) |
| Val | 5.4 ( 6) | 4.4 ( 5) |
| Met | 0.1 ( 0) | 0.0 ( 0) |
| Cys | 1.2 ( 2) | 0.8 ( 2) |
| Ile | 5.1 ( 6) | 4.9 ( 6) |
| Leu | 13.5 (14) | 13.3 (14) |
| Phe | 1.2 ( 1) | 1.2 ( 1) |
| Lys | 13.2 (13) | 12.7 (13) |
| Trp | N.D. ( 2) | N.D. ( 2) |
| Total | 130 | 117 |

[a] Values in parentheses correspond to the actual number of amino acids in recombinant p18.

[b] Values in parentheses correspond to the "best-fit" number of amino acids, as determined for p18LM (see Experimental).

protein gives a value of $pI_c = 9.47$, which is slightly more basic than the value of $pI_c = 9.23$ calculated for a complete recombinant p18. This finding correlates well with the fact that p18LM is eluted at higher salt concentrations than recombinant p18 from the Sulfoethyl Aspartamide HPLC column (Fig. 2).

Fig. 8 shows the immunostain (II) of purified recombinant p18 and p18LM obtained with a monoclonal antibody which was selected for binding to natural p18 after a mouse had been injected with whole HIV-1[14]. Recombinant p18 is recognized in *E. coli* homogenate (Fig. 8II, lane E), in supernatant after freezing–thawing cycles (Fig. 8II, lane D), and as the purified protein (Fig. 8II, lane B).

No immunostain was detected in *E. coli* wild-type homogenate (Fig. 8II, lane F) or in the non-adsorbed fraction of S Sepharose (Fig. 8II, lane C), thus showing that the 17.4-kDa protein in this fraction (visible in the Coomassie Brilliant Blue-stained SDS-polyacrylamide gel, Fig. 8I) is not identical with recombinant p18.

In spite of the missing C-terminus, p18LM still contains the epitope(s) which is/are recognized by the monoclonal antibody, as is shown by the immunostain (Fig. 8II, lane A). The calculated Hopp index for antigenicity shows a maximum for $Glu_{104}$–$Lys_{112}$. an amino acid sequence which, according to our analysis (amino acid composition, $pI_c$), is present in p18LM.

We have isolated recombinant p18 from *E. coli* in an overall yield of *ca.* 20% and a final purity of 90% as determined by SDS-PAGE and >95% as determined by reversed-phase HPLC. Owing to the absence of a post-translational myristylation

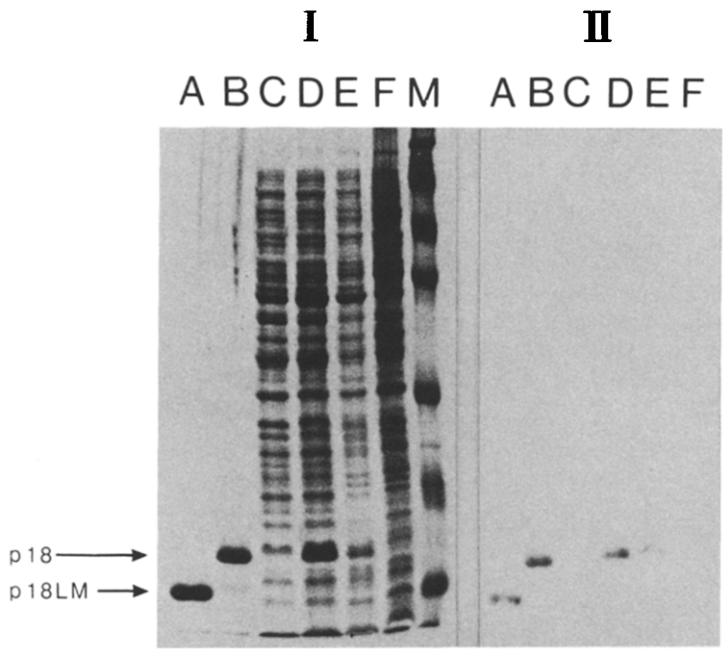


Fig. 8. Comparison of Coomassie Brilliant Blue-stained SDS-polyacrylamide gel (I) and immunostain after Western blotting (II). A = Purified p18LM; B = purified recombinant p18; C = S Sepharose flow-through; D = supernatant after freezing–thawing cycles; E = *E. coli* homogenate; F = *E. coli* wild-type homogenate; M = molecular mass markers (see Experimental).

pathway in *E. coli*, the protein lacks the myristic acid at the N-terminal glycine. As the codon UAC in position 611–613 of the *gag* gene[1] was point-mutated to generate a UAG stop codon, the C-terminal tyrosine of natural p18 is missing from recombinant p18. This was confirmed by amino acid sequence analysis (Fig. 5, Table II).

A monoclonal antibody, which had been generated against natural p18 from HIV-1, recognized purified recombinant p18, thus making the protein appear suitable for preliminary tests for diagnostic and immunological applications.

ACKNOWLEDGEMENTS

We thank J. C. Gluckman for a generous gift of anti-p18 monoclonal antibody, F. Daul for the photographic work and F. Vernon for the bibliography. This work was supported by Pasteur Vaccins and Diagnostics Pasteur.

REFERENCES

1 S. Wain-Hobson, P. Sonigo, O. Danos, S. Cole and M. Alizon, *Cell*, 46 (1986) 63.
2 M. G. Sarngadharan, F. D. Veronese, S. Oroszlan, S. Arya and R. C. Gallo, *Acquired Immunodeficiency Syndrome*, Elsevier, Paris, 1987, p. 43.
3 H. R. Gelderblom, E. H. S. Hausmann, M. Ozel, G. Pauli and M. A. Koch, *Virology*, 156 (1987) 171.
4 G. Schmidt, K. Amiraian, H. Frey, R. W. Stevens and D. S. Berns, *J. Clin. Microbiol.*, 25 (1987) 1993.

5 D. S. Burke, R. R. Redfield, P. Putman and S. S. Alexander. *J. Clin. Microbiol.*, 25 (1987) 81.
6 R. L. Shoeman, D. Young, R. Pottathil, J. Victor, R. R. Conroy, R. M. Crowl, T. Coleman, E. Heimer, C. Y. Lai, K. Ganguly, E. P. Reddy, A. M. Skalka, P. R. Pine, F. R. Kahn and H. Weissbach, *Anal. Biochem.*, 161 (1987) 370.
7 P. H. Naylor, C. W. Naylor, M. Badamchian, S. Wada, A. L. Goldstein, S. S. Wang, D. K. Sun, A. H. Thornton and P. S. Sarin, *Proc. Natl. Acad. Sci. U.S.A*, 84 (1987) 2951.
8 P. S. Sarin, D. K. Sun, A. H. Thornton, P. H. Naylor and A. L. Goldstein, *Science (Washington, D.C.)*, 232 (1986) 1135.
9 G. Rautmann, M.-P. Kieny, R. Brandely, K. Dott, M. Girard, L. Montagnier and J.-P. Lecocq, *Aids Research and Human Retroviruses*, (1989) in press.
10 B. Kadenbach, *Biochem. Z.*, 344 (1966) 49.
11 M. Bradford, *Anal. Biochem.*, 72 (1976) 248.
12 H. V. J. Kolbe, D. Costello, A. Wong, R. C. Lu and H. Wohlrab, *J. Biol. Chem.*, 259 (1984) 9115.
13 H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.*, 76 (1979) 4350.
14 J. Chassagne, P. Verrelle, C. Dionet, F. Clavel, F. Barre Sinoussi, J. C. Chermann, L. Montagnier, J. C. Gluckman and D. Klatzmann, *J. Immunol.*, 136 (1986) 1442.