Journal of Chromatography, 521 (1990) 267–277
Elsevier Science Publishers B.V., Amsterdam

CHROMSYMP. 1962

Purification of a recombinantly produced transmembrane protein (gp41) of HIV I

ERWIN SOUTSCHEK*, BRIGITTE HÖFLACHER and MANFRED MOTZ
Mikrogen GmbH, Westendstrasse 125, 8000 München 2 (F.R.G.)

ABSTRACT

The transmembrane protein gp41, a component of the viral envelope of HIV I, and its analogue gp36 of HIV II are important antigens for the sensitive and specific detection of anti-HIV antibodies. The immunodominant region of the protein gp41, which reacts with 100% of sera of infected persons, was produced by gene technological means in *Escherichia coli*. The protein accumulates in the form of insoluble inclusion bodies in the bacterial cell. Purification strategies for this aggregated material depend mainly on the isolation of these "inclusion bodies" and subsequent washing procedures. Growth conditions of the recombinant *E. coli* cells and the method of the cell disruption are important for the efficiency of purification and the recovery of the antigen. Owing to the insolubility of the expressed antigen, a significant concentration of recombinant gp41 was possible by extracting the soluble cell components. For this purpose, mild detergent solutions and low-molarity chaotropic buffer solutions were used. After final solubilization in 8 *M* urea buffer at pH 12.5, further chromatographic purification steps followed. The reduction of disulphide bridges with β-mercaptoethanol or dithiothreitol was important. Gel filtration on a Sephacryl S-200 or Superose 12 column and/or ion-exchange chromatography on a DEAE-Sepharose Fast Flow or Mono Q HR (5/5) column finally resulted in the desired purity of the antigen.

INTRODUCTION

A group of closely related retrovirus isolates designated human T-cell lymphotrophic virus type III (HTLV-III), AIDS-related virus (ARV) and lymphadenopathy-associated virus (LAV), collectively referred to as human immunodeficiency virus (HIV), has been identified as the primary etiologic agent of the acquired immunodeficiency syndrome (AIDS) and AIDS-related complex (ARC) [1,2]. The infectious virus particle contains RNA wherein the genetic information is encoded. On invasion into susceptible cells the genetic information is transcribed into DNA. Integration into the chromosome occurs along with the production of virus-specific proteins. Known viral proteins are the group-specific antigens (gag), the pol encoded antigens (protease, reverse transcriptase, integrase), the envelope (env) antigens (gp120, gp41) and the regulatory proteins TAT, REV, VIF and NEF [3].

The transmission of the HI virus occurs in most instances by blood-to-blood contact. After a period of a few weeks or months seroconversion can be observed by detection of antibodies against viral proteins, namely p24 and env (gp120, gp41) derived antigens.

0021-9673/90/$03.50

Anti-gp41 antibodies are currently the most important marker for HIV-I, as this protein reacts with virtually all sera positive for HIV-I. To date, no cross-reactivity has been found and relatively few variances in the amino acid sequence of differing isolates are known [4–8]. A recombinant antigen was produced in *Escherichia coli* by expressing almost the complete nucleotide sequence encoding the outwardly directed region of the transmembrane protein gp41 of HIV-I [9,10]. The purified recombinant protein is intended to be used as diagnostic reagent in a new confirmatory assay based on the Western blot principle.

EXPERIMENTAL

Extraction of the recombinant protein

An overnight culture of *E. coli* (strain JM 83 or JM 109) in L-broth [10 g of tryptone (Difco), 5 g of yeast extract (Difco), 5 g of NaCl, pH 7.4] harbouring the expression plasmid (pUC18RSp, pLIN41 or pDS2BPs) was diluted (1:50) and grown further for 3 h at 37°C to an absorbance at 578 nm of 0.7–0.8. After this the lac promoter was induced with 1 m*M* isopropyl-*β*-D-thiogalactopyranoside (IPTG) (Biomol) for an additional 3–4 h to express the recombinant antigen. The cells were harvested by centrifugation for 10 min at 8000 *g* (4°C). After washing in lysis buffer [50 m*M* Tris–HCl (Boehringer) (pH 8.0 at 25°C)–2 m*M* EDTA–1 m*M* dithiothreitol (DTT) (Biomol) 5% glycerol] the cells were resuspended (4 ml/g wet weight) in the same buffer [with the addition of 0.4 mg/ml lysozyme and 1 m*M* phenylmethylsulphonyl fluoride (PMSF) (Sigma)]. The suspension was incubated for 30 min and centrifuged for 45 min at 12 000 *g*. The pellet was resuspended in lysis buffer containing 0.1% Triton X-100 (Serva) and 0.1 m*M* PMSF (25 ml/g wet weight). The suspension was stirred for 30 min and Benzonase (DNase) (Merck) (1 *μ*l per 25 ml) was added to reduce the viscosity due to the DNA release. NaCl was added to a final concentration of 1 *M* and the suspension was stirred for 2 h prior to centrifugation for 30 min at 12 000 *g*. The pellet fraction containing the "inclusion bodies" was suspended (10 ml/g wet weight) in detergent buffer [50 m*M* Tris–HCl (pH 8.0)–2 m*M* EDTA–2 m*M* DTT 0.1 m*M* PMSF] containing 1.5% of octyl *β*-D-glucopyranoside (OGP) (Biomol) and stirred for 2 h. The suspension was centrifuged for 45 min at 12 000 *g* and the pellet fraction was washed either with 2, 4 or 6 *M* urea solution [in 50 m*M* Tris–HCl (pH 8.0)–2 m*M* EDTA–2 m*M* DTT–0.1 m*M* PMSF]. Finally, the pellet was solubilized in 8 *M* urea (Merck) in 20 m*M* Tris–HCl (pH 8.0)–2 m*M* DTT. To achieve complete solubilization the pH was increased to 12.5 by adding 1 *M* NaOH.

Purification of the solubilized antigen

For gel filtration, an S-200 Sephacryl HR column (100 × 1.6 cm I.D.) or a Superose 12 HR 10/30 column (300 × 10 mm I.D.) (Pharmacia) in 8 *M* urea buffer [50 m*M* Tris–HCl (pH 8.0)–2 m*M* DTT] were used. Fractions containing the recombinant protein were then purified to apparent homogeneity by reversed-phase high-performance liquid chromatography (HPLC) on a ProRPC 15-*μ*m HR 16/10 or Mono Q HR 5/5 column (Pharmacia). The ProRPC column was developed with an organic solvent gradient [0 to 100% acetonitrile–isopropanol (2:1); flow-rate 1 ml/min]. The protein-containing solution was acidified prior to injection with 0.1% trifluoroacetic acid (TFA). The Mono Q column was developed with a 20 m*M*

ethanolamine buffer containing 8 *M* urea (pH 9.8) with a gradient from 0 to 1 *M* NaCl (flow-rate 1 ml/min). The fast protein liquid chromatographic (FPLC) columns (Pharmacia–LKB) were attached to an inert HPLC system consisting of a Model 2249 gradient pump and a Model 2150 UV monitor from Pharmacia–LKB.

Alternatively, the solubilized antigen was loaded directly on a DEAE-Sepharose Fast Flow column (25 × 2.3 cm I.D.). The column was developed with a gradient from 0 to 1 *M* NaCl in 8 *M* urea [20 m*M* ethanolamine (pH 9.8)–2 m*M* DTT 0.1 m*M* PMSF; flow-rate 0.5 ml/min].

Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and Western blotting

Discontinuous SDS-PAGE was performed according to Laemmli [11] using a 3% stacking gel and a 17% running gel. Separated proteins were transferred by electroblotting between two graphite plates in blotting buffer [0.2 *M* glycine–25 m*M* Tris base–20% (v/v) methanol (pH 8.2)] for 1 h onto nitrocellulose. The transferred proteins can be renderd visible by staining the nitrocellulose with Ponceau S (Serva). The nitrocellulose sheets were blocked with dry milk powder solution [5% dry milk powder–0.05% Tween 20–10 m*M* Tris–HCl (pH 7.5)–150 m*M* NaCl–0.1% NaN_3]. For detection of specific bands an HIV serum pool was used (inactivated for 45 min at 58°C). Non-specific bands were rendered visible by incubation with a high-titered anti-*E. coli* rabbit serum. Bound antibodies were detected by incubation with peroxidase conjugated anti-human immunoglobulin G (IgG) (Dakopatts) or anti-rabbit IgG (Dakopatts) followed by peroxidase-catalysed conversion of 3,3'-diaminobenzidine tetrahydrochloride (Sigma) and H_2O_2 (0.05%) to an insoluble brown pigment [10].

RESULTS AND DISCUSSION

The plasmid vector pUC18st was used to express an Rsa–Ssp-restriction fragment from the envelope gene of an HIV-1 isolate (WF1.13) closely resembling the BH10 [12] isolate. The expressed fragment encodes the exterior part of the transmembrane protein gp41 omitting the amino-terminal fusion peptide, including amino acid residues 531–674 [12]. It was ligated into pUC or pLIN vectors and *E. coli* (strain JM83 or JM109) cells were transformed with this plasmid construct. The expressed antigen should have an approximate molecular mass of about 17 000 dalton and, depending on the plasmid construct, from 4 to 15 foreign amino acids attached to the amino terminus.

For efficient expression of the protein and good recovery, different growth conditions were examined. Double concentrated L-broth supplemented with 0.2 0.5% glycerine was finally used for cell growth. The lac promoter was induced by IPTG as indicated under Experimental. The cells were washed in lysis buffer and then lysed by addition of lysozyme and Triton X-100 with constant stirring. Alternatively, sodium deoxychlolate was used for destabilizing the cell membrane [13,14]. The cell pellet of 1–2 l fermentation broth could be easily disrupted with the Triton X-100 method. Cell pellets from up to 10 l are more conveniently handled by mechanical (French press) or ultrasonification methods with additional lysozyme treatment. The cell disruption should be nearly complete (>99%) otherwise the following purification

steps are hindered. To this viscous suspension DNase (Benzonase) was added to reduce the viscosity due to the release of chromosomal DNA. The addition of a protease inhibitor such as PMSF throughout the purification procedure was important to reduce degradation of the expressed protein. After adding NaCl to a final concentration of 1 *M*, the suspension was centrifuged for 30 min at 12 000 *g* (4°C) and the expressed protein was found in the pellet fraction. With this step most of the soluble impurities can be separated (Fig. 1).

The pellet fraction was then extracted with octyl β-D-glucopyranoside solution to remove *E. coli* membrane proteins from the inclusion bodies (Fig. 1). Neither the extractions with solutions containing increasing Triton X-100 concentrations (0.1–2%) nor octyl β-D-thioglucopyranoside (1.5%) showed the same good efficiency (data not shown). The suspension was centrifuged again at 12 000 *g* and the enriched "inclusion bodies" were suspended in chaotropic buffer solutions with increasing molarity [up to 8 *M* urea or 7 *M* guanidinium hydrochloride (GHCl) in 50 m*M* Tris–HCl (pH 8.5)–5 m*M* DTT].

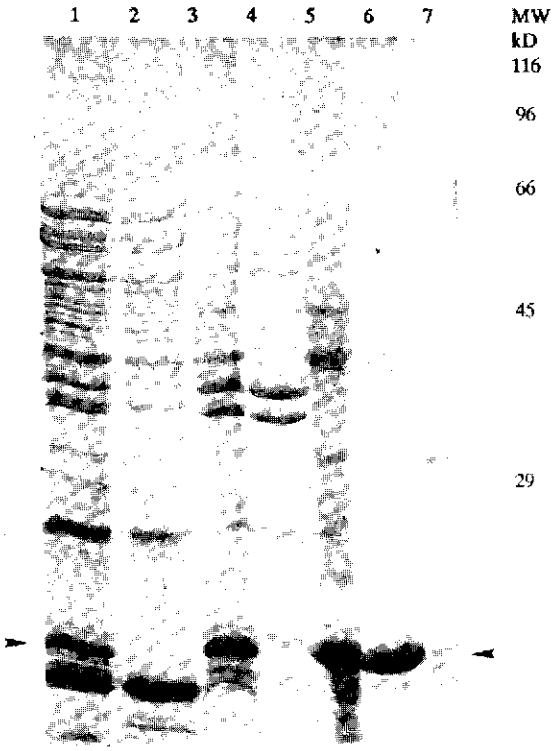

Fig. 1. Coomassie brilliant blue-stained 17% SDS-polyacrylamide gel of total *E. coli* proteins and isolated inclusion bodies. Molecular masses (MW) are indicated on the right, in kilodalton (kD). Lanes: 1 = total cell proteins from JM109/pDS2BPs resuspended in lysis buffer with addition of lysozyme; the position of the recombinant gp41 is marked with an arrow; 2 = lysis supernatant after 1 *M* NaCl extraction; 3 = lysis pellet resuspended in octyl β-D-glucopyranoside (OGP) buffer; 4 = supernatant of the OGP extraction; 5 = pellet of OGP extraction resuspended in 6 *M* urea buffer; 6 = pellet of 6 *M* urea extraction resuspended in 8 *M* urea buffer (pH 8.5); 7 = supernatant of 6 *M* urea extraction.

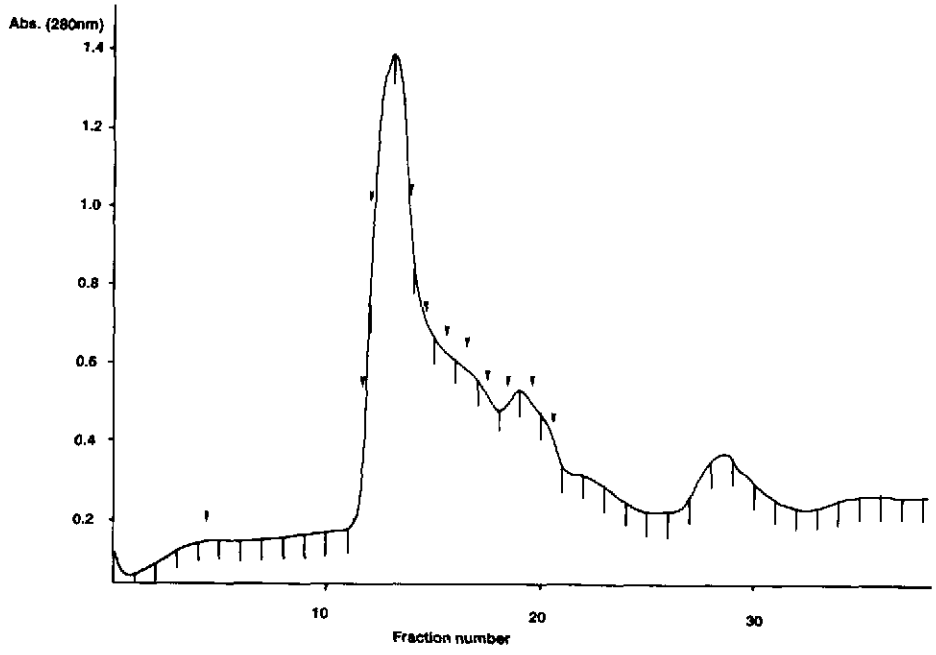

Fig. 2. Gel filtration chromatography of solubilized recombinant gp41 on an S-200 Sepharose column (100 cm × 1.6 cm I.D.) equilibrated with 8 *M* urea buffer [50 m*M* Tris–HCl (pH 8.5)–2 m*M* EDTA–2 m*M* DTT–0.1 m*M* PMSF]. Applied to the column were 6 ml of solubilized protein-containing solution (corresponding to 2 g wet weight of cells) and the proteins were eluted with 8 *M* urea buffer (flow-rate 0.5 ml/min). Fractions indicated with arrows were analysed by SDS-PAGE (Fig. 3).

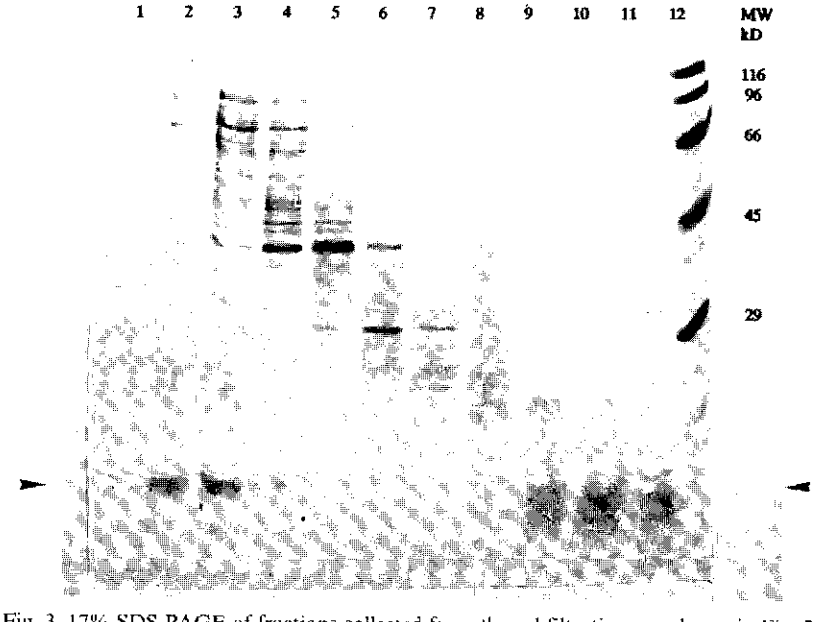

Fig. 3. 17% SDS-PAGE of fractions collected from the gel filtration run shown in Fig. 2. The molecular masses are indicated on the right. In lanes 1–11, the aliquots separated are those marked with arrows in Fig. 2 (from left to right). The position of the recombinant gp41 is marked with an arrow.

Impurities could be removed by washing the pellet with 6 *M* urea buffer. Finally, the resulting pellet was solubilized in 8 *M* urea or 6 *M* GHCl buffer by increasing the pH to 12.5 with 1 *M* NaOH. Some remaining impurities could be excluded by solubilization in 6 *M* GHCl buffer and dialysing extensively against 8 *M* urea buffer. Precipitated material was removed by centrifugation at 20 000 *g* for 45 min at 4°C. The solubilized antigen was chromatographed on a Sephacryl S-200 column. The protein was eluted with 8 *M* urea buffer [50 m*M* Tris–HCl (pH 8.5)–5 m*M* DDT–0.1% PMSF]. As shown in Figs. 2 and 3, the protein elutes mainly as a high-molecular-mass aggregate even in the presence of 8 *M* urea. A small part of the protein elutes as a low-molecular-mass protein as judged by SDS-PAGE and Western blot analysis (not shown). Neither the addition of 1% SDS nor running the column in 6 or 7 *M* GHCl buffer gave a better resolution. The reason for this anomalous behaviour could be due in part to the content of hydrophobic amino acids in this recombinant protein [15]. However, expression constructs omitting the entire transmembrane region and the hydrophobic amino-terminal fusion peptide behaved in nearly the same way. The presence of an extended amphipathic, α-helical fibre [7] might be responsible for the aggregation of gp41 by intermolecular bonding between these hydrophobic stretches. Gel filtration on a Superose 12 column in the above-described buffer resulted in a comparable elution profile.

Fractions containing the recombinant protein were further purified by reversed-phase chromatography on a ProRPC column. The recombinant protein eluted at >65% acetonitrile–isopropanol (2:1) (Fig. 4). This further emphasizes the hydro-

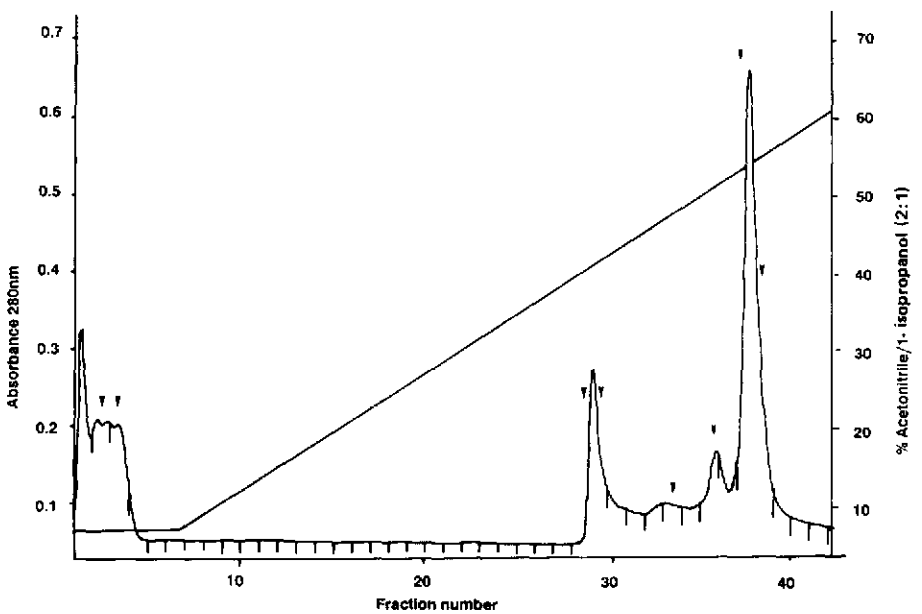

Fig. 4. Reversed-phase column run on a ProRPC column developed with a linear acetonitrile isopropanol (2:1) gradient. A 1-ml aliquot of the high-molecular-mass fraction from the gel filtration was injected after acidifying with 0.1% TFA. The proteins were eluted at a flow-rate of 1 ml/min. Aliquots of fractions indicated with arrows were analysed by SDS-PAGE (Fig. 5).

phobic behaviour of this protein. In SDS-PAGE the protein appears as a homogeneous and pure band (lane 7, Fig. 5a). Western blot analysis with HIV pool serum demonstrated the immunoreactivity of this reversed-phase purified protein (Fig. 5b). A protein band at about 34 000 dalton could be recognized as a dimer (in lane 7), which is seen more or less in every preparation even under highly reducing conditions (β-mercaptoethanol or DTT and SDS). However, the incubation with the high-titered anti-*E. coli* serum revealed some impurities which could only be seen by overloading the gel with a concentrated purified protein and silver staining (not shown).

Another approach which resulted in a more enriched gp41 with respect to impurities (up to 90%) was to solubilize the antigen in 8 *M* urea buffer at pH 12.5 and to titrate back to pH 9.4, which resulted in precipitation of impurities, leaving the recombinant antigen in solution. After centrifugation of the suspension and filtration, the antigen-containing solution was applied to a Mono Q HR 5/5 column in 8 *M* urea (containing 20 m*M* ethanolamine) buffer (pH 9.4). The column was developed with 8 *M* urea (containing 20 m*M* ethanolamine) buffer (pH 9.4) using a linear NaCl gradient from 0.1 to 1 *M*. The protein eluted in a pure form at >80% buffer B (8 *M*

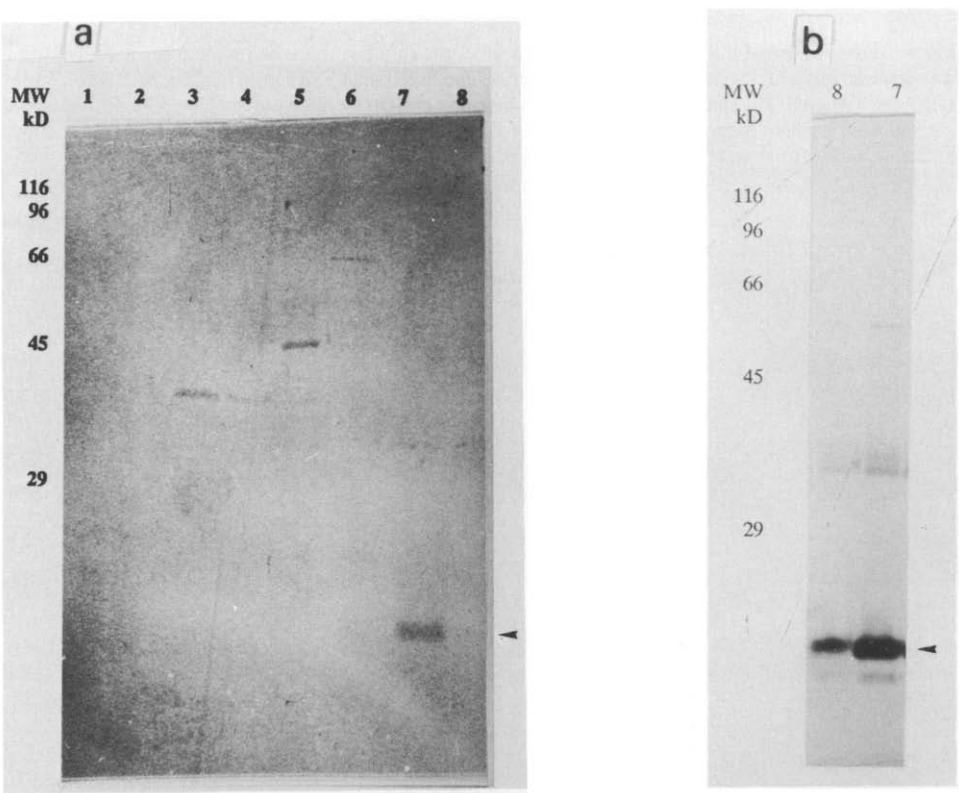

Fig. 5. (a) SDS-PAGE (17% gel) of fractions derived from the reversed-phase column run (Fig. 4). Molecular masses are indicated on the left. In lanes 1–8, the aliquots separated are those marked with arrows in Fig. 4 (from left to right). The recombinant gp41 is marked with an arrow. (b) Western blot analysis of the fractions shown in lanes 7 and 8 in (a) containing the recombinant gp41 protein.

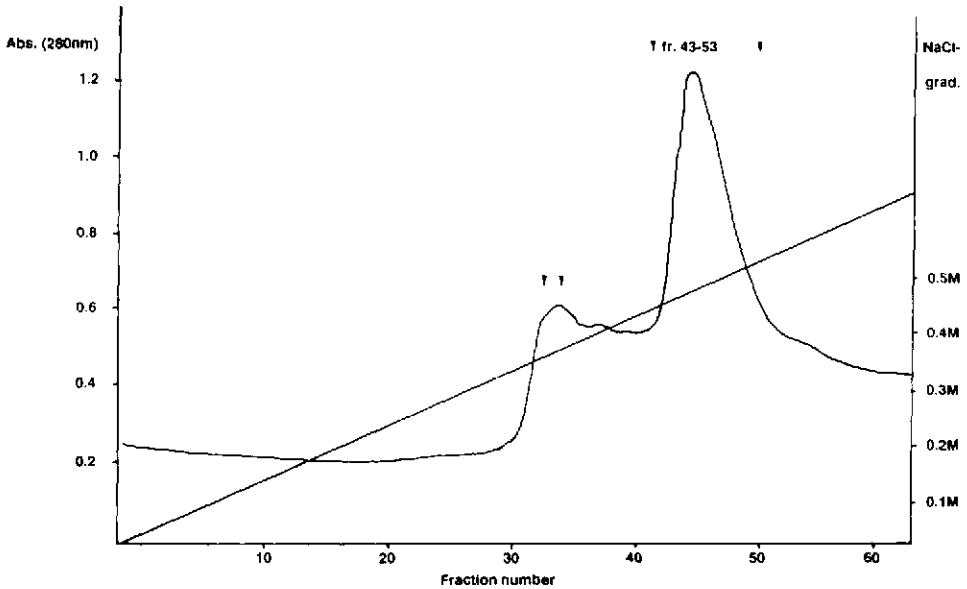

Fig. 6. Anion-exchange chromatography on a DEAE-Sepharose Fast Flow column (25 cm × 2.3 cm I.D.). The column was equilibrated with 20 m*M* ethanolamine buffer (pH 9.4) containing 2 m*M* EDTA, 5 m*M* DDT and 8 *M* urea. The solubilized protein fraction was applied in the same buffer. The proteins were eluted with a linear gradient from 0 to 1 *M* NaCl. The recombinant gp41 elutes at about 0.5 *M* NaCl. Fractions indicated with arrows were analysed by SDS-PAGE (Fig. 7).

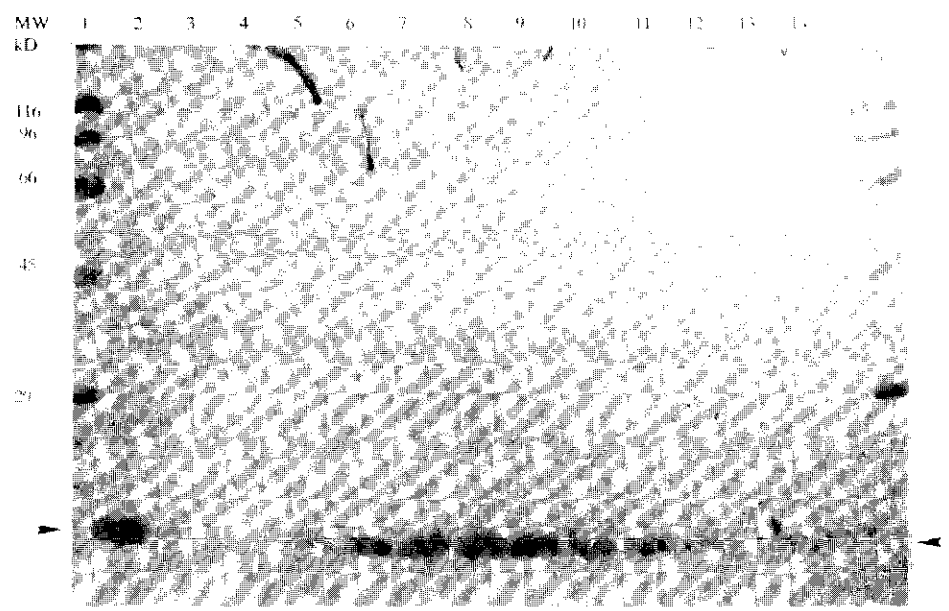

Fig. 7. SDS-PAGE of aliquots from the anion-exchange run on the DEAE-Sepharose Fast Flow column (Fig. 6). Molecular masses are indicated on the left. In lane 2 an aliquot of the filtered solution which was applied to the column is shown. In lanes 3–14, the aliquots separated are those marked with arrows in Fig. 6. The recombinant gp41 is marked with an arrow.

urea, pH 9.4, 20 m*M* ethanolamine, 1 *M* NaCl). Western blot analysis showed an immunoreactive antigen free of contaminants (similar as shown in Fig. 7).

For producing greater amounts of up to 10 mg per column run we finally used a DEAE-Sepharose Fast Flow column (25 × 2.6 cm I.D.) with 8 *M* urea buffer (20 m*M* ethanolamine–5 m*M* DTT–0.1% PMSF) and a linear NaCl gradient from 0 to 1 *M* (Figs. 6 and 7). The recombinant protein eluted at >0.5 *M* NaCl. Western blot analysis showed a pure and immunoreactive protein. No impurities could be detected by incubation with the anti-*E. coli* rabbit serum (Fig. 8). Depending on the amount of gp41 applied to the SDS gel, multiple banding and smearing occurs, together with a prominent band at the size of the dimer. These bands could only be seen in the more sensitive Western blot analysis and depend also on the serum dilution and the anti-gp41 antibody content. These minor bands are not seen after staining with Coomassie brilliant blue. The multiple banding could be due to the tendency of the gp41 protein to aggregate; even in the presence of SDS a minor portion appears in an aggregated form of variable size.

The purified antigen could now be used in a diagnostic assay system, based on the Western blot principle. For this, antigens such as recombinantly produced and purified gp41, p24 and reverse transcriptase from HIV I, and gp36 from HIV II were

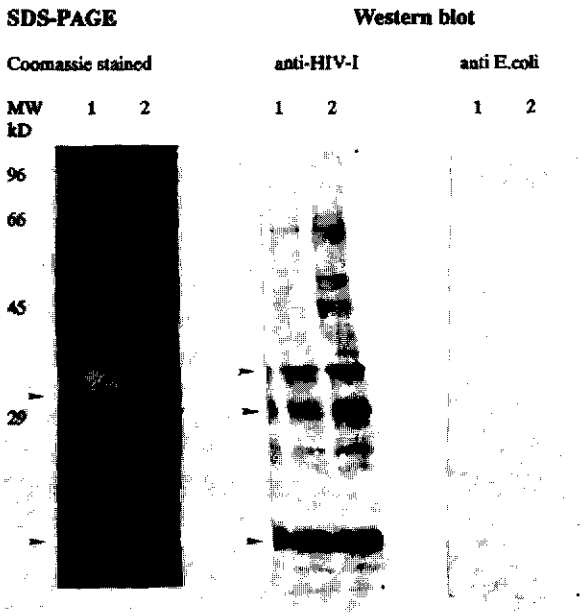

Fig. 8. SDS-PAGE and Western blot analysis of purified recombinant gp41. Molecular masses are indicated on the left. In lanes 1 and 2 purified antigen preparations of different batches are separated. The Coomassie brilliant blue-stained gel shows a homogeneous band at 16 000 dalton. There is also a band at about 32 000 dalton, which could be identified as a dimer of the recombinant protein. The Western blot incubated with HIV pool serum showed the immunoreactivity of the purified antigen. As can be seen, even under denaturing conditions in the SDS-PAGE aggregation of the protein occurs, which resulted in a smear over the whole lane. The incubation with anti-*E. coli* rabbit serum demonstrated the purity of the antigen preparations.

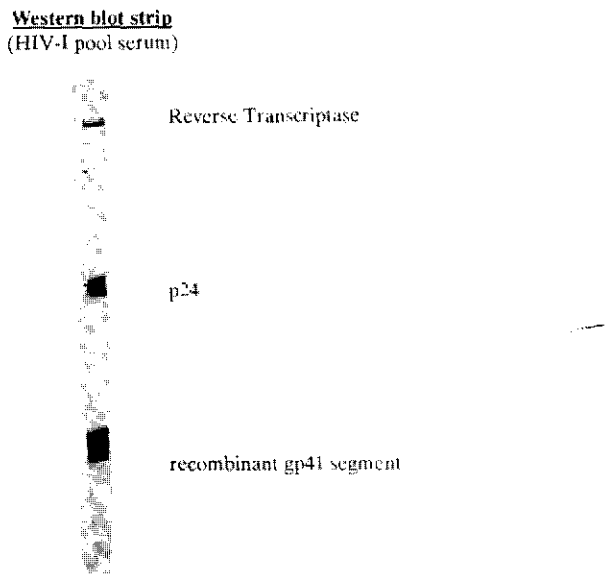

Fig. 9. Usage of the purified recombinant gp41 in a confirmation test based on the Western blot assay principle. A nitrocellulose strip was incubated with an HIV I-positive patient's serum and bound antibodies were made visible by immunodetection as described under Experimental. On the strip are indicated electrophoretically separated a recombinant-produced reverse transcriptase (p66), a p24 equivalent and the purified gp41 antigen.

applied to SDS-PAGE and transferred to nitrocellulose. After incubation with a patient's serum [diluted 1:100 with Tris-buffered saline containing Tween 20 (0.5 *M* NaCl, 20 m*M* Tris–HCl, pH 7.5, 8.01% Tween 20)], as shown for example in Fig. 9, and immunological detection with a rabbit serum containing anti-human IgG antibodies conjugated with peroxidase, a clear test result confirming the seropositivity was obtained. In comparison to the conventional Western blot, using whole virus lysates, a more sensitive detection of anti-gp41 antibodies could be demonstrated. This is due to the absence of glycosylation of the recombinant gp41 and, therefore, sharp banding in SDS-PAGE and unrestricted availability, whereas especially the anti-gp41 reactivity is hampered for several reasons in the conventional Western blot assay system [9,10], including the small amount of gp41 present in every individual preparation, glycosylation, which resulted in diffuse banding in SDS-PAGE, and under-representation compared with other viral components.

REFERENCES

1 M. Popovic, M. G. Sarngadharan, E. Read and R. C. Gallo, *Science (Washington, D.C.)*, 224 (1984) 497–500.
2 F. Barré-Sinoussi, J. C. Chermann, F. Rey, M. T. Nuyere, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vézinet-Brun, C. Rouzioux, W. Rozenbaum and L. Montagnier, *Science (Washington, D.C.)*, 220 (1983) 868 – 871.

3 R. C. Gallo, F. Wong-Staal, L. Montagnier, W. A. Haseltine and M. Yoshida, *Nature (London)*, 333 (1988) 504.
4 B. R. Starcich, B. H. Hahn, G. M. Shaw, P. D. McNeely, S. Modrow, H. Wolf, E. S. Parks, W. P. Parks, S. F. Josephs, R. C. Gallo and F. Wong-Staal, *Cell*, 45 (1987) 637–648.
5 J. J. G. Wang, S. Steel, R. Wiesniewolski and C. Y. Wang, *Proc. Natl. Acad. Sci. U.S.A.*, 83 (1986) 6159–6163.
6 J. W. Gnann, J. A. Nelson and M. A. B. Oldstone, *J. Virol.*, 61 (1987) 2639–2641.
7 W. R. Gallaher, J. M. Ball, R. F. Garry, M. C. Griffin and R. C. Montelaro, *AIDS Res. Hum. Retroviruses*, 5 (1989) 431–449.
8 J. W. Gnann, P. L. Schwimmbeck, J. A. Nelson, A. B. Truax and M. B. A. Oldstone, *J. Infect. Dis.*, 156 (1987) 261–267.
9 M. Motz, E. Soutschek, G. G. Frösner, L. Gürtler, M. Schall and H. Wolf, *Lancet*, ii (1987) 1093.
10 M. Motz, E. Soutschek, G. G. Frösner and H. Wolf, *AIDS-Forsch.*, 1 (1988) 10–17.
11 U. K. Laemmli, *Nature (London)*, 227 (1970) 680–685.
12 G. Myers, *Human Retroviruses and AIDS 1989*, Los Alamos National Laboratory, Los Alamos, 1989.
13 B. Krippl, B. Ferguson, M. Rosenberg and H. Westphal, *Proc. Natl. Acad. Sci. U.S.A.*, 81 (1984) 6988–6992.
14 F. A. O. Marston, P. A. Lowe, M. T. Doel, J. M. Schoemaker, S. White and S. Angal, *Biotechnology*, 2 (1984) 800–804.
15 G. C. Dubois, K. P. Samuel, C. A. Hanson, M. Zweig, S. D. Showalter and T. S. Papas, *AIDS Res. Hum. Retroviruses*, 4 (1988) 419–431.