The Cloning and Expression in Escherichia coli of Sequences Coding for p24, the Core Protein of Human Immunodeficiency Virus, and the Use of the Recombinant Protein in Characterizing a Panel of Monoclonal Antibodies against the Viral p24 Protein

By R. P. SPENCE,* W. M. JARVILL, R. B. FERNS, R. S. TEDDER and D. PARKER

1 Department of Molecular Biology, Wellcome Biotech, and 2 Wellcome Diagnostics R & D, Langley Court, Beckenham, Kent BR3 3BS and 3 Virology Section, Department of Medical Microbiology, Middlesex Hospital Medical School, Riding House Street, London W1P 7PN, U.K.

(Accepted 29 June 1989)

SUMMARY

The sequences encoding the p24 core protein of human immunodeficiency virus type 1 were identified in a cDNA library made from infected CEM cells. The nucleotide sequence of the DNA coding for p24 was shown to be very similar but not identical to the sequences of lymphadenopathy virus and human T-cell leukaemia virus type IIIb. These sequences were expressed in Escherichia coli at the amino terminus of β-galactosidase and used to screen a panel of monoclonal antibodies raised against virus-expressed p24. Regions containing the epitopes of five of the monoclonal antibodies were located using a series of amino- and carboxy-terminal deletion mutants of the recombinant p24 protein.

INTRODUCTION

The human immunodeficiency virus (HIV) is the causative agent of AIDS (Barré-Sinoussi et al., 1983; Gallo et al., 1984; Pinching & Weiss, 1986). HIV is a member of the retrovirus subfamily Lentivirinae (Montagnier et al., 1984); however, its genome arrangement differs from that of other retroviruses in that it encodes a collection of proteins involved in the control of viral replication, in addition to the gag, pol and env genes (Allan et al., 1985; Arya et al., 1985; Lee et al., 1986; Sodroski et al., 1985 a, b). The gag gene encodes p55, a protein which is cleaved by a product of the pol gene (Essex et al., 1985; Montagnier, 1985; Ratner et al., 1985) to produce the p24 protein which forms the viral core, p18, an envelope-associated protein and p15, a nucleic acid-binding protein. The last protein is cleaved further (Veronese et al., 1987). Recently Ferns et al., (1987) produced a panel of monoclonal antibodies (MAbs) against the gag proteins p24 and p18 of a British isolate of HIV named CBL-1 (Weiss et al., 1985). Using cross-competition assays they identified at least two independent epitopes on p24.

In this paper we describe the cloning of the sequences encoding the p18 and p24 proteins from the CBL-1 isolate, the expression of p24 sequences in Escherichia coli and the mapping of the epitopes of five of the p24 reactive MAbs. In the accompanying paper (Spence et al., 1989) we describe the expression of the p18 protein and the analysis of different HIV isolates using the anti-p18 MAbs.

METHODS

Preparation and screening of cDNA from CBL-1-infected CEM cells. The CBL-1 isolate of HIV-1 was grown in CEM cells as previously described by Weiss et al. (1985). RNA from the infected cells was extracted using the guanidinium–cesium chloride method of Chirgwin et al. (1979), and the poly(A)⁺ fraction was selected on oligo(dT)-cellulose (Maniatis et al., 1982). cDNA was made from oligo(dT)-primed RNA using avian myeloblastosis virus reverse transcriptase (Anglian Biotechnology) essentially by the method of Allison et al. (1981)
and cloned using EcoR1 linkers into lambda gt10 (Huynh et al., 1985). Approx. 3 x 10^4 recombinants were obtained after in vitro packaging (Northumbria Biologicals) and these were screened using an adaptation of the method of Maniatis et al. (1982). Filters were prehybridized in 50% deionized formamide, 5 x SSPE (740 mM-NaCl, 44 mM-NaH2PO4, 5 mM-EDTA pH 7.4), 5 x Denhardt’s solution [1% Ficol, 1% polyvinylpyrrolidone, 1% bovine serum albumin (BSA)] and 200 μg/ml salmon sperm DNA at 37 °C for 4 h. Hybridization was performed in the above solution with 0.1% sodium pyrophosphate and the probe for at least 18 h at 37 °C. The probe consisted of a 737 bp Sacl–PstI fragment (nucleotide numbers 224 to 961) from pB2, a lymphadenopathyvirus (LAV)-like isolate (Benn et al., 1985) (kindly supplied by Dr M. Martin, Bethesda, Md., U.S.A.) which was labelled by random priming (Feinberg & Vogelstein, 1983, 1984) to a specific activity of 1 x 10^8 to 2 x 10^8 c.p.m./μg. After three rounds of screening, single positive clones were subcloned into the EcoR1 site of pUC19 and transformed into E. coli DH1.

Expression of p24 as a fusion protein in E. coli. A clone, designated pDM608, containing the coding regions for both p18 and p24 was selected by sequence analysis of its 5' and 3' ends. From this clone a 710 bp PvuII–NcoI fragment (nucleotide numbers 691 to 1401) encoding most of the p24 protein was gel-purified and the NeII 3' recessed ends were filled in with Klenow DNA polymerase before ligation into the Smal site of the expression vector pXY460, a derivative of pXY410 (Winther et al., 1986). After transformation of the ligation mixture into TG1 cells, a derivative of E. coli JM101 (Gibson, 1984), recombinants expressing β-galactosidase were selected by plating out on LB agar containing 100 μg/ml ampicillin and 40 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside, followed by a further selection of those plasmids containing the 710 bp fragment by small-scale DNA preparation (Maniatis et al., 1982) and restriction analysis. Fusion proteins were induced by diluting an overnight liquid culture fivefold in fresh medium containing 60 μg/ml isopropyl-β-D-thiogalactoside and 100 μg/ml ampicillin followed by incubation at 37 °C for 3 h. The cells were harvested by centrifugation at 5000 r.p.m., washed in TE buffer (25 mM-Tris-HCl pH 8.0, 1 mM-EDTA), three times in high salt RIPA buffer (1 M-NaCl), and finally three times in RIPA buffer. After denaturation by boiling in sample buffer (3% SDS, 1.6 M-urea, 33 mM-mercaptoethanol) for 2 min the antibody complex was subjected to SDS–PAGE and the proteins were immunoblotted.

Western blot analysis of expressed proteins. Bacterial cell lysates were run on 10% or 12% polyacrylamide gels (Laemmli, 1970), transferred to nitrocellulose filters and probed with antibodies by a modification of the method of Matlashewski et al. (1986). The filters were blocked using 10% foetal calf serum at 37 °C for 1 h. The MAbS were diluted 1:50 in phosphate-buffered saline (PBS) containing 0.3% BSA and 0.1% Tween 20 and then incubated with the filters for 2 h using approx. 2 ml per 50 cm². Blots were washed four times in PBS/Tween then incubated with goat anti-mouse-biotin conjugate (Amersham) followed by streptavidin-peroxidase conjugate (Amersham). Blots were developed in the presence of 0.1% hydrogen peroxide and 1-5 mg/ml Hanker–Yates reagent (Polysciences).

Western immune precipitation assays. Immune precipitation assays were done essentially by the method of Androphy et al. (1987). The bacterial cell lysates were incubated with constant agitation at 4 °C with 500 μl of 20% Protein A-Sepharose for 30 min. After centrifugation, the supernatant was incubated with the appropriate antibody for 45 min at 4 °C. Most of the mouse MAbs were of the IgG1 subclass, which bind Protein A weakly; therefore a goat anti-mouse antibody was added 30 min before the addition of 15 μl Protein A-Sepharose followed by incubation for 1 h at as above. Protein A-Sepharose was then pelleted and washed three times in RIPA buffer (50 mM-Tris–HCl pH 7.4, 150 mM-NaCl, 1 mM-EDTA), three times in high salt RIPA buffer (1 M-NaCl), and finally three times in RIPA buffer. After denaturation by boiling in sample buffer (3% SDS, 1.6 M-urea, 33 mM-mercaptoethanol) for 2 min the antibody complex was subjected to SDS–PAGE and the proteins were immunoblotted.

Preparation of deletion mutants of p24. Deletions were carried out on the pDM614 construct using exonuclease Bal31, essentially by the method of Maniatis et al. (1982), after digestion with either EcoR1 or BamHI. The Bal31-digested vector was digested with BamHI (if previously digested with EcoR1) or EcoR1 (if previously digested with BamHI) and the 5' or 3' deleted fragments were purified from agarose gels. These samples were then religated into EcoR1–Smal-digested pXY460 and transformants expressing fusion proteins were selected.

Sequence analysis. This was carried out using Sequenase (United States Biochemical) in a modification (Tabor & Richardson, 1987) of the dideoxynucleotide chain terminator procedure of Sanger et al. (1987). DNA cloned into pUC19 was sequenced using forward and reverse M13 sequencing primers (Pharmacia). Oligonucleotides were synthesized complementary to the sequences 5' and 3' to and including the polycloning site of pXY460. For the p24 expression fragment, a series of oligonucleotides complementary to both strands of this fragment were synthesized.

RESULTS

Identification of a clone containing the p24 coding region and the expression of p24 in E. coli

Recombinants (10⁵) were screened for clones containing the p24 coding region using the 737 bp SacI–PstI fragment (nucleotide numbers 224 to 961) which covered the 5' region of the
**Recombinant HIV p24**

Fig. 1. The shaded bar represents the gag polyprotein p55 divided into its cleavage products p18, p24 and p15, the unshaded bar represents the sequences proteolytically removed when p55 is processed. Part (a) represents the HIV genome, (b) shows the position of the 737 bp SacI-PstI probe used to identify gag coding sequences, (c) indicates the limits of the HIV-1 fragment in pDM608 and (d) shows the position of the restriction sites used to construct pDM614.

**gag open reading frame (ORF)** (Fig. 1). Eighteen positive clones were identified, one of which contained a fragment of approx. 1400 bp which was subcloned into pUC19 to produce the construct pDM608. Sequencing showed that the complete p18 and p24 coding regions were contained within this clone. Using the sequence and nucleotide numbering derived for LAV (Wain-Hobson et al., 1985), used throughout this paper, we cleaved at a PvuII site (position 691) and an NciI site (position 1401) to generate a 710 bp fragment which consisted of 38 nucleotides from the 3'end of p18 plus the p24 sequence except for 21 bp at the extreme 3' end. This fragment was cloned into Smal-cut pXY460, which donated the initiating methionine codon, to produce the expression construct pDM614 (Fig. 1). The predicted fusion protein would contain 12 amino acids from the carboxy terminus of p18, while at the carboxy terminus of the p24 fragment seven amino acids would be missing. Induced *E. coli* containing the pDM614 construct expressed a fusion protein of approximately 142K, the amino terminus of which represented 26K of gag products with 116K of active β-galactosidase at the carboxy terminus (Fig. 2).

**Sequence of CBL-1 p24**

Using a series of 15-mer oligonucleotides, the sequence of the insert of pDM614 construct was obtained and found to be very similar to ones in both LAV (Wain-Hobson et al., 1985) and human T-cell leukaemia virus type IIIb (HTLV-IIIb) (Ratner et al., 1985). Variations were seen at position 839 (CBL-1, G; HTLV-IIIb, A), position 980 (CBL-1, G; HTLV-IIIb, A), position 1212 (CBL-1, T; LAV, C), position 1355 (CBL-1, A; HTLV-IIIib, G) and position 1394 (CBL-1,
Fig. 2. A Coomassie Brilliant Blue-stained 10% SDS-polyacrylamide gel of induced protein lysates. Lane 1, Mr markers; lane 2, β-galactosidase expressed from pXY461, a derivative of pXY460 in which the sequence of the polycloning site is changed to allow expression of pure β-galactosidase; lane 3, E. coli lysate of the induced product of plasmid pDM614.

Table 1. Reactivity of MAbs against the fusion protein expressed by pDM614

<table>
<thead>
<tr>
<th>MAb</th>
<th>Virus expressed protein, Western blot*</th>
<th>Recombinant protein, Western blot</th>
<th>Recombinant protein, WIPA†</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D3</td>
<td>p55, p24</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1E8G2</td>
<td>p55, p24</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DF9C3</td>
<td>p55, p24 (weak)</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>CD12B4</td>
<td>p55, p24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1D8F6</td>
<td>p55, p24 (weak)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EH12E1</td>
<td>p55, p24</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>EB1A9</td>
<td>p55, p24</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4H2B1</td>
<td>p55, p18</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>1D9</td>
<td>p18</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

* Data from Ferns et al. (1987).
† Western blot detection immunoprecipitation assay.

G; HTLV-IIIb, A). In each case the CBL-1 sequence is either the same as LAV or HTLV-IIIb and none of the differences alters the amino acid composition of the CBL-1 p24, as compared to LAV and HTLV-IIIb.

Characterization of the anti-HIV-1 MAb using the pDM614 fusion protein

A panel of seven MAbs against p24/p55 previously described by Ferns et al. (1987) were screened against the pDM614 fusion protein in both Western blot and immunoprecipitation assays (Table 1). Only two of the seven p24/p55-specific MAbs, DF9C3 and 1D8F6, failed to
Table 2. The deletion mutants of pDM614

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Position of deletion</th>
<th>No. of base pairs deleted</th>
<th>No. of p24 coding amino acids deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDM614</td>
<td>Carboxy</td>
<td>0</td>
<td>7*</td>
</tr>
<tr>
<td>pDM640a</td>
<td>Carboxy</td>
<td>134</td>
<td>52</td>
</tr>
<tr>
<td>pDM640b</td>
<td>Carboxy</td>
<td>313</td>
<td>112</td>
</tr>
<tr>
<td>pDM640c</td>
<td>Carboxy</td>
<td>487</td>
<td>170</td>
</tr>
<tr>
<td>pDM640d</td>
<td>Carboxy</td>
<td>496</td>
<td>173</td>
</tr>
<tr>
<td>pDM640e</td>
<td>Carboxy</td>
<td>547</td>
<td>190</td>
</tr>
<tr>
<td>pDM644a</td>
<td>Amino</td>
<td>99</td>
<td>33</td>
</tr>
<tr>
<td>pDM644b</td>
<td>Amino</td>
<td>128</td>
<td>43</td>
</tr>
<tr>
<td>pDM644c</td>
<td>Amino</td>
<td>262</td>
<td>88</td>
</tr>
<tr>
<td>pDM644d</td>
<td>Amino</td>
<td>292</td>
<td>98</td>
</tr>
<tr>
<td>pDM644e</td>
<td>Amino</td>
<td>343</td>
<td>115</td>
</tr>
<tr>
<td>pDM644f</td>
<td>Amino</td>
<td>352</td>
<td>118</td>
</tr>
<tr>
<td>pDM644g</td>
<td>Amino</td>
<td>436</td>
<td>146</td>
</tr>
</tbody>
</table>

* The pDM614 plasmid does not express the carboxy-terminal seven amino acids so the sizes of the carboxy amino acid deletions have been adjusted accordingly.

react with the recombinant protein on Western blots. Their failure to react may be due to either low avidity for their target epitopes (both gave weak signals on Western blots against viral proteins) (Ferns et al., 1987), or the lack of a suitable epitope on the recombinant protein. As all the anti-p24 MAbs also reacted with p55, the 12 p18 amino acids at the amino terminus of the fusion protein should not inhibit their binding. However due to the position of the 3' cloning site of the expression fragment, the pDM614 construct cannot express the seven carboxy terminus amino acids of p24. It is therefore possible that their lack of reactivity is due to the location of the epitopes of these MAbs at the carboxy terminus of p24. The immunoprecipitation assays for 3D3, 1E8G2 and EB1A9 supported the Western blot data presented here and the immunofluorescence data of Ferns et al. (1987). However, CD12B4 and EH12E1 failed to react in this assay although they reacted on Western blots. It is possible that this is due to the lower sensitivity of the immunoprecipitation assay as compared with Western blots.

Location of the regions containing the epitopes of the reactive MAbs using deletion mutants of pDM614

The production of HIV-1 p24 immunoreactive proteins in E. coli can be a valuable tool for the analysis of epitopes of MAbs since the manipulation of the DNA encoding these regions is relatively easy.

A series of mutants were constructed from pDM614 by progressive Bal 31 deletions from either the 5' or 3' end of the 710 bp fragment. The end of each deletion was sequenced to determine the exact size and to ensure that the reading frame was correct. Table 2 lists the deletion constructs and Fig. 3 shows the pattern of reactivity of their products with the p24-specific MAbs. It is worth noting that the antibodies react with a series of bands below the 142K pDM614 fusion protein and the deleted fusion proteins. Control antibodies fail to react with any of these bands (Fig. 3f) so we conclude that they must be carboxy proteolytic cleavage products of the fusion protein since amino proteolytic cleavage products would not react with the anti-p24 MAbs. The MAbs 3D3, EH12E1 and EB1A9 all react with the 43 amino acid (128 bp, 5') amino deletion product (pDM644b) but fail to react with any larger deletions, 88 amino acids (262 bp, 5') and greater (pDM644c to pDM644g). 3D3 reacts with the 173 amino acids (496 bp, 3') carboxy deletion product (pDM640d) but not with the 190 amino acid (547 bp, 3') deletion product (pDM640e). This would delimit the epitope for this MAb to 17 amino acids (within a 51 bp coding region, positions 854 to 905). The MAb EB1A9 did not react after the deletion of a region of 60 amino acids (134 to 313 bp, 3') (see pDM640a and pDM640b, Fig. 3) indicating that
Fig. 3. Western blots of 10% SDS-polyacrylamide gels of protein lysates produced from the deletion constructs. Lanes 1 to 5, pDM640a to e respectively; lanes 6 to 12, pDM644a to g respectively. MAbs used were (a) EB1A9, (b) 1E8G2, (c) CD12B4, (d) EH12E1, (e) 3D3 and (f) G7H9 (negative control MAb).

a site required for binding falls within this region. However, a site required for its binding also falls within a 45 amino acid region at the amino end of the protein (Fig. 4). This would suggest that both these sequences, and probably the region between them, are required for the binding of this MAb. If this is so, then this epitope must be highly stable to be reactive after SDS-PAGE and Western blotting, although some proteins can become renatured after binding to nitrocellulose (Towbin & Gordon, 1984). The carboxy and amino deletion data locate the epitope of EH12E1 in the same region as that recognized by 3D3. The epitopes for 1E8G2 and CD12B4 are nearer the carboxy terminus of p24, that for CD12B4 falling within a 52 amino acid region at the carboxy end of p24 (Fig. 4).

DISCUSSION

E. coli expressing CBL-1 p24 sequences produced a protein which was reactive with MAbs raised against the viral protein. The advantages of using recombinant proteins for the analysis of immune reactions to HIV are, first, that one does not have to handle infectious material, and secondly that the protein sequence can be easily manipulated, allowing a direct approach to epitope mapping.
Recombinant HIV p24

The cross-competition analysis of Ferns et al. (1987) indicated that the MAbs EH12E1 and CD12B4 recognized closely related, although probably discrete, epitopes. By deletion analysis we have shown that the epitope for CD12B4 lies near the carboxy terminus of p24 while the epitope for EH12E1 falls more within the body of the protein, towards the amino terminus. Our analysis cannot place the epitopes any more precisely than shown in Fig. 4, although further analysis with peptides spanning this region probably would. However, it is worth noting that epitope mapping using peptides alone would not have identified the conformational epitope recognized by EB1A9. It has recently been reported that a series of peptides spanning the whole of the p24 protein failed to allow identification of the epitopes for some MAbs (Helling et al., 1988).

Ferns et al. (1987) used the panel of MAbs to distinguish between different HIV isolates. ARV-2, an isolate from the U.S.A. (Sanchez-Pescador et al., 1985), did not bind the MAb CD12B4 which binds to p24 near its carboxy terminus. Examination of the protein sequence of the p24 product of ARV-2 shows only one amino acid change (an asparagine residue for a glutamine residue) within the 52 amino acid epitope region for CD12B4. It is interesting to observe that the LAVEU isolate from Zaire contains the same amino acid substitution as ARV-2 (Alizon et al., 1986). A peptide spanning this region would confirm whether this is the epitope for CD12B4. Computer predictions of the antigenic sites of gag p24 agree closely with the regions located using our deletion analysis. It is of particular interest that a suggested six amino acid epitope (at amino acid position 158) corresponds precisely with the region we would have
predicted as the target for CD12B4 (Argos, 1989). As more sequence data become available and more isolates are analysed immunologically, it may be possible to gather together a panel of MAbs capable of distinguishing between HIV isolates by changes in their gag as well as their env proteins.

The authors would like to thank Dr Peter Highfield for his advice during the course of this work and Jo Keenan for the production of the figures.

REFERENCES


(Received 17 October 1988)