Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 2449-2453, April 1988
Biochemistry

An 11-kDa form of human immunodeficiency virus protease expressed in *Escherichia coli* is sufficient for enzymatic activity

(retrovirus maturation/*pol* and *gag* open reading frames)

MARY C. GRAVES*†, J. JOHANN LIM*, EDGAR P. HEIMER‡, AND RICHARD A. KRAMER*

Departments of *Molecular Genetics and ‡Peptide Research, Roche Research Center, Hoffmann-La Roche, Nutley, NJ 07110

Communicated by Herbert Weissbach, December 3, 1987

ABSTRACT **In order to define the protease domain of human immunodeficiency virus 1, various regions of the *pol* open reading frame were cloned and expressed in *Escherichia coli*. Antiserum directed against the conserved retroviral protease active site was used to identify pol precursor and processed species containing the presumed protease domain. The smallest product that accumulates is about 11 kDa as measured by NaDodSO$_4$/PAGE. This size agrees with that predicted from the presence in this region of two Phe-Pro sequences, which is one of the cleavage sites recognized by HIV protease. DNA encoding only the predicted 11-kDa protein was cloned, bypassing the need for autoprocessing, and the protein was expressed to a high level in *E. coli*. This form is active as demonstrated by its ability to specifically cleave protease-deficient pol protein *in vivo* in *E. coli*. Extracts of *E. coli* containing the 11-kDa protease also process human immunodeficiency virus gag substrates *in vitro*. These results demonstrate that the 11-kDa protease is sufficient for enzymatic activity and are consistent with a major role for this form in virus maturation.**

In order to combat human immunodeficiency virus (HIV), the causative agent of acquired immunodeficiency syndrome (AIDS), we must understand the synthesis, assembly, and maturation pathways of retroviruses in general and of HIV in particular. As in all retroviruses, HIV replication requires the synthesis and processing of polyproteins encoded by the *gag* and *pol* open reading frames (1). The HIV gag precursor (p56, a 56-kDa protein) gives rise to the major structural proteins p17, p24, p9, and p7 (2–6). The *pol* region is expressed initially as a gag-pol fusion protein (p170), which most likely results from an infrequent ribosomal frameshift (7) as in the case of Rous sarcoma virus (8) and yeast Ty elements (9). The gag-pol fusion protein is processed to yield the enzymatic activities of protease (PR), reverse transcriptase (RT; RNA-directed DNA polymerase), and endonuclease/integrase (IN). In Rous sarcoma virus, reliance on this rare frameshifting event achieves the molar balance of 20–50 to 1 of structural gag protein to enzymatic pol protein (8).

Comparison of the assembly pathway in mammalian type C retroviruses reveals common events that serve as a model for HIV assembly (1, 10), including the crucial role of viral PR in the maturation pathway. During infection, precursor gag and gag-pol molecules accumulate at the cell membrane, most likely targeted there by the myristic acid moiety that is covalently linked to the precursor amino termini (11). Precursors assemble with viral RNA into viral particles, which bud from the envelope protein (env)-studded cellular membrane, and immature particles are released. Appearance of mature particles, visualized by electron microscopy as containing electron-dense cores, is correlated with the PR-dependent processing of gag and pol proteins in murine leukosis virus (1, 12) and yeast Ty particles (13). In retroviruses, this morphological change correlates with the conversion of noninfectious viral particles to infectious ones (14). Some unknown stimulus presumably activates the PR domain of the gag-pol precursor concomitant with or after budding, leading to the production of infectious, mature virus.

This maturation pathway, then, raises several questions about the form of PR that is the active species. One possibility is that the PR domain frees itself from the gag-pol precursor to give a small, more freely diffusible enzyme that is better able to locate and process the more numerous gag proteins. Determination of amino-terminal sequences of HIV gag p24 (2–5) and RT (15, 16) proteins and knowledge of the surrounding sequences have revealed a common sequence [Asn-(Phe or Tyr)-Pro-Ile] at the cleavage site. A similar sequence (Asn-Phe-Pro-Gln-Ile) occurs within the PR domain. If cleavage occurs here as well, then a PR species of ≈11 kDa could be produced.

In order to address these questions, we sought to define the limits of the PR domain within *pol* and to study its catalytic activity. Since *Escherichia coli* successfully served as a host for production of active Rous sarcoma virus PR (17), it was used as the host for HIV PR production. Several sections of the HIV *pol* gene were cloned and expressed in *E. coli*, and we report here the intra- and intermolecular proteolytic activity of the resulting HIV proteins.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Plasmids, and Growth Conditions. *E. coli* K-12 strain MC1061 transformed with plasmid pRK248cIts was used for the work described. pRK248cIts is compatible with pBR322 derivatives and encodes a temperature-sensitive λ repressor (18). The pEV-vrf plasmids used below are pBR322 derivatives containing the bacteriophage λ P_L promoter and translation initiation signals (19). Transformants were grown in supplemented M9 medium and heat-induced at 40°C (19) for the times indicated. Cells were stored at −20°C at a concentration of 20 OD_{600} units per ml in 10 mM Tris·HCl, pH 8.0/10% sucrose.

Plasmid Constructions. DNA preparations, transformations, and analyses were performed as described (20). Restriction enzymes were from Boehringer Mannheim, and Klenow fragment of DNA polymerase I, bacteriophage T4 DNA ligase, and T4 polynucleotide kinase were from New England Biolabs. Plasmid pPol3 encodes most of the *pol* open reading frame (Fig. 1) and was constructed by ligating the 2433-base-pair (bp) *Bgl* II–*Xmn* I fragment of the λHXB-3 HIV genome (ref. 21; nucleotides 2094–4526) to the 2570-bp *Pvu* II–*Bam*HI DNA of the pEV-vrf3 expression

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.

Abbreviations: HIV, human immunodeficiency virus; PR, protease; RT, reverse transcriptase; IN, endonuclease/integrase.

†To whom reprint requests should be addressed.

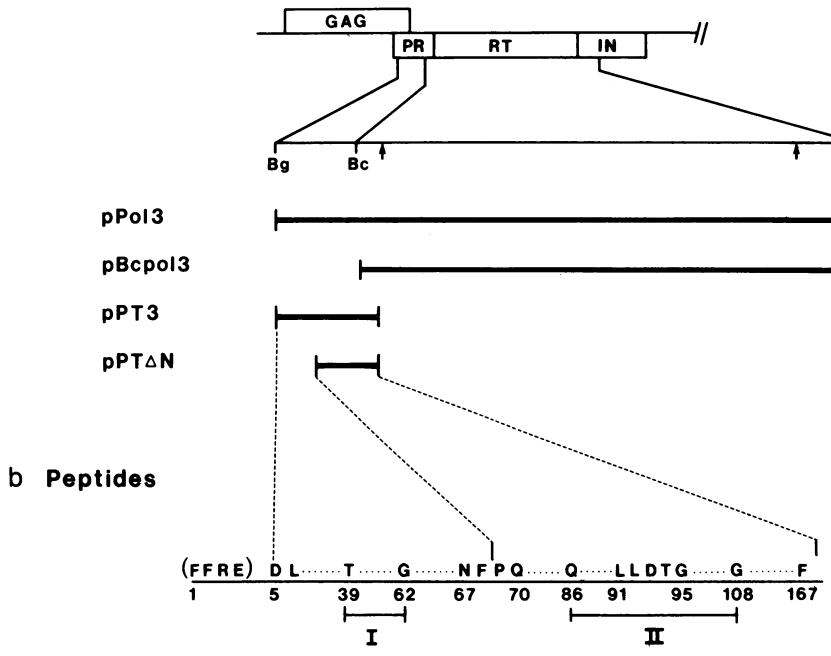

FIG. 1. Construction of PR expression plasmids and peptides. (*a*) Top map represents a section of the λHXB-3 HIV genome. Part of the *pol* region (encoding PR, RT, and IN) is expanded to show relevant restriction sites (Bg, *Bgl* II; Bc, *Bcl* I; X, *Xmn* I) and previously mapped sites of protease cleavage (arrows) that generate the amino termini of RT and IN. A cryptic cleavage site in RT, not indicated, generates two RT forms of about 66 and 51 kDa with identical amino termini (15, 16). Bold lines indicate the *pol* region inserted into expression vectors. (*b*) Amino acid sequence of HIV PR region. The FFRE sequence is encoded in the intact *pol* open reading frame but is absent in the plasmids above. The conserved aspartyl active site residues are shown by L91 to G95. P69 is the proposed amino terminus and F167 the carboxyl terminus of the 11-kDa PR form, p11. Peptides defined by sequences I and II were used to raise antisera as described in *Experimental Procedures*. One-letter amino acid symbols: F, Phe; R, Arg; E, Glu; D, Asp; L, Leu; T, Thr; G, Gly; N, Asn; P, Pro; Q, Gln.

plasmid. The *Bgl* II/*Bam*HI hybrid site thus generated is no longer recognized by either restriction enzyme, and the plasmid retains a single *Bgl* II site 5′ to the P_L promoter. pPT3, a derivative of pPol3, contains *pol* information upstream of the RT sequence and was made in the following manner. pPol3 was treated with *Bgl* II and *Dra* I restriction endonucleases to produce a 780-bp fragment containing the P_L promoter and all but the final 7 nucleotides of the PR coding region. Complementary oligonucleotides were synthesized to regenerate the missing nucleotides, provide a translational termination codon, and add a *Sal* I restriction site. This 785-bp DNA was then ligated to the 3711-bp *Sal* I–*Bgl* II fragment of pEV-vrf1. pPT3 was modified to create pPTΔN, which encodes only an 11-kDa protein that includes the consensus aspartyl protease active site. It was constructed by the oligonucleotide-directed deletion of pPT3 double-stranded DNA (22) to delete *pol* sequences between Phe-4 and Pro-69 (Fig. 1). pBcpol3 was constructed from pPol3 and contains most of the *pol* region except for the sequence encoding the PR active site. It was produced by digesting pPol3 with *Bcl* I and *Pst* I and ligating the 3608-bp fragment to the 1033-bp *Pst* I–*Bam*HI DNA of pEV-vrf3. pBcpol3 was modified to produce pBPT11, which includes information for the 11-kDa (p11) form of PR. This was accomplished by digesting pBcpol3 with *Bgl* II, treating with *Bal*-31 exonuclease (IBI) to create blunt ends, and digesting with *Pst* I. The resulting 3893-bp fragment was ligated to the 2748-bp *Pst* I–*Pvu* II DNA of pPTΔN, which contains the entire p11 transcriptional unit.

Antibody Production. Two peptides corresponding to two hydrophilic regions of the probable PR domain were synthesized using a Vega model 1000 peptide synthesizer and purified by HPLC. Peptide I (Thr-Arg-Arg-Glu-Leu-Gln-Val-Trp-Gly-Arg-Asp-Asn-Asn-Ser-Pro-Ser-Glu-Ala-Gly-Ala-Asp-Arg-Gln-Gly) represents amino acids 39–62, and peptide II (Gln-Leu-Lys-Glu-Ala-Leu-Leu-Asp-Thr-Gly-Ala-Asp-Asp-Thr-Val-Leu-Glu-Glu-Met-Ser-Leu-Pro-Gly) represents amino acids 86–108 of the *pol* open reading frame (Fig. 1*b*). Peptide II includes the consensus sequence found in aspartyl protease active sites. Each peptide was conjugated to a carrier protein, keyhole limpet hemocyanin (23). Antisera were raised in rabbits and used without purification. Polyclonal rabbit serum against a RT precursor was kindly provided by J. Mous and S. LeGrice and gag p24 monoclonal antibody was obtained from J. J. Burckhardt, all of F. Hoffmann–La Roche (Basel).

Gel Analysis of Proteins. pol protein synthesis and processing in *E. coli* were examined by $NaDodSO_4$/PAGE and subsequent immunoblotting. Cell suspensions were mixed with an equal volume of sample buffer (24) and treated at 100°C for 3 min. The indicated amounts were subjected to $NaDodSO_4$/PAGE (24) in 10% or 15% gels as noted in the figure legends. Proteins were electrophoretically transferred to nitrocellulose (Schleicher & Schuell) and probed with the indicated antibody. In some cases, rabbit serum was pretreated to remove endogenous anti-*E. coli* antibodies by incubating the serum with a 20-fold excess of an *E. coli* lysate (containing 10 mg of protein per ml and obtained from Bio-Rad) for 30 min at 4°C and then diluting prior to use. Second antibody was the appropriate IgG coupled to horseradish peroxidase (Boehringer Mannheim). Development was with 4-chloro-1-naphthol (Sigma) and H_2O_2.

***In Vitro* Assay for PR Activity.** Bacterial cultures containing pRK248cIts with or without pPTΔN were induced for 80 min and stored as described above. A 200-μl aliquot of each was brought to 1 mM dithiothreitol, 1 mM EDTA, 50 mM NaCl, and 2.3 mM phenylmethylsulfonyl fluoride in a final volume of 220 μl. The cells were broken by five 20-sec sonication bursts, and cellular debris was removed by centrifugation at 15,000 × *g* for 15 min in a microcentrifuge at 4°C. HIV gag precursor was produced in *E. coli* as described (25). After purification, gag consists primarily of three species as observed by $NaDodSO_4$/PAGE: full-length p56, p53, and p44. The amino termini of p56 and p53 are identical (25). Upon immunoblotting with monoclonal antibody to virally produced gag p24, many more species are apparent (Fig. 4, lane A). The antibody reacts most strongly with a truncated p26, which is barely visible on Coomassie blue-stained gels. Poor reactivity of the antibody to the major species is most likely due to masking of the p24 epitope in the larger proteins. Protease digestions were carried out with 12 μg of gag protein and 50 μg of crude extract protein in 100 mM sodium phosphate, pH 6.0/5 mM dithiothreitol for 2 hr at 37°C, fractionated by $NaDodSO_4$/15% PAGE, and immunoblotted as above with monoclonal antibody recognizing gag p24.

RESULTS

Several lines of evidence place the coding region of the PR domain in *pol* (Fig. 1) just 5′ of the RT sequence. First, a region identical to the active site of known cellular, retroviral, yeast Ty, and hepatitis B virus aspartyl protease (26–28) is encoded just upstream of the *Bcl* I site (Fig. 1*b*). Also, several groups have demonstrated that PR-dependent processing requires information between the *Bgl* II site and the RT coding region (29–32). Our initial plasmid, pPol3, includes the coding region of the PR domain and begins *pol* sequences at the *Bgl* II site just four codons into the *pol* open reading frame; four different codons are contributed by the vector 5′ to the *pol* sequences. The primary product includes PR, RT, and part of IN in a p96 precursor.

Additional plasmids were made to help define the limits of the PR domain (Fig. 1*a*). In pBcpol, the region encoding the consensus aspartyl protease active site has been deleted, thereby retaining pol sequences that presumably are incapable of autoproteolysis (see below). pPT3 contains a translational stop codon following the codon for Phe-167, the site of cleavage that produces Pro-168 as the RT amino terminus, and thus encodes an 18-kDa protein. Within this 18-kDa region is another sequence similar to the common cleavage site mentioned above (^{67}Asn-Phe-Pro-Gln-Ile^{71}). If PR recognizes this site, then the conserved aspartyl protease region would remain on a protein with a predicted M_r = 10,780 (Pro-69 to Phe-167; Fig. 1*b*). Another plasmid, pPTΔN, was constructed that encodes only this 10.8-kDa peptide preceded by an amino-terminal methionine.

Analysis of Processed PR Species. Using the set of PR expression plasmids, we determined the size of precursor and processed pol species and directly assayed for PR activity. In order to identify PR species in *E. coli* extracts, rabbit antisera were raised to two synthetic peptides (Fig. 1*b*) located on either side of the putative cleavage site (Phe-69/Pro-70). Peptide II antiserum (serum II) is predicted to recognize precursors, intermediates, and the limit digestion product that contain the conserved PR active site. Serum I should recognize any protein that contains the region just preceding the putative cleavage site at Phe-69.

E. coli harboring the PR plasmids were heat-treated for various times to induce transcription. PR precursor–product relationships were analyzed by immunoblotting of *E. coli* cellular proteins resolved by $NaDodSO_4$/PAGE. As expected, p96 and p18 precursors and several smaller products crossreact with the two sera (Fig. 2). The rabbit sera also react with many *E. coli* proteins, resulting in a background pattern in every lane. The most obvious product is the abundant 10- to 11-kDa protein that is recognized by serum II but not by serum I (p11; Fig. 2*a*, arrow 5). Note that the immunoblot was purposely overdeveloped with respect to this species in order to visualize intermediates. Large amounts of p11 accumulate regardless of whether the precursor was p96 (lanes A–E), p18 (lanes F–J), or just p11 (lanes K–O). p11 takes longer to appear from the p96 pol precursors (80 min; Fig. 2*a*, lane D) than in the other two cases (20 min, lanes H and M), as one would expect for a precursor–product relationship. It is produced to the highest level from pPTΔN plasmid and represents a few percent of the total cellular protein as judged by Coomassie blue staining of polyacrylamide gels (data not shown). The stable p11 species in the three strains appear to comigrate (this is more obvious in normally developed immunoblots).

Several other intermediates are observed in the pPol3 stain but are labile (Fig. 2*a*, lanes B–E, arrows). Note the rapid disappearance of proteins containing peptide I (Fig. 2*b*). This observation suggests that cleavage between peptide I and peptide II occurs rapidly or that proteins containing peptide I are unstable in *E. coli* (see *Discussion*).

Activity of p11 PR. Although the smallest stable protein containing the putative PR active site is 11 kDa, is it the active form or only a remnant lacking activity? To determine whether p11 is active on viral substrates, we examined the PR activity in the pPTΔN strain on pol and gag substrates both *in vivo* and *in vitro*. For the *in vivo* analysis, the substrate was a portion of the pol molecule defined by the *Bcl* I–*Xmn* I fragment (Fig. 1, pBcpol3). A mutation at this

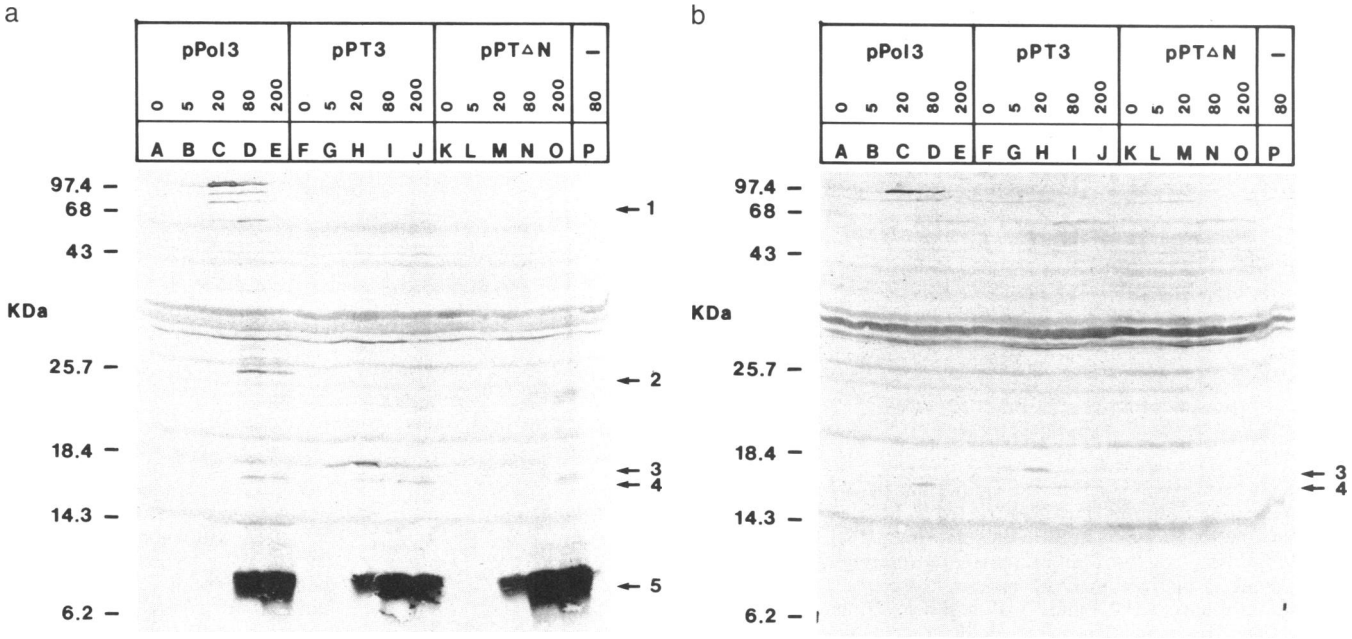

FIG. 2. Identification of the 11-kDa form of PR. pPol3 (lanes A–E), pPT3 (lanes F–J), and pPTΔN (lanes K–O) strains were induced for 0, 5, 20, 80, or 200 min as indicated above each lane. Aliquots (25 μl) of *E. coli* carrying one of the indicated plasmids were subjected to $NaDodSO_4$/15% PAGE and immunoblotted. Lane P: *E. coli* lacking a PR expression plasmid. The primary antibody used was 1:5000 dilution of serum II (*a*) or a 1:2000 dilution of serum I (*b*). Arrows indicate processed intermediates and products of PR as described in the text. Mobilities of molecular size standards are noted.

Bcl I site abolishes PR activity (7). The consensus PR active site is not contained in this region, but several processing sites are (PR/RT and RT/IN junctions and the unmapped additional site in RT; Fig. 1). Expression of the *pol* region of pBcpol3 over time was examined by using antibodies to HIV RT (Fig. 3*a*). By 15 min (lane B) a predominant 83-kDa precursor and minor fragments are evident. The proteins are relatively stable as reflected by only minor changes until a generalized proteolysis is observed at 120 min (lane E). When the pBcpol3 plasmid is modified (pBPT11, see *Experimental Procedures*) to include the independent PR transcriptional unit from pPTΔN, the p11 PR is expressed as a separate protein (Fig. 3*b*). In *E. coli* containing pBPT11, the fate of the p83 precursor is quite different: the precursor is quickly processed (<15 min), and major species at 66 and 51 kDa appear to be stable. These sizes are consistent with those for authentic active RT previously shown to be produced in HIV virions (15, 16), *E. coli* (29–31), *Bacillus subtilis* (32), and yeast (33). In fact, when expressed in this fashion, all RT products are indistinguishable from those made in pPol3 strains (Fig. 3*a*, lane L).

The activity of p11 PR also was detected *in vitro* by incubation of *E. coli* crude extracts with HIV gag substrate (Fig. 4). gag purified from *E. coli* consists primarily of p56 precursor and two truncated forms, p53 and p44 (ref. 25 and *Experimental Procedures*). They carry gag p24 epitopes but react poorly with a p24 monoclonal antibody (lane A). A family of minor species are also visualized by reaction with p24 monoclonal antibody; the smaller gag species may arise from cryptic premature termination and/or internal reinitiation events during protein synthesis or by proteolysis during purification. When control extracts of heat-induced *E. coli* are incubated with gag, only minor degradation is observed (Fig. 4, lane B). On the other hand, extracts containing PR specifically process gag (arrows), most notably the full- and near-full-length precursors. The full-length p24 (top arrow) and other p24 truncated proteins are stable under these conditions up to 20 hr at 37°C, as are the vast majority of *E. coli* proteins (data not shown), again attesting to the specificity of PR and the general lack of nonspecific proteolysis by bacterial proteases. Similar observations on the stability of avian retroviral precursors synthesized in *E. coli* were made by Alexander *et al.* (34).

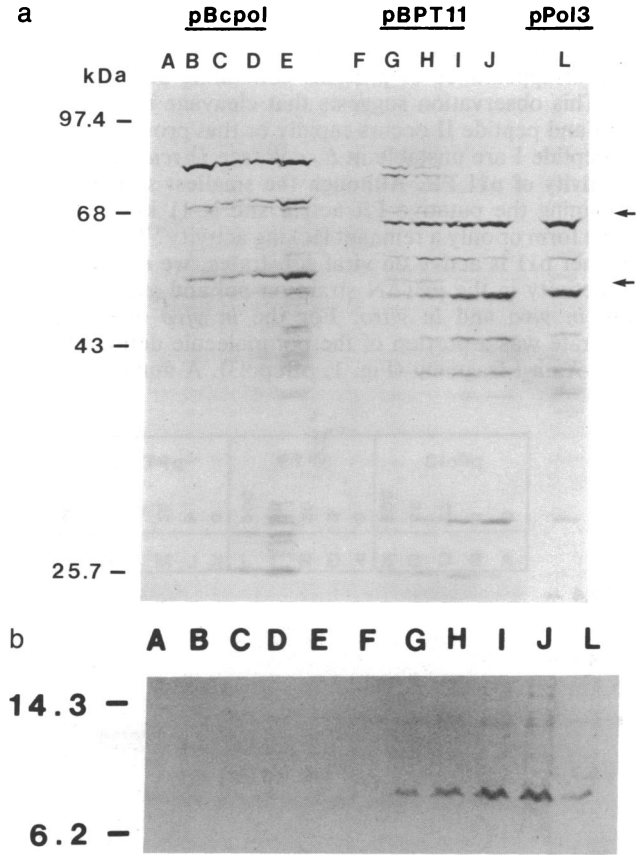

FIG. 3. pol precursor is processed in trans by p11 PR. (*a*) The indicated strains were induced and treated as before. Aliquots (15 μl) were electrophoresed in 10% gels and immunoblotted with a 1:1500 dilution of anti-RT rabbit antiserum (pretreated to reduce background as outlined in *Experimental Procedures*). Cells were induced for 0 min (lanes A and F), 15 min (lanes B and G), 30 min (lanes C and H), 60 min (lanes D, I, and L), or 120 min (lanes E and J). Arrows mark the positions consistent with the size of authentic RT. (*b*) p11 PR is expressed from pBPT11 but not from pBcpol. Aliquots (5 μl) of the above cells were electrophoresed in 15% minigels. Serum II was pretreated as described prior to being used in a 1:5000 dilution. Lane designations are the same as in *a*. Mobilities of prestained molecular size standards are noted.

DISCUSSION

The important role of retroviral PR in the maturation of infectious virus is clear; however, we do not yet understand how the intrinsic protease activity in gag-pol is suppressed prior to viral assembly or how it is subsequently activated after budding. Identifying the requirements for PR activity and specificity may help to elucidate these mechanisms. The data presented here demonstrate that pol can be synthesized and processed in *E. coli* to produce an 11-kDa protein possessing an activity similar to that responsible for processing gag and pol in HIV. This size is similar to that reported for PR from murine leukemia virus (35) and avian leukosis virus (36). The HIV p11 PR demonstrates apparent fidelity and specificity both *in vivo* and *in vitro* on viral substrates. Correct pol processing *in vivo* in *E. coli* is dependent on HIV PR and is identical whether PR is provided in cis (intramolecular cleavage) or in trans (intermolecular cleavage) as p11.

The same p11 is responsible for specific proteolysis of gag substrates *in vitro*. *E. coli*-derived p11 specifically cleaves synthetic peptide substrates containing the Phe-Pro con-

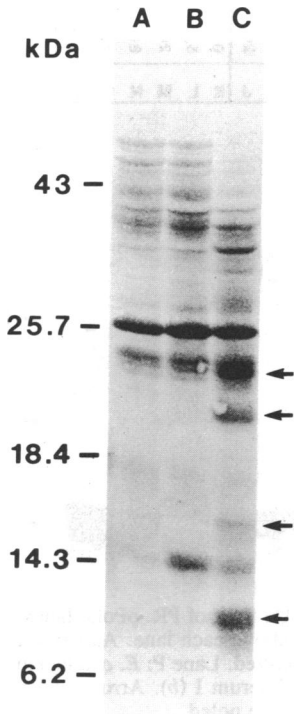

FIG. 4. gag is processed by *E. coli* extracts containing p11. The samples were treated and immunoblotted with monoclonal antibody specific for gag p24. Equivalent amounts of gag were incubated alone (lane A), with extract lacking p11 PR (lane B), or with extract including p11 PR (lane C). Arrows indicate species generated when p11 is expressed. The top arrow marks p24. Positions of prestained molecular size standards are noted.

served sequences (M.C.G. and M. Kotler, unpublished observations). The information presented here leads us to conclude that p11 corresponds to the active form present in virions and that p11 produced in *E. coli* may serve as a viable model system for further studies. This conclusion is supported by the observation of 11.5-kDa (6) and 10-kDa (33) species in HIV virions that are recognized by serum directed to peptides in the PR domain.

The appearance and disappearance of immunoreactive intermediates in pPol3 strains may suggest the temporal order of processing of gag-pol protein; however, any interpretation must be viewed with caution because the precursor is not the full-length gag-pol found in virions. Furthermore, differential stability of intermediates in *E. coli* will affect their relative levels. For instance, the p18 precursor encoded in pPT3 (Fig. 2, lanes G and H, arrow 3) reacts with both sera, as expected. Processing should yield equimolar amounts of p11 and an additional p7 product carrying peptide I. However, p7 is not visualized by serum I in either pPT3 or pPol3 strains, suggesting either that it is unstable or that it is poorly recognized by the polyclonal antisera (low amounts have been observed in other immunoblots; data not shown).

Further experiments will be needed to confirm the structure of the 11-kDa form. Amino- and carboxyl-terminal sequencing of p11 purified from the PR strains described here will unequivocally determine whether p11 is identical to PR found in HIV particles. In addition, it will be interesting to determine the oligomeric structure of the active form. Cellular aspartyl PR analogs require two copies of the conserved active region in order to bind one molecule of substrate, as recently confirmed by x-ray crystallographic data (37). Retroviral PR and yeast Ty PR (26), on the other hand, only contain a single consensus sequence. However, evidence has been observed in murine leukosis virus for PR dimer formation (38), which would yield a form reminiscent of the cellular proteases. A model for retroviral PR dimers has recently been proposed (28), in agreement with these observations.

In conclusion, the HIV PR expressed in *E. coli* is active in the 11-kDa form or a multiple form thereof. The activity is specific and can be observed on viral substrates both *in vivo* and *in vitro*. Such results will aid in the purification and eventual elucidation of the mechanism of this important enzyme.

Note Added in Proof. C. Debouck *et al.* (39) have shown that HIV pol precursor synthesized in *E. coli* can autocatalytically produce a 10-kDa species that is immunoreactive with PR antiserum.

We thank Bryan Cullen for valuable discussions and help in antibody production, C.-Y. Lai for providing gag precursor, Warren McComas and Fran Erlitz for oligonucleotide synthesis, Moshe Kotler for valuable discussions, Andrew Howard for the antiserum-pretreatment suggestion, and Lisa Nieves for expert manuscript preparation. We thank Louis Henderson for sharing data on p11 in virions and Stuart LeGrice for his pol processing information prior to publication. We acknowledge Robert Crowl for thoughtful comments on the work and manuscript and for providing the pEV-vrf vectors.

1. Dickson, C., Eisenman, R., Fan, H., Hunter, E. & Teich, N. (1984) in *RNA Tumor Viruses*, eds. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed., pp. 513–648.
2. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafaski, 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.
3. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. (1985) *Cell* **40**, 9–17.
4. Sanchez-Pescador, R., Power, M. D., Barr, P. J., Steimer, K. S., Stempien, M. M., Brown-Shimer, S. L., Gee, W. W., Renard, A., Randoph, A., Levy, J. A., Dina, D. & Luciw, P. A. (1985) *Science* **227**, 484–492.
5. Muesing, M. A., Smith, D. H., Cabradilla, C. D., Benton, C. V., Lasky, L. A. & Capon, D. J. (1985) *Nature (London)* **313**, 450–458.
6. Henderson, L. E., Copeland, T. D., Sowder, R. C., Schultz, A. M. & Oroszlan, S. (1987) in *UCLA Symposium on Molecular and Cellular Biology*, ed. Bolognesi, D. (Liss, New York), Vol. 71, pp. 135–147.
7. Kramer, R. A., Schaber, M. D., Skalka, A. M., Ganguly, K., Wong-Staal, F. & Reddy, E. P. (1986) *Science* **231**, 1580–1584.
8. Jacks, T. & Varmus, H. E. (1985) *Science* **230**, 1237–1242.
9. Wilson, W., Malim, M. H., Mellor, J., Kingsman, A. J. & Kingsman, S. M. (1986) *Nucleic Acids Res.* **14**, 7001–7016.
10. Teich, N. (1984) in *RNA Tumor Viruses*, eds. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed., pp. 25–207.
11. Rein, A., McClure, M. R., Rice, N. R., Luftig, R. B. & Schultz, A. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7246–7250.
12. Crawford, S. & Goff, S. P. (1985) *J. Virol.* **53**, 899–907.
13. Adams, S. E., Mellor, J., Gull, K., Sim, R. B., Tuite, M. F., Kingsman, S. M. & Kingsman, A. J. (1987) *Cell* **49**, 111–119.
14. Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T. & Oroszlan, S. (1985) *Virology* **145**, 280–292.
15. Veronese, F. D. M., Copeland, T. D., DeVico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C. & Sarngadharan, M. G. (1986) *Science* **231**, 1289–1291.
16. Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A. & Venkatesan, S. (1986) *J. Virol.* **60**, 771–775.
17. Mermer, B., Malamy, M. & Coffin, J. M. (1983) *Mol. Cell. Biol.* **3**, 1746–1758.
18. Bernard, H.-U. & Helinski, D. R. (1979) *Methods Enzymol.* **68**, 482–492.
19. Crowl, R., Seamans, C., Lomedico, P. & McAndrew, S. (1985) *Gene* **38**, 31–38.
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
21. Shaw, G. M., Hahn, B. H., Arya, S. K., Groopman, J. E., Gallo, R. C. & Wong-Staal, F. (1984) *Science* **226**, 1165–1171.
22. Morinaga, Y., Franceschini, T., Inouye, S. & Inouye, M. (1984) *Bio/Technology* **2**, 636–639.
23. Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., Sutcliffe, J. G. & Lerner, R. A. (1982) *Cell* **28**, 477–487.
24. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
25. Shoeman, R. L., Young, D., Pottathil, R., Victor, J., Conroy, R. R., Crowl, R. M., Coleman, T., Heimer, E., Lai, C.-Y., Ganguly, K., Reddy, E. P., Skalka, A. M., Pine, P. R., Khan, F. R. & Weissbach, H. (1987) *Anal. Biochem.* **161**, 370–379.
26. Toh, H., Ono, M., Saigo, K. & Miyata, T. (1985) *Nature (London)* **315**, 691.
27. Miller, R. H. (1987) *Science* **236**, 722–725.
28. Pearl, L. H. & Taylor, W. R. (1987) *Nature (London)* **329**, 351–354.
29. Tanese, N., Sodroski, J., Haseltine, W. A. & Goff, S. P. (1986) *J. Virol.* **59**, 743–745.
30. Farmerie, W. G., Loeb, D. D., Casavant, C., Hutchison, C. A., III, Edgell, M. H. & Swanstrom, R. (1987) *Science* **236**, 305–308.
31. Hansen, J., Schulze, T. & Moelling, K. (1987) *J. Biol. Chem.* **262**, 12393–12396.
32. LeGrice, S. F. J., Beuck, V. & Mous, J. (1987) *Gene* **55**, 95–103.
33. Barr, P. J., Power, M. D., Lee-Ng, C. T., Gibson, H. L. & Luciw, P. A. (1987) *Bio/Technology* **5**, 486–489.
34. Alexander, F., Leis, J., Soltis, D. A., Crowl, R. M., Danho, W., Poonian, M. S., Pan, Y.-C.E. & Skalka, A. M. (1987) *J. Virol.* **61**, 534–542.
35. Yoshinaka, Y., Shames, R. B., Luftig, R. B., Smythers, G. W. & Oroszlan, S. (1985) *J. Gen. Virol.* **66**, 379–383.
36. Vogt, V. M., Wight, A. & Eisenman, R. (1979) *Virology* **98**, 154–167.
37. Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., Sibanda, B. L., Hemmings, A., Wood, S. P., Blundell, T. L., Valler, M. J., Norey, C. G., Kay, J., Boger, J., Dunn, B. M., Leckie, B. J., Jones, D. M., Atrash, B., Hallett, A. & Szelke, M. (1987) *Nature (London)* **327**, 349–352.
38. Yoshinaka, Y. & Luftig, R. B. (1980) *J. Gen. Virol.* **48**, 329–340.
39. Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W. & Rosenberg, M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8903–8906.