JOURNAL OF VIROLOGY, Nov. 1992, p. 6781–6783
0022-538X/92/116781-03$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 66, No. 11

Human Immunodeficiency Virus Type 1 *gag*-Protease Fusion Proteins Are Enzymatically Active

MOSHE KOTLER,[1,2] GILA ARAD,[2] AND STEPHEN H. HUGHES[1]*

ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, P. O. Box B, Frederick, Maryland 21702-1201,[1] and Department of Molecular Genetics, Hadassah Medical School, The Hebrew University, Jerusalem, Israel[2]

Received 10 March 1992/Accepted 27 July 1992

We have introduced mutations into the region of the genome of human immunodeficiency virus type 1 (HIV-1) that encodes the cleavage sites between the viral protease (PR) and the adjacent upstream region of the polyprotein precursor. Segments containing these mutations were introduced into plasmids, and the retroviral proteins were expressed in *Escherichia coli*. The mutations prevented cleavage between the PR and the adjacent polypeptide; however, other PR cleavage sites in the polyprotein were cleaved normally, showing that the release of free PR is not a prerequisite for the appropriate processing of HIV-1 precursors.

The mature *gag* and *pol* proteins found in virions are cleaved from polyprotein precursors by a viral protease (PR) (1). The PR of the human immunodeficiency virus type 1 (HIV-1) is encoded at the beginning of the *pol* gene and is synthesized only as part of the *gag-pol* polyprotein (2, 9, 10). The synthesis of the *gag-pol* polyproteins of HIV-1 requires specific frameshift events that occur approximately 5% of the time that unspliced viral RNA is translated (5).

All known retroviral PRs are dimers, and it is generally accepted that the requirement for dimerization plays an important role in the regulation of PR function (8, 13–15, 17, 19). The *gag* and *gag-pol* polyproteins are cleaved during or after the process of virion assembly (1, 6, 7, 18). It is believed that the delay in the processing of the viral polyproteins is an essential feature of the assembly process; the *pol* proteins are directed to the interior of the virion because they are a part of a *gag-pol* polyprotein. Since viral assembly is dependent on sequences present in *gag*, premature cleavage would result in a virion deficient in *pol* proteins (22).

The three-dimensional structure of the mature HIV-1 PR is known, and there is considerable information about the biochemical properties of this enzyme (for reviews, see references 10 and 18). Much less is known about the structure and the function of PR in the polyprotein precursors. It has been suggested that the adjacent polypeptides influence the structure and function of the retroviral PRs (11, 21). It is also possible that a significant amount of the processing of the polyproteins is done not by free PR but by a form (or forms) of the PR that is still linked to *gag-pol* polypeptides. To investigate the structure and biochemical properties of partially processed precursors of the PR of HIV-1, we have made mutations in the region of the HIV-1 genome that encodes the cleavage site between PR and the adjacent upstream polypeptide transframeshift protein (TF) (3, 4).

Segments of the HIV-1 genome containing the mutations were introduced into *Escherichia coli* expression vectors. Although the mutations in the cleavage site prevented cleavage between PR and the adjacent polypeptides, other sites in the polyprotein were cleaved normally.

The organization of the HIV-1 genome in the region near the boundary between *gag-pol* and the relationship of this region to the segments we inserted into the expression plasmids are shown in Fig. 1. Since we wished to express HIV-1 *gag-pol* fusion proteins efficiently in *E. coli*, it was necessary to modify the region where frameshifting usually occurs (5). Using the polymerase chain reaction, we introduced a mutation that would give rise to a polyprotein with Asn Phe Leu Arg at the *gag-pol* junction into this region. This is the sequence that is believed to be present at the *gag-pol* junction in the major frameshift read-through polyprotein (5).

The polymerase chain reaction was used to convert the segments of the HIV-1 genome that encode the N terminus of capsid protein (CA), TF, and PR into *Nco*I sites. The *Nco*I site (CCATGG) contains an ATG that can serve as an initiator in eukaryotic and prokaryotic expression systems. To preserve the *Nco*I site, a second codon (GTC, Val) was also introduced upstream of the codons that specify the N-terminal amino acids of CA, TF, and PR normally produced by proteolytic cleavage. To express PR by itself, the amplification reaction introduced the Met-Val codons immediately upstream of the Pro codon at position 1832. A termination codon was introduced just downstream of the Phe codon at position 2116. To express the TF protein, the Met-Val codons that contain the *Nco*I site were inserted immediately in front of the Arg codon at position 1634 (20). To permit efficient translation through the *gag-pol* junction, a frameshift mutation was introduced just downstream of the A at position 1634 by inserting a single base, A. This single-base insertion puts *gag* and *pol* into the same reading frame. The Phe codon normally found at the C terminus of TF (position 1829) was converted to Ile. The termination codon for TF was supplied by the T7 expression plasmid pT5. The CA/NC fusion protein truncated at the frameshift site (CA-ΔNC) was linked to *Nco*I Met-Val at a Pro codon (position 729). The last HIV-1 codon was Phe (position 1649). The T7 expression plasmid supplied the termination codon.

These three amplified DNA segments were all cleaved with *Nco*I. The CA-ΔNC, TF, and PR segments were also cleaved with *Eco*RI, *Sac*I, and *Bam*HI, respectively. The segments were initially cloned into pUC12N (16, 24). For expression, these segments were subcloned in the T7 expression plasmid pT5. Larger segments were built up by combining the individual cloned segments. Appropriate choice of

* Corresponding author.

FIG. 1. Expression of HIV-1 proteins in *E. coli*. The upper part of the figure shows the organization of the genome of HIV-1 in the region near the junction (arrow) of *gag* and *pol*. The relative positions of the encoded polypeptides (CA, NC, P6, TF, and PR) and the molecular weights (in thousands) of the polypeptides (24, 7, 6, and 11) are indicated. The lower part of the figure depicts the various proteins that were expressed in *E. coli* and indicates the approximate positions of the mutation that was introduced in the cleavage site between TF and PR (CS) and the mutation that puts *gag* and *pol* in the same reading frame (FS). The reactions of specific antisera with polypeptides contained in the extracts are shown on the right. Immunological detection in Western transfers of proteins of the expected molecular weights are indicated (+). Immunological detection of a precursor in the absence of free immunoreactive product is indicated (PC).

FIG. 2. Expression and processing of HIV-1 polypeptides in *E. coli*. Lysates of *E. coli* strains that express the indicated polypeptides (see text and Fig. 1 for a more complete description of the HIV-1 polypeptides) were fractionated on sodium dodecyl sulfate–15% polyacrylamide gels and transferred to nitrocellulose paper. HIV-1 polypeptides were detected by reaction with a monoclonal antibody to CA (A) or antiserum specific for PR (B). Reaction products were detected chemically with alkaline phosphatase. Molecular weights (in thousands) are indicated on the left.

segments allowed us to build versions of the T7 expression plasmids which expressed wild-type versions of the HIV-1 proteins and versions that had the mutation at the proteolytic cleavage site between TF and PR (Fig. 1). The resulting plasmids were transferred into a BL21 strain of *E. coli* which contains an inducible T7 polymerase (23). Induction of T7 polymerase expression results in the production of an HIV-1 polyprotein which is cleaved by the PR it contains (see Fig. 2). Individual *gag* and *pol* proteins (CA and PR) were identified with specific antisera, showing that the frameshift mutation we introduced gives rise to a polyprotein that can be appropriately processed in *E. coli*. (The anti-PR serum, a generous gift of Lily Babe and Charles Craik [University of California, San Francisco], is a rabbit polyclonal antiserum made against recombinant HIV-1 protease; monoclonal antibody against HIV-1 CA was kindly provided by S. N. Nigida [PRI]). The precursor-product relationship between the larger precursors and the smaller cleavage products was confirmed by analyzing samples at various times after induction (data not shown).

Although other researchers have previously shown that the appropriate processing of HIV-1 precursors is absolutely dependent on the presence of active HIV-1 PR (for reviews, see references 10 and 18), we have expressed CA linked to a truncated form of NC as a control (CA-ΔNC). Although this fusion protein is made in *E. coli*, it is not cleaved to produce free CA (Fig. 2A). We have also made strains that synthesize free TF, free PR, and a TF-PR fusion protein. These strains produce proteins that serve as size markers and provide controls to demonstrate the specificities of the antisera we have used.

To test the enzymatic activity of the TF-PR fusion protein, a mutation was introduced into the PR cleavage site that joins these two polypeptides in the polyprotein. The Phe residue at the C terminus of TF was changed to Ile (TTC to ATC). This mutation was introduced into the plasmid that expresses a large polyprotein that includes the sequences between CA and the PR. Although CA is released by specific PR cleavage, just as it was with the expression plasmid containing the normal cleavage site, the HIV-1 PR remains linked to TF (Fig. 2B). The TF-PR fusion protein can be identified because it specifically reacts with anti-PR antiserum. This same antiserum did not detect any free PR in the extract.

These data suggest, but do not prove, the hypothesis that the PRs would be active in the context of the complete *gag* or *gag-pol* polyprotein. The PR of HIV-1 is part of the *gag-pol* polyprotein, so it remains possible that the linkage to *pol* protein(s) would interfere with PR function.

Loeb et al. (11) produced HIV-1 *pol* polyproteins in *E. coli* and showed that mutations that interfered with cleavage between PR and reverse transcriptase or between reverse transcriptase and integration protein prevented PR from cleaving other sites in the polyprotein. Louis et al. (12) found that an HIV-1 fusion protein that contained the *E. coli* maltose-binding protein and the HIV-1 *pol* gene was not properly processed if PR was prevented, by a mutation in the cleavage site, from releasing the maltose-binding protein. However, Loeb et al. (11) found that some (but not all) of their polyproteins, which contain 23 N-terminal amino acids from *E. coli lacZ* and a polylinker, show PR activity even when PR is not allowed to cleave off the N-terminal segment which contains the *lacZ*-polylinker sequences. We suggest that the difference between the results of Louis et al. (12) and those of Loeb et al. (11) may have to do with the sizes of foreign sequences fused to PR. Since both sets of experiments involved the addition of foreign sequences, it was not possible to decide on the basis of the published data whether

to expect a *gag* fusion protein which contains PR to have significant proteolytic activity.

Although we can now say that an HIV-1 *gag*-PR fusion protein is enzymatically active, our data do not allow us to decide whether the PR of HIV-1 is as active when linked to *gag* proteins as it is when it has been released from the polyprotein. We have monitored the accumulation of the cleaved HIV-1 CA protein at various times after induction, and there is no striking difference in the rate of accumulation of CA whether or not the HIV-1 PR can be cleaved from TF (data not shown). Unfortunately, monitoring the time course of the accumulation of the free CA protein does not necessarily provide an accurate measurement of the activity of the PR, since other factors, for example, the folding or the dimerization of the precursor, could be the rate-limiting step in processing the polyprotein.

We thank Hilda Marusiodis for typing the manuscript.

This work was supported by a grant from the Chief Scientist of the Israeli Ministry of Health and by the Israel Cancer Research Fund and was sponsored in part by the National Cancer Institute, DHHS, under contract NO1-CO-74101 with ABL.

REFERENCES

1. **Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich.** 1984. Protein biosynthesis and assembly, p. 513–648. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Graves, M. C., J. J. Lim, E. P. Heimer, and R. A. Kramer.** 1988. An 11-kDa form of human immunodeficiency virus protease expressed in *Escherichia coli* is sufficient for enzymatic activity. Proc. Natl. Acad. Sci. USA **85:**2449–2453.
3. **Henderson, L. E., R. E. Benveniste, R. Sowder, T. D. Copeland, A. M. Schultz, and S. Oroszlan.** 1988. Molecular characterization of *gag* proteins from simian immunodeficiency virus (SIV_{Mne}). J. Virol. **62:**2587–2595.
4. **Henderson, L. E., R. C. Sowder, T. D. Copeland, R. E. Benveniste, and S. Oroszlan.** 1988. Isolation and characterization of a novel protein (X-orf product) from SIV and HIV-2. Science **241:**199–201.
5. **Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, and H. E. Varmus.** 1985. Characterization of ribosomal frameshifting in HIV-1 *gag-pol* expression. Nature (London) **331:**280–283.
6. **Kaplan, A. H., and R. Swanstrom.** 1991. Human immunodeficiency virus type 1 *gag* proteins are processed in two cellular compartments. Proc. Natl. Acad. Sci. USA **88:**4528–4532.
7. **Katoh, I., Y. Yoshinaka, A. Rein, M. Shibuya, T. Okada, and S. Oroszlan.** 1985. Murine leukemia virus maturation: protease region is required for conversion from "immature" to "mature" core form and for virus infectivity. Virology **145:**280–292.
8. **Kotler, M., W. Danho, R. A. Katz, J. Lies, and A. M. Skalka.** 1989. Avian retrovirus protease and cellular aspartic protease are distinguished by activities on peptide substrates. J. Biol. Chem. **264:**3428–3435.
9. **Kramer, R. A., M. D. Schaber, A. M. Skalka, K. Ganguly, F. Wong-Staal, and E. P. Reddy.** 1986. HTLV III *gag* protein is processed in yeast cells by the virus *pol*-protease. Science **231:**1580–1584.
10. **Krausslich, H. G., and E. Wimmer.** 1988. Viral proteinases. Annu. Rev. Biochem. **57:**701–754.
11. **Loeb, D. D., C. A. Hutchison III, M. H. Edgell, W. G. Farmerie, and R. Swanstrom.** 1989. Mutational analysis of human immunodeficiency virus type 1 protease suggests functional homology with aspartic proteinases. J. Virol. **63:**111–121.
12. **Louis, J. M., R. A. McDonald, N. T. Nashed, E. M. Wondrak, D. M. Jerina, S. Oroszlan, and P. T. Mora.** 1991. Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in *Escherichia coli*. Eur. J. Biochem. **199:**361–369.
13. **Meek, T. D., B. D. Dayton, B. W. Metcalf, G. B. Dreyer, J. E. Strickler, J. G. Gorniak, M. Rosenberg, M. Moore, V. W. Magaard, and C. Debouck.** 1989. Human immunodeficiency virus-1 protease expressed in *Escherichia coli* behaves as a dimeric aspartic protease. Proc. Natl. Acad. Sci. USA **86:**1841–1845.
14. **Miller, M., M. Jaskolsky, J. K. M. Fao, J. Leis, and A. Wlodawer.** 1989. Crystal structure of a retroviral protease proves relationship to aspartic protease family. Nature (London) **337:**576–579.
15. **Navia, M. A., P. M. D. Fitzgerald, B. M. McKeever, C.-T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer.** 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature (London) **337:**615–620.
16. **Norrander, J. M., J. Vieira, I. Rubenstein, and J. Messing.** 1985. Manipulation and expression of the maize zein storage proteins in *Escherichia coli*. J. Biotechnol. **2:**157–175.
17. **Nutt, R. F., S. F. Brady, P. L. Darke, T. M. Ciccarone, C. D. Colton, E. M. Nutt, J. A. Rodkey, C. D. Bennett, L. H. Waxman, I. S. Sigal, P. S. Anderson, and D. F. Veber.** 1988. Chemical synthesis and enzymatic activity of a 99-residue peptide with a sequence proposed for the human immunodeficiency virus protease. Proc. Natl. Acad. Sci. USA **85:**7129–7133.
18. **Oroszlan, S., and R. B. Luftig.** 1990. Retroviral proteinases. Curr. Top. Microbiol. Immunol. **157:**153–185.
19. **Pearl, L. H., and W. R. Taylor.** 1987. A structural model for the retroviral protease. Nature (London) **329:**351–354.
20. **Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Joseph, E. R. Doran, J. A. Rafalski, K. B. Whitehorn, L. Ivanoff, S. R. Petteway, Jr., M. L. Pearson, J. A. Lautenberg, T. S. Papas, J. Ghrayeb, N. T. Chang, R. C. Gallo, and F. Wong-Staal.** 1985. Complete sequence of the AIDS virus, HTLV-III. Nature (London) **313:**277–284.
21. **Sedlacek, J., P. Strop, F. Kapralek, V. Pecenka, V. Kostka, M. Travnicek, and J. Riman.** 1988. Processed enzymatically active protein ($p15^{gag}$) of avian retrovirus obtained in an *E. coli* system expressing a recombinant precursor ($Pr^{lac\Delta gag}$). FEBS Lett. **237:**187–190.
22. **Shields, A., O. N. Witte, E. Rothenberg, and D. Baltimore.** 1978. High frequency of an aberrant expression of Moloney murine leukemia virus in clonal infections. Cell **14:**601–609.
23. **Studier, F. W., A. H. Rosenberg, J. J. John, and J. W. Bubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
24. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.