Volume 314, number 3, 449–454 FEBS 11911 December 1992

Intrinsic activity of precursor forms of HIV-1 proteinase

Lowri H. Phylip[a], John S. Mills[b], Benne F. Parten[c], Ben M. Dunn[c] and John Kay[a]

[a]*Department of Biochemistry, University of Wales College of Cardiff, Cardiff CF1 1ST, Wales, UK,* [b]*Roche Research Centre, Welwyn Garden City, Herts AL7 3AY, UK* and [c]*Department of Biochemistry and Molecular Biology, J. Hillis Miller Health Centre, University of Florida, Gainesville, FL 32610, USA*

Received 5 November 1992

The wild-type -Phe*Pro- bond located at the N-terminus of the mature aspartic proteinase of HIV-1 was replaced by -Ile-Pro- or -Val-Pro-. By this means, processing at this cleavage junction was prevented and so, extended or precursor forms of HIV-proteinase were generated. These constructs were expressed in *Escherichia coli*, purified therefrom, and their specificity, activity at different pH values and susceptibility to the potent inhibitor, Ro31-8959, was assessed. A hitherto unobserved cleavage junction (at ~Ala-Phe*Leu-Gln~) in the frame-shift region of the gag–pol viral genome was identified and confirmed by demonstrating cleavage of a synthetic peptide corresponding to this region. The implications for viral replication of self-processing at neutral pH by proteinase whilst still present (in a precursor form) as a component of the polyprotein are considered; such reactions, however, are still blocked even at pH values as high as 8.0 by Ro31-8959.

HIV-proteinase precursor; Activity and pH dependence; Inhibition by Ro31-8959; Novel polyprotein cleavage junction

1. INTRODUCTION

Mammalian and fungal aspartic proteinases (PRs) are synthesized in the form of zymogens. Conversion to generate each mature PR is frequently autocatalytic and requires the removal of a propart segment consisting of 40–50 residues from the N-terminus of the molecules [1]. In the human immunodeficiency virus (HIV), synthesis of viral gag and pol proteins as the gag–pol polyprotein necessitates cleavage to release the mature forms of each protein component. The dimeric aspartic PR encoded therein is responsible for these cleavages, including its own autocatalytic release [2]. The 65 residues (from the start of the pol orf) immediately upstream from the mature HIV-1 PR (Fig. 1) have not as yet been ascribed a function but correspond to the prosegment found in the archetypal aspartic PR [3]. Processing of the gag–pol polyprotein may be governed by the form in which its intrinsic PR is present, i.e. initially, at least, as a larger precursor which may be different from 'mature' PR in terms of specificity, catalytic efficiency, pH dependence and, in turn, susceptibility to blockage by inhibitors designed against the mature PR. In order to study this, precursor or extended forms of the mature (99 residue) HIV-1 PR were engineered and their properties were examined.

Correspondence address: J. Kay, Department of Biochemistry, University of Wales College of Cardiff, P.O. Box 903, Cardiff CF1 1ST, Wales, UK. Fax: (44) (222) 874 116.

Abbreviations: HIV, human immunodeficiency virus; PR, proteinase; MBP, maltose-binding protein. The nomenclature system of Schechter and Berger [11] i.e. P_2–P_1* P_1'–P_2' is used to depict amino acids adjacent to the residues in the P_1 and P_1' positions which contribute to the scissile peptide bond (indicated by an asterisk*).

2. MATERIALS AND METHODS

The *Bam*HI–*Eco*RV fragment of the BH10 strain of HIV-1 encoding the prosegment, PR, and the N-terminal region of reverse transcriptase was used in all constructs. In order to delineate the PR components from reverse transcriptase, an in-frame stop codon was introduced at the C-terminus of PR using the oligonucleotide, 5′–AGTCTCAATAGGGCGAATTCAAAAATTTAAAGTGCA–3′, in conjunction with the Oligonucleotide-directed in vitro mutagenesis system (Amersham International, Bucks, UK). PCR mutagenesis by overlap extension [4] was used to introduce the mutations, -Ile-Pro- or -Val-Pro- in place of the wild-type -Phe*Pro- cleavage junction at the N-terminus of PR (Fig. 1). Comparable results were derived for both mutations so, for brevity, only the data for the -Val-Pro- constructs are described. The mutated PR gene was subcloned into the following expression vectors; (i) pDS56 [5]; (ii) pDS56, 6His [6] which encodes 6 His residues upstream from the N-terminus of the protein (construct H_V in Fig. 1); and (iii) pMAL Maltose-binding protein (MBP) fusion and purification system (New England Biolabs. Inc. Beverly, MA, USA) (construct M_V in Fig. 1). The corresponding wild-type constructs were also engineered. Following all mutagenesis and reconstruction, each nucleotide sequence was confirmed by dideoxy sequencing.

E. coli strain M15 was used throughout for expression studies [7]. Following induction with IPTG (between 70 and 400 μg/ml), culture aliquots were removed after 2 h and analyzed by Western blotting using a monoclonal antibody directed against the active site region of mature PR, as described previously [8].

For large-scale purification, the MBP-extended PR fusion (product of M_V in Fig. 1) was purified in a single step from lysates of 500 ml M15 cultures by affinity chromatography on an amylose column as described in the manufacturer's protocol. Autocatalytic conversion of the purified fusion protein was analyzed at 37°C in 10 mM Tris-HCl buffer, pH 8.0, containing 100 mM NaCl, or in the same buffer adjusted to pH 5.0, 6.0 or 7.0 by the addition of sodium phosphate

buffers, as appropriate. Protein bands from SDS-gels were blotted directly onto PVDF membranes and loaded into an Applied Biosystems ABI 470A Sequenator for N-terminal sequence analysis by Edman degradation. The peptide, GSGDLAF*LQGKAR, was synthesized by solid-phase methodologies and its hydrolysis was monitored by reverse-phase FPLC, as described previously [9].

3. RESULTS AND DISCUSSION

We have shown previously with synthetic peptide [9,10] and engineered protein [8] substrates that introduction of a β-branched amino acid into the P_1 position (nomenclature of Schechter and Berger [11]) prevents substrate cleavage by HIV PR. Consequently, to engineer an extended form of HIV PR that would not be capable of self-processing at its N-terminus to release mature PR, the N-terminal junction between the prosegment and PR was mutated from the wild-type -Phe*Pro- (Fig. 1) to -Ile-Pro- or -Val-Pro-. Upon expression of these mutant constructs in *E. coli*, a larger form of PR (M_r ~18 kDa) was detected in the cell extracts by Western blotting (data not shown). Only low levels of the two extended PR accumulated, however. In order to generate larger amounts for further study and to facilitate purification by Ni-chelate chromatography [12], a fusion construct was made (H_V in Fig. 1) in which 6 histidine residues were positioned immediately upstream from the start of the pol orf. Upon expression in *E. coli*, the equivalent wild-type -Phe*Pro- construct processed to release the anticipated mature (11 kDa) PR (Fig. 2) whereas the -Val-Pro- mutant construct generated two bands of approx. 19 and 18 kDa when cell lysates were examined by SDS-PAGE and Western blotting (Fig. 2). The 19 kDa band had a comparable mobility to that of a distinct form of an extended HIV PR (Fig. 2, cf. lanes 2 and 3); this was produced by mutation of the active site Asp^{25} residue (to Ala [13]), thus abolishing the autolytic processing ability to release the prosegment from the mature 11 kDa proteinase. The 18 kDa band may have arisen by autocatalytic clipping near the N-terminus of the extended PR molecules (see below). An identical result was obtained with the mutant containing -Ile-Pro- (data not shown). There was, thus, no indication with these mutations of processing occurring at the N-terminus of mature PR to generate this 11 kDa protein. The strategy of generating -Ile-Pro- or -Val-Pro- sequences in place of the wild-type -Phe*Pro- would appear, then, to have been effective in preventing autolytic cleavage at this junction. However, extensive attempts to extract and further purify these forms of extended PR revealed that the histidine-tagged proteins were predominantly insoluble.

Consequently, a further construct was engineered in which the -Val-Pro- mutant PR was placed as an in-

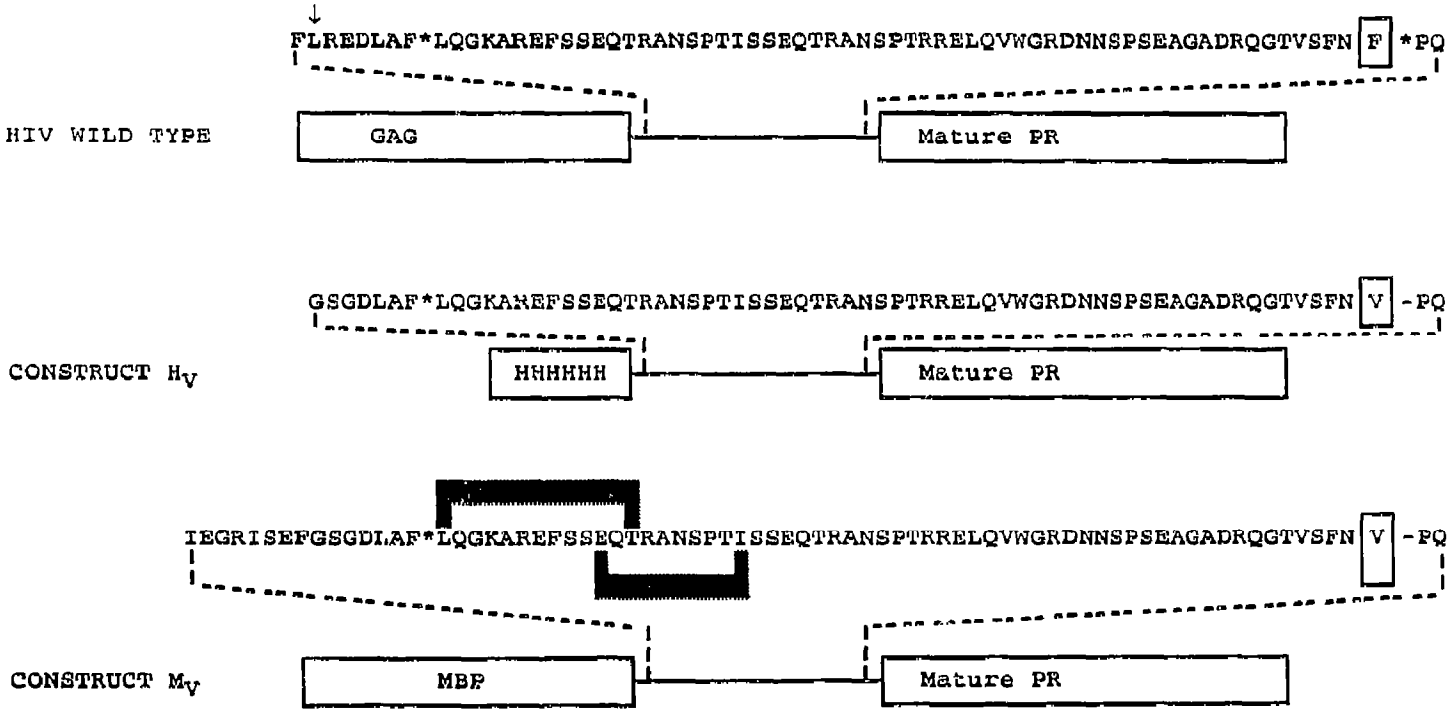

Fig. 1. Extended HIV-1 proteinase constructs. The wild-type viral gag–PR sequence linking the C-terminus of gag with the N-terminus of PR in the pol region of the genome is detailed in the top panel. The frame shift location is indicated by an arrow. In construct H_V (middle panel), the F*P bond at the N-terminus of mature PR was mutated to V-P and a H_6 tag was introduced in front of the extended PR molecules to facilitate affinity purification. In construct M_V (bottom panel), a MBP domain was used instead. N-Terminal sequences that were determined (see text) are indicated in this construct by the bars.

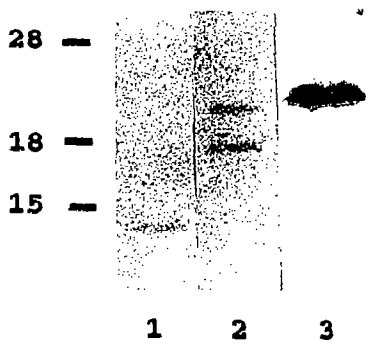

Fig. 2. Characterization of the products derived from expression in *E. coli* of plasmids harbouring H_6-tagged extended HIV-1 PR. Construct H_V (see Fig. 1) containing -V-P- at the N-terminus of the mature PR, is depicted in lane 2. Lanes 1 and 3 illustrate controls expressing the corresponding wild-type constructs with -F*P- at the N-terminal cleavage junction (lane 1) and with mutation of the active site Asp residue PR to Ala (lane 3). Immunodetection was carried out using an anti-PR monoclonal antibody. Mol. wt. markers (kDa) migrated as indicated.

frame fusion behind the soluble MBP (mol. wt. 42 kDa) linked through the sequence -IEGRISEF- to the start of the prosegment (construct M_V in Fig. 1). Following IPTG induction, two bands (at ~60 and 18 kDa) on Western blots were stained with the anti-PR monoclonal antibody, whereas only the 60 kDa band reacted with antibody to MBP. The upper band corresponded in size to that anticipated for the fusion between MBP and extended PR. Reasonable quantities (10 mg/l) of soluble product were generated in *E. coli* and further purification was accomplished by affinity chromatography on an amylose–agarose column. The material retained by the amylose resin displayed the same two 60 kDa and 18 kDa bands on SDS-PAGE (Fig. 3), as had been observed in the cell lysates. The 18 kDa band was eluted from a preparative SDS-gel and Edman degradation indicated it to have the N-terminal sequence, LQGKAREFSSEQT~ (see Fig. 1, bottom). Since affinity chromatography on amylose resin requires the presence of a MBP component, the presence of this 18 kDa band in the purified material is most readily explained by the formation of heterodimeric molecules through association of one subunit consisting of full-length MBP-extended PR with another comprising extended PR but devoid of MBP. The formation of such a heterodimeric structure suggests that the protein must have been folded correctly in order to achieve dimerization.

The column-purified material was dialysed against 10 mM Tris-HCl buffer, 100 mM NaCl, pH 8.0 and then incubated in this buffer at 37°C. The fusion protein disappeared, as monitored both by the antiserum to MBP and the anti-PR monoclonal (Fig. 4, left panels, upper and lower, respectively) with generation of free MBP (at ~42 kDa). This conversion was slow initially (within the first 24 h) but accelerated thereafter (to be complete between 24 and 30 h), reminiscent of an autocatalytic reaction. A parallel incubation was thus performed in the presence of 100 nM Ro31-8959, an inhibitor that is completely specific for HIV PR [14]. No conversion occurred under these conditions (Fig. 4, right panels), confirming that an autocatalytic cleavage is indeed responsible. When the fate of the 18 kDa extended PR was monitored, this band disappeared (Fig. 4, bottom left panel). Two new bands (of around 14 and 13 kDa) were generated that were immunoreactive with the anti-PR antibody. The upper of these (at ~14 kDa) was eluted from a preparative SDS-gel and subjected to N-terminal sequencing. The sequence began EQTRANSPTI~ (Fig. 1, bottom). After incubation for 30 h and longer, no immunoreactive PR species could be detected (Fig. 4, bottom left panel). This autolytic destruction was prevented by inclusion of Ro31-8959 (Fig. 4, bottom right panel).

A further set of incubations was then performed at pH 5.0, 6.0, 7.0 and 8.0, and the extent of conversion was assessed by SDS-PAGE and silver staining. No autolytic release of the MBP from the fusion protein was observed at pH 5.0 (Fig. 5), whereas conversion did occur at all the higher pH values. This implies that the extended HIV-1 PR does not require an acidic environment to be active. This is in contrast with the mature form of the enzyme which is maximally active at pH values of 5.0 and below [9].

From these data, it is apparent that the extended or precursor forms of HIV PR have intrinsic activity which can be directed in an autolytic fashion to ensure self-processing. Mature HIV PR is capable of hydrolysing two types of cleavage junction (-Aromatic*Pro- and -Hydrophobic*Hydrophobic- bonds [15]) and many

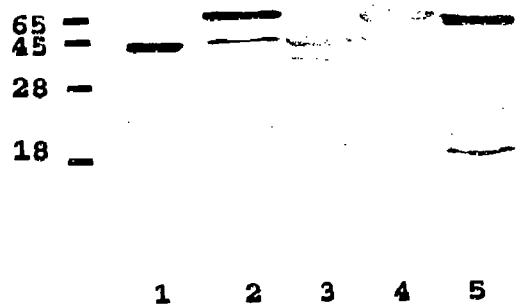

Fig. 3. SDS-PAGE of protein purified by affinity chromatography of the soluble extract from *E. coli* harbouring the plasmid containing the MBP-extended HIV-1 PR fusion. After centrifugation, soluble lysates were applied to a column of amylose–agarose in 10 mM sodium phosphate buffer, pH 7.0 containing 0.5 M NaCl, 10 mM mercaptoethanol and 1 mM EGTA. After extensive washing, desorption was achieved using the same buffer containing 10 mM maltose. Aliquots of this desorbed material were analyzed by SDS-PAGE and visualized by staining with Coomassie blue (lane 2) or, in Western blots, by immunostaining with an antiserum to MBP (lane 4) or with an anti-PR monoclonal antibody (lane 5). Samples of authentic MBP (supplied by New England Biolabs. Inc.) were stained with Coomassie blue (lane 1) or with the antiserum to MBP (lane 3).

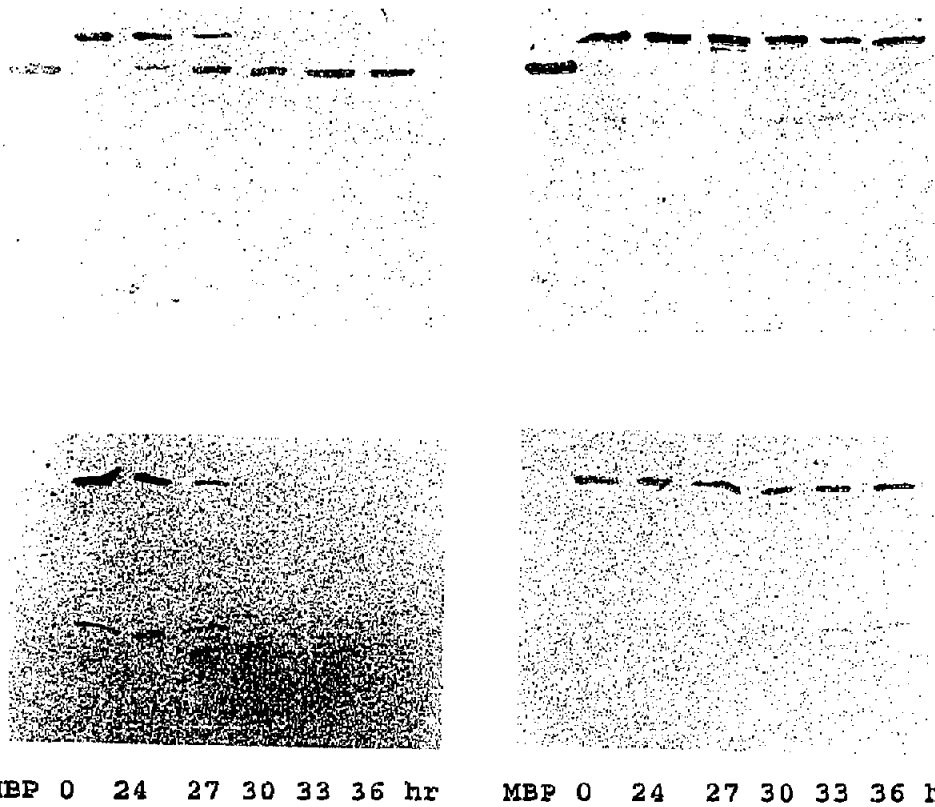

Fig. 4. Self-processing of the affinity-purified products from expression of construct M_V in *E. coli*. The column-purified material was incubated at pH 8.0 at 37°C in the absence (left panels) and in the presence (right panels) of the inhibitor, Ro31-8959 (100 nM). Samples were removed at the indicated times, run on Western blots and stained with an antiserum to MBP (upper panels) or the anti-PR monoclonal antibody (lower panels).

previous investigations by us and by others [9,10,16] have revealed that, for the latter, the sequence -Ile/Val/Ala-Hydrophobic*Hydrophobic-Glu/Gln- is optimal. The sequence -Ala-Phe*Leu-Gln- found at the N-terminal end of our constructs (Fig. 1) matches this consensus pattern, and its position, immediately downstream from the gag–pol frame shift region of the viral genome (indicated by the arrow in Fig. 1), would suggest another cleavage junction, hitherto unobserved, in this location in the viral polyprotein. In order to test this assertion further, a synthetic peptide (GSGDLAF*LQGKAR), spanning this putative cleavage site and corresponding to the sequence in our H_V and M_V constructs, was incubated with homogeneous mature HIV-1 PR under identical conditions (at pH 4.7) to those we have described previously for other synthetic peptide substrates [9,10]. Complete cleavage to generate two products was observed; these were seperated by reverse-phase FPLC and their amino acid compositions were determined. These were Asp 1.0, Ser 0.9, Gly 2.2, Ala 1.1, Leu 1.1, Phe 1.1, and Glu 1.1, Gly 1.0, Ala 1.0, Leu 1.0, Lys 1.0, Arg 1.1, respectively, confirming that cleavage had indeed taken place at the projected -Phe*Leu- bond. Values of 680 μM and 2 s^{-1} were determined, respectively, for the kinetic parameters, K_m and k_{cat}, for the hydrolysis of this peptide. These are comparable to the values obtained previously with other peptide substrates for HIV-1 PR containing Ala and Gln in P_2 and P_2', respectively [10,15,16]. Thus, it would appear that an additional cleavage junction may well be present in the gag–pol polyprotein of HIV-1. Hydrolysis at this location would ensure that gag molecules were 'disconnected' from the encumbrance of a pol tail despite the frame shift; and in the downstream direction, with cleavage also occurring in the wild-type genomic product at the N-terminus (Phe*Pro) of the mature PR (Fig. 1, top line), a peptide of approx. 65 residues should be released. Indeed, a component of just such a size has been visualized (by an antibody which recognizes this segment) in virion extracts (S.F.J. Le Grice, personal communication) but its functional significance remains to be elucidated.

The other implication which arises from this study is that the extended or precursor forms of HIV PR (albeit in the form of a heterodimer with one subunit attached to MBP) may not require acidic conditions for activity. Considerable thought (see e.g. in [17]) has been given as to how the developing HIV virion may become acidified in order to accomplish its essential proteolytic processing. But acidification may not be necessary; if the char-

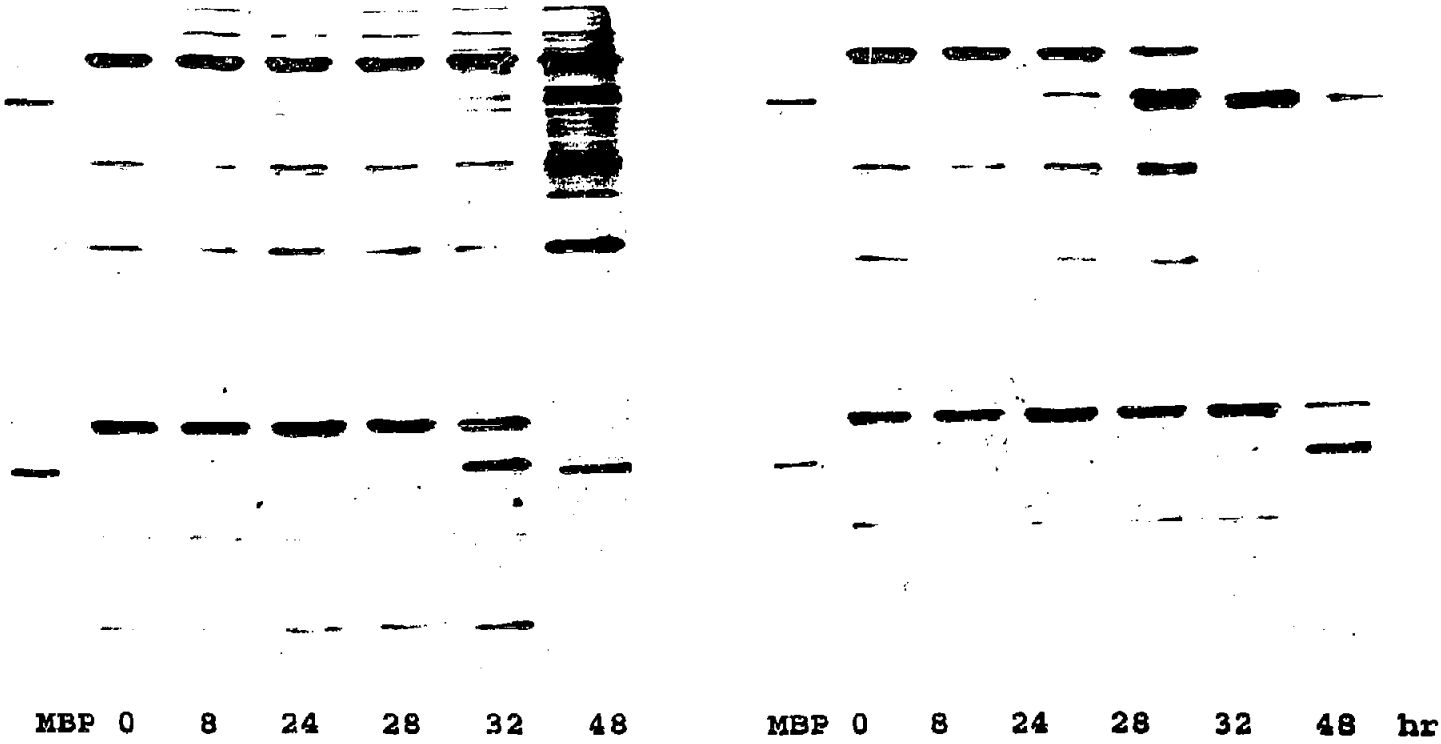

Fig. 5. pH dependence of the self-processing of the MBP-extended HIV-1 PR. The affinity-purified products from expression of construct M_Y in *E. coli* were incubated at pH 5.0 or 6.0 (upper panels, left and right, respectively) and pH 7.0 or 8.0 (lower panels, left and right, respectively). Samples were removed at the indicated times and analyzed by SDS-PAGE followed by silver staining. The *absence* of processing at pH 5.0 is illustrated from the increased loading of material in the 48 h sample. Samples of authentic MBP are included as markers in all panels.

acteristics of the fusion protein (an artificial construct for ease of manipulation) studied herein should be reflected in the wild-type gag–PR–pol polyprotein fusion, then the precursor form(s) with PR still present as a component in the polyprotein, may be capable of performing the requisite cleavages *at neutral pH*. At the same time, this processing would result in the release of the mature form of PR itself, which would be dormant under the neutral pH conditions of the infected cell. This may be advantageous to viral replication overall since the PR would be less able to attack host cell proteins and thus incapacitate the host cell; under *acidic* conditions, PR does digest cellular proteins e.g. [18–20].

Clearly, further investigations are necessary to examine this concept in more detail, and experiments are under way to purify sufficient quantities of the extended forms of PR to permit comparison of their kinetic parameters at different pH values with those determined previously [9,10] for the mature form of HIV-1 PR. However, it is evident from Fig. 4, that, irrespective of the precise nature of the form of the PR that is responsible for polyprotein processing and the pH at which this takes place, the inhibitor, Ro31-8959, is effective in blocking proteolysis even at pH values as high as 8.0. This compound is currently undergoing clinical trials in Italy, France and the UK for the treatment of AIDS; their outcome is thus being awaited with justifiable interest.

Acknowledgements: This work was supported by the AIDS-Directed Programme of the Medical Research Council and by NIH Grant AI 28571. We are very grateful to Ray Jupp, Stuart Le Grice and Jonathan T. Griffiths for contributing their enthusiasm and valuable advice, and particularly to Dr. Carol Carter, Stony Brook, for many exciting discussions and for sharing with us her own results in this area prior to publication.

REFERENCES

[1] Kay, J. (1985) in: Aspartic Proteinases and Their Inhibitors (Kostka, V. ed.) pp. 1–18, Walter de Gruyter, Berlin.
[2] Kay, J. and Dunn, B.M. (1990) Biochim. Biophys. Acta 1048, 1–18.
[3] Partin, K., Zybarth, G., Ehrlich, L., De Crombrugghe, M., Wimmer, E. and Carter, C. (1991) Proc. Natl. Acad. Sci. USA 88, 4776–4780.
[4] Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Gene 77, 51–59.
[5] Stüber, D., Ibrahimi, I., Cutler, D., Dobberstein, B. and Bujard, H. (1984) EMBO J. 3, 3143–3148.
[6] Bujard, J., Gentz, R., Lanzer, M., Stüber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.T. and Dobberstein, B. (1987) Methods Enzymol. 155, 416–433.
[7] Certa, U., Bannwarth, W., Stueber, D., Gentz, R., Lanzer, M., Le Grice, S.F.J., Guillot, F., Wendler, I., Hunsmann, G., Bujard, H. and Mous, J. (1986) EMBO J. 5, 680–685.

[8] Jupp, R.A., Phylip, L.H., Mills, J.S., Le Grice, S.F.J. and Kay, J. (1991) FEBS Lett. 283, 180–184.
[9] Richards, A.D., Phylip, L.H., Farmerie, W.G., Scarborough, P.E., Alvarez, A., Dunn, B.M., Hirel, Ph.-H., Konvalinka, J., Strop, P., Pavlickova, L., Kostka, V. and Kay, J. (1990) J. Biol. Chem. 265, 7733–7736.
[10] Phylip, L.H., Richards, A.D., Kay, J., Konvalinka, J., Strop, P., Kostka, V., Ritchie, A.J., Broadhurst, A.V., Farmerie, W.G., Scarborough, P.E. and Dunn, B.M. (1990) Biochem. Biophys. Res. Commun. 171, 439–444.
[11] Schechter, I. and Berger, A. (1967) Biochim. Biophys. Res. Commun. 27, 157–162.
[12] Hochuli, E., Bannwarth, W., Dubeli, H., Gentz, R. and Stüber, D. (1988) Biotechnology 6, 1321–1325.
[13] Le Grice, S.F.J., Mills, J.S. and Mous, J. (1988) EMBO J. 7, 2547–2553.
[14] Roberts, N.A., Martin, J.A., Kinchington, D., Broadhurst, A.V., Craig, J.C., Duncan, I.B., Galpin, S.A., Handa, B.K., Kay, J., Krohn, A., Lambert, R.W., Merrett, J.H., Mills, J.S., Parkes, K.E.B., Redshaw, S., Ritchie, A.J., Taylor, D.L., Thomas, G.L. and Machin, P.J. (1990) Science 248, 358–361.
[15] Griffiths, J.T., Phylip, L.H., Konvalinka, J., Strop, P., Gustchina, A., Wlodawer, A., Davenport, R.J., Briggs, R., Dunn, B.M. and Kay, J. (1992) Biochemistry 31, 5193–5200.
[16] Poorman, R.A., Tomasselli, A.G., Heinrikson, R.L. and Kezdy, F.J. (1991) J. Biol. chem. 266, 14554–14561.
[17] Pearl, L.H. (1990) Retroviral Proteases: Maturation and Morphogenesis, pp. 1–180, Macmillan Press, London.
[18] Wallin, M., Deinum, J., Goobar, L. and Danielson, U.H. (1990) J. Gen. Virol. 71, 1985–1991.
[19] Tomasselli, A.G., Hui, J.O., Adams, L., Chosay, J., Lowery, D., Greenberg, B., Yem, A., Deibel, M.R., Zurcher-Neely, H. and Heinrikson, R.L. (1991) J. Biol. Chem. 266, 14548–14553.
[20] Honer, B., Shoeman, R.L. and Traub, P. (1992) Cell Biol. Int. Rep. 16, 603–612.