Volume 237, number 1,2, 187-190 FEB 06301 September 1988

Processed enzymatically active protease ($p15^{gag}$) of avian retrovirus obtained in an *E. coli* system expressing a recombinant precursor ($Pr25^{lac\text{-}\Delta gag}$)

Juraj Sedláček, Petr Štrop*, František Kaprálek, Vladimír Pečenka, Vladimír Kostka*, Miloslav Trávníček and Josef Říman

Czechoslovak Academy of Sciences, Institute of Molecular Genetics and **Institute of Organic Chemistry and Biochemistry, 166 37 Prague 6, Czechoslovakia*

Received 5 July 1988; revised version received 28 July 1988

Processing proteases of avian and mammalian retroviruses cut the polyprotein precursors encoded by the retroviral genes into mature functional proteins. Retroviral processing proteases are still a rather poorly characterized group as to their relation to other proteases, specificity, and mechanism of enzymatic action. In avian retroviruses the generation of the processing protease itself comprises a processing cleavage event – the protease $p15^{gag}$ is cut off the carboxy-terminus of a *gag* polyprotein precursor, $Pr76^{gag}$. We report here that direct and efficient production of the avian retrovirus processing protease $p15^{gag}$ (required for structure-function studies and rational design of inhibitors) was obtained in an *E. coli* system, where massive expression of a size-reduced, recombinant precursor ($Pr25^{lac\text{-}\Delta gag}$) was accompanied by its structurally accurate processing.

Retroviral protease; Polyprotein precursor processing; Recombinant product accumulation; (Myeloblastosis associated virus, *E. coli*)

1. INTRODUCTION

In spite of the crucial importance of retrovirus-encoded proteases in the processing of retroviral protein components, very little is known about their enzyme specificity and structure. Their relatedness to aspartic proteases is indicated by short stretches of high sequence homologies, but the retroviral proteases are half their size [1]. The supply of retroviral proteases obtainable from natural material is often inadequately low. In natural formation, processing proteases of avian retroviruses, $p15^{gag}$, are synthesized as a part of a common polyprotein precursor $P76^{gag}$, where they map at the carboxy-terminus and form a boundary with $p12^{gag}$ [2-4]. The $p15^{gag}$ itself acts as the *gag*-specific processing protease and the intriguing question of how the first catalytic processing act arises has not yet been answered conclusively: a controversy persists as to whether initial amounts of the processing protease may come from autocatalysis of $Pr76^{gag}$ or from the imitating action of a specific host-cell protease. We attempted to obtain a preparative supply of the processing protease of an avian retrovirus (myeloblastosis associated virus, MAV [5]) from an *E. coli* sysem expressing an artificial (size-reduced and fused) precursor. Formation of $p15^{gag}$ from the precursor was found to proceed in the *E. coli* and the molecular entities formed were characterized; the mechanistic aspects of the findings will be discussed.

Correspondence address: J. Sedláček, Czechoslovak Academy of Sciences, Institute of Molecular Genetics, 166 37 Prague 6, Czechoslovakia

2. EXPERIMENTAL

The proviral DNA originated from a plasmid subclone of a bacteriophage clone [5] carrying the full-sized proviral DNA of MAV type 1. The expression plasmid constructed, pMG45, is related to pMG225 [6], a construct for the expression of another gene product (fusion calf prochymosin), and the actual construction was via removal of the previous insert, conversion of the linker, elimination of the *Bam*HI site in front of the *tac*

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

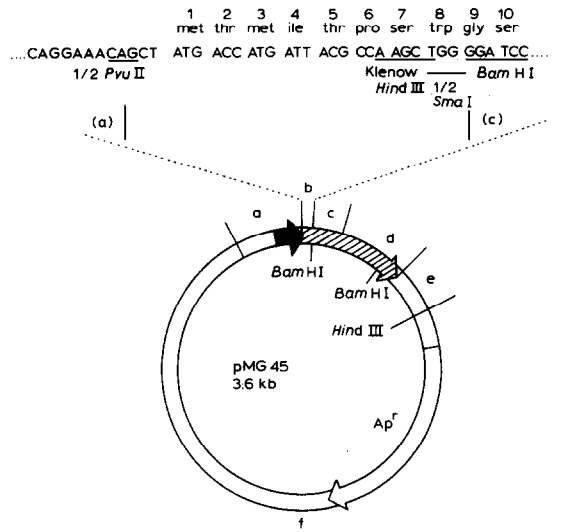

Fig.1. Expression plasmid pMG45. Segment map and (the extension, top) sequences in the *lac-gag* fusion region; the segment encoding $Pr25^{lac\text{-}\Delta gag}$ is shown by the hatched arrow. (a) pKK233-3-derived segment [6] containing the *tac* promoter/operator (black arrow); (b) pUC9-derived portion of the *E. coli lacZ* gene and linker; (c) proviral DNA encoding a portion of $p12^{gag}$; (d) proviral DNA encoding $p15^{gag}$; (e) proviral DNA of the *gag-pol* junction and a portion of the *pol* gene; (f) the pUC-series provided the rest of the vector part.

promoter-containing segment, and two-step insertion of 0.82 kb *Bam*HI-*Bam*HI-*Hin*dIII segment of the MAV proviral DNA. The resulting expression plasmid pMG45 (fig.1), a pUC-series [7] derived replicon, places the 0.82 kb *Bam*HI-*Bam*HI-*Hin*dIII insert of proviral DNA of MAV under transcription control of the *tac* promoter and translation control of the initiation signal of *E. coli* β-galactosidase. The insert contains codons for 65 amino acids antecedent to the *gag* p12/p15 boundary and codons for all the 124 amino acids of $p15^{gag}$, followed by a stop codon (plus a downstream *pol*-specific sequence lacking a signal for reinitiation in prokaryotes). The segment of *E. coli* DNA and linker in front of the proviral DNA adds five amino acids of β-galactosidase and three amino acids of the linker translate to the amino-terminus of the viral protein. The precursor thus encoded has been designated $Pr25^{lac\text{-}\Delta gag}$.

Host cells of *E. coli* strain MT [6] which provide endogenous repression by the F′-encoded $lacI^Q$ gene product, were used in constructing intermediate plasmids and pMG45. For expression, pMG45-harbouring *E. coli* MT were grown in the presence of lactose in a complete medium for 24 h at 37°C in flasks on a rotary shaker; the culture densities reached up to 4 g wet biomass per l. The cytoplasmic inclusions were collected from sonicated cell lysates, washed with 0.5% Triton X-100 [8], and dissolved in 9 M urea, 0.05 M Tris-HCl (pH 7.0) and 0.1% β-mercaptoethanol (omitted where specified). The preparations continued with dialysis against buffer containing either 0.02 M Tris-HCl (pH 7.1) or 0.03 M sodium acetate (pH 5.2) and always 0.03 M NaCl, 0.05% β-mercaptoethanol and 0.005% Triton X-100; after dialysis the pH was adjusted where specified and sediments were separated by centrifugation.

Assays of proteolytic activities were performed essentially as in [9]; each mixture contained in 0.1 ml final volume: 0.06 M sodium citrate (pH 5.0), 0.08% SDS, 200 μg bovine serum albumin substrate and 5 μg protein from different preparations of $p15^{gag}$.

Amino-terminal amino acid sequences were determined with an Applied Biosystems 470A protein sequencer.

Authentic viral $p15^{gag}$ protease was prepared as in [10].

3. RESULTS

Coding for most of the non-p15 moiety of the natural precursor, $Pr76^{gag}$, was eliminated in our *E. coli* system by constructing a size-reduced precursor that retained about 24 kDa of $Pr76^{gag}$ [2, 3], including the $p15^{gag}$ cut-off site [4]. In induced cultures, all *E. coli* cells harbouring pMG45 formed polar cytoplasmic inclusions that filled up to 24% of the cellular volume. The inclusions contained p15 as a major protein component: it co-migrated exactly with a reference sample of authentic viral p15 protease in SDS-PAGE (fig.2). Another prominent band migrated as was expected for $Pr25^{lac\text{-}\Delta gag}$.

Contamination of the product by a great variety of less abundant proteins was probably a consequence of scavenging connected with the cross-linking capacity of the four cysteine residues in the p12-specific part of the precursor. Separation of the proteins solubilized with urea, very remarkable from a practical point of view, was obtained upon removal of the urea by dialysis. Heavy precipitates formed, composed of all of the inclusion proteins, but partially depleted of p15. The solutions above the precipitates ('supernatants' in fig.2) contained, by contrast, nearly homogeneous p15, with only traces of some contaminants. Typical recoveries per g wet biomass were: 180 mg washed wet inclusions; 25 mg total protein of inclusions; 6.5 mg protein in supernatant after dialysis, most of it p15; and, as based on activity measurements, 4 mg enzymatically active p15 protease.

Obtainment of active p15 was favoured at neutral pH and in the presence of β-mercaptoethanol and/or Triton X-100. The preparations of p15 from *E. coli* displayed a proteolytic activity of the same specificity as did viral p15 protease. A characteristic pattern of the cleavage of bovine serum albumin is shown in fig.3. Precursor

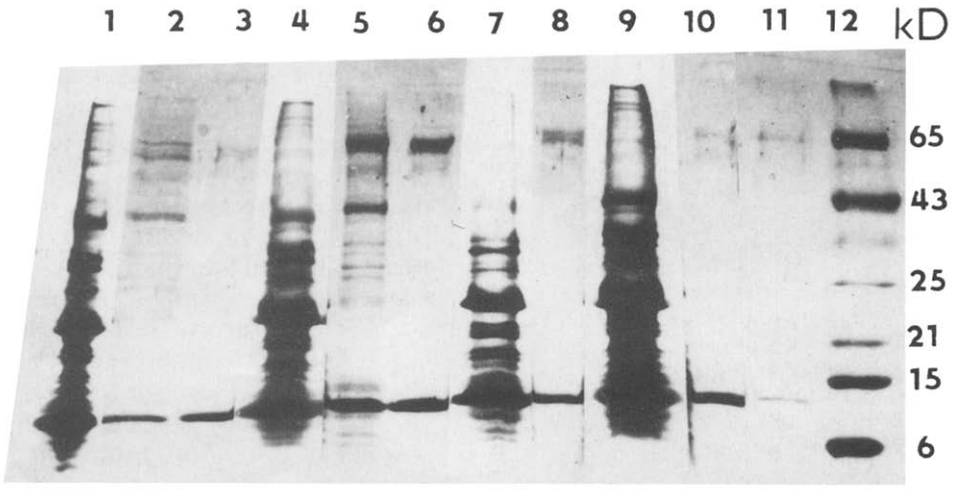

Fig.2. Protein composition of inclusions and preparations of p15. Coomassie blue stain; SDS, 10–25% polyacrylamide gel electrophoresis (SDS-PAGE); sample buffer containing β-mercaptoethanol was used throughout. Lanes: (1) dissolved inclusions, no further treatment; (2) inclusions dissolved without β-mercaptoethanol and dialysed at pH 7.1, supernatant; (3) dissolved without β-mercaptoethanol, dialysed at pH 7.1, adjusted to pH 5.2, supernatant; (4) dissolved without β-mercaptoethanol, dialysed at pH 5.2, sediment; (5) dialysed at pH 7.1, supernatant; (6) dialysed at pH 7.1, adjusted to pH 5.2, supernatant; (7) dialysed at pH 7.1, sediment; (8) dissolved without β-mercaptoethanol, dialysed at pH 5.2, supernatant; (9) dissolved without β-mercaptoethanol, dialysed at pH 5.2, sediment; (10) dialysed at pH 5.2, supernatant; (11) authentic viral $p15^{gag}$; (12) molecular mass standards.

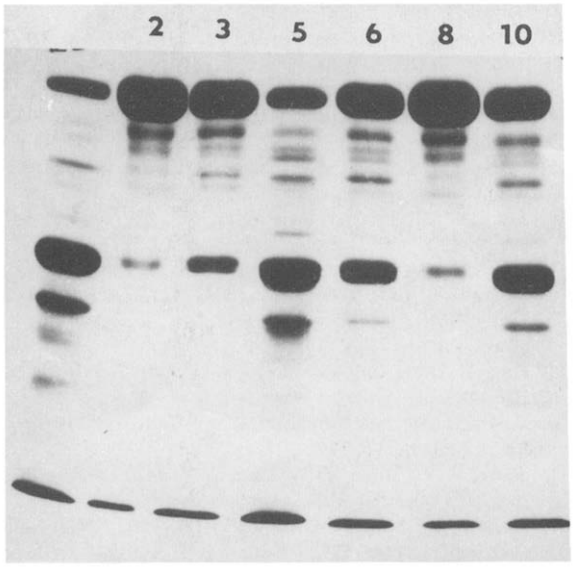

Fig.3. Enzymatic activity of different preparations of $p15^{gag}$. SDS-PAGE (10% gel) of the cleavage products of bovine serum albumin treated with $p15^{gag}$. Left lane: authentic viral protease; other lanes: supernatants from dialysis of solubilized inclusion proteins prepared as detailed in the legend to fig.2 and numbered correspondingly.

$Pr25^{lac\text{-}\Delta gag}$ electroeluted from gels did not display enzyme activity.

Accuracy of the processing cleavage in the *E. coli* expression system and identity of the recombinant product were established: a uniform Leu-Ala-Met-Thr-Met-Glu-His amino-terminal sequence was determined in the major HPLC-separated protein (fig.4A,B; peak 3). It corresponded both with the downstream part of the p12/p15 boundary of homologous $Pr76^{gag}$ of Rous sarcoma virus [4] and with the amino acid sequence determined in the authentic MAV protein [11]. This allows one to conclude that Ser(73)/Leu(74) cleavage had occurred, which represents genuine processing. Analysis of the amino acid composition (not listed) confirmed the integrity of the recombinant product. Two amino-terminal sequences identified in the HPLC-separated proteins (fig.4A, peaks 1,2) were an effect of Met(1)/Thr(2) cleavage in the β-galactosidase-specific part and of Arg(45)/Lys(46) cleavage in the p12-specific part of the precursor.

4. DISCUSSION

Processing schemes for *gag*-encoded polypeptides of avian retroviruses have to explain how in-

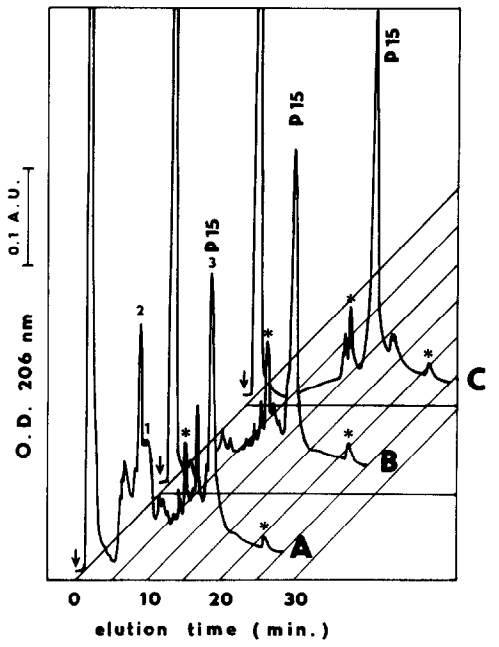

Fig.4. Reverse-phase HPLC and N-terminal sequence analysis of the recombinant products. On an RPSC Ultrapore column was applied (A) material obtained as described in the legend to fig.2 for lane 5; (B) as described above for lane 6; (C) authentic viral protease; the elution gradients were 0–50% acetonitrile/25 min. Asterisks indicate elution of buffer components.

itial amounts of $p15^{gag}$ might be generated. Schemes have been proposed wherein $p15^{gag}$ is removed from $Pr76^{gag}$ by the action of a specific host-cell protease or, alternatively, $Pr76^{gag}$ has the capacity to cleave its own p12/p15 boundary or that of a neighbouring precursor molecule [2,3]; neither alternative has been established or ruled out conclusively (cf. [12]). The first component possessing the specific catalytic processing activity in our particular system has not been rigorously identified either. The most obvious candidate is, however, the Arg(45)/Lys(46) cleavage derivative of the primary translate; it represents a molecular entity differing from the mature $p15^{gag}$ only by an extension of 28 amino acids at the amino terminus. Active forms of aspartic proteases containing extensions of tens of amino acids at the amino-termini are not uncommon. Similar catalytically active intermediates of the processing of the natural precursor have possibly escaped attention so far and deserve to be sought for deliberately (cf. [13]). The experimental data obtained earlier with the full-sized precursor in *E. coli* [12] are fully compatible with such a mechanism and they could have been interpreted accordingly had signalling information on size-reduced entities been available.

The *E. coli* system presented can be very efficient in providing a preparative supply and has already allowed the above structure determinations and in vitro measurements to be performed; it also offers possibilities of straightforward use in the expression of site-mutated proteins, as a tool for the elucidation of structure-function relations.

REFERENCES

[1] Pearl, L.H. and Taylor, W.R. (1987) Nature 329, 351–354.
[2] Dickson, C., Eisenman, R., Fan, H., Hunter, E. and Teich, N. (1984) in: RNA Tumor Viruses 1 (Weiss, R. et al. eds) pp. 528–547, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[3] Dickson, C., Eisenman, R. and Fan, H. (1985) in: RNA Tumor Viruses 2 (Weiss, R. et al. eds) pp. 136–146, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[4] Schwartz, D.E., Tizard, R. and Gilbert, W. (1983) Cell 32, 853–869.
[5] Perbal, B., Lipsick, J.S., Svoboda, J., Silva, R.F. and Baluda, M.A. (1985) J. Virol. 56, 240–244.
[6] Sedláček, J., Fábry, M., Kašpar, P., Zadražil, S. and Kaprálek, F. (1986) Biotechnol. Bioind. 6, 17–19.
[7] Vieira, J. and Messing, J. (1982) Gene 19, 259–268.
[8] Marston, F.A.O., Lowe, P.A., Doel, M.T., Schoemaker, J.M., White, S. and Angal, S. (1984) Biotechnology 2, 800–804.
[9] Dittmar, K.J. and Moelling, K. (1978) J. Virol. 28, 106–118.
[10] Trávníček, M. and Říman, J. (1980) Biochem. Biophys. Res. Commun. 96, 1768–1777.
[11] Sauer, R.T., Allen, D.W. and Niall, H.D. (1981) Biochemistry 20, 3784–3791.
[12] Mermer, B., Malamy, M. and Coffin, J.M. (1983) Mol. Cell. Biol. 3, 1746–1758.
[13] Trávníček, M., Malý, A., Šulová, A. and Říman, J. (1982) Fol. Biol. 28, 145–159.