The EMBO Journal vol.12 no.7 pp.2977 – 2987, 1993

Polyprotein processing in African swine fever virus: a novel gene expression strategy for a DNA virus

Carmen Simón-Mateo, Germán Andrés and Eladio Viñuela[1]

Centro de Biología Molecular 'Severo Ochoa' (CSIC-UAM), Facultad de Ciencias, Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain

[1]Corresponding author

Communicated by E.Viñuela

This report shows that African swine fever virus (ASFV)—a large DNA-containing virus—synthesizes a polyprotein to produce several of its structural proteins. By immunoprecipitation analysis, we have found that ASFV polyprotein is a 220 kDa myristoylated polypeptide (pp220) which, after proteolytic processing, gives rise to four major structural proteins: p150, p37, p34 and p14. Processing of the ASFV polyprotein takes place at the consensus sequence Gly-Gly-X and occurs through an ordered cascade of proteolytic cleavages. So far, polyprotein processing as a mechanism of gene expression had been found only in positive-strand RNA viruses and retroviruses. According to the results presented here, ASFV is the first example of a DNA virus that synthesizes a polyprotein as a strategy of gene expression.

Key words: ASFV/polyprotein/proteolytic processing

Introduction

The replication of many animal and plant viruses is dependent on proteolytic processing by virus-encoded proteinases that are involved in functions such as separation of structural and non-structural proteins, generation of specific enzymes, coordinated assembly of the virion, and maturation (for a review see Kräusslich and Wimmer, 1988). In general, proteolytic processing by viral proteinases can be considered as playing two general roles. Their first role is as a mechanism for gene expression in which a viral polyprotein (a precursor polypeptide) is proteolytically processed to yield the mature proteins. Synthesis of such polyproteins has only been described for positive-strand RNA viruses and retroviruses. In these viruses, whose genomic RNAs are of the same polarity as the viral mRNAs, polyprotein processing as a mechanism of gene expression has been explained to comply with the apparent rule that eukaryotic translation systems only handle monocistronic messengers (Kozak, 1983). Furthermore, the synthesis of a polyprotein in these small RNA viruses allows for 'genetic economy' since several proteins are produced as the result of a single transcription/translation event. A second general role for proteolytic processing is that related to the viral morphogenesis in the activation of structural proteins to form mature virus particles. This proteolytic processing also takes place during the replication of positive-strand RNA viruses and, in the case of DNA viruses, has been found to occur only in the maturation of single proteins and not in polyproteins. Adenovirus (Bhatti and Weber, 1979), herpesvirus (Liu and Roizman, 1991) and poxvirus (Katz and Moss, 1970) are examples of large DNA viruses showing this characteristic. African swine fever virus (ASFV) is also among the complex DNA viruses in which proteolytic processing of proteins has been described during viral replication (López-Otín *et al.*, 1988, 1989).

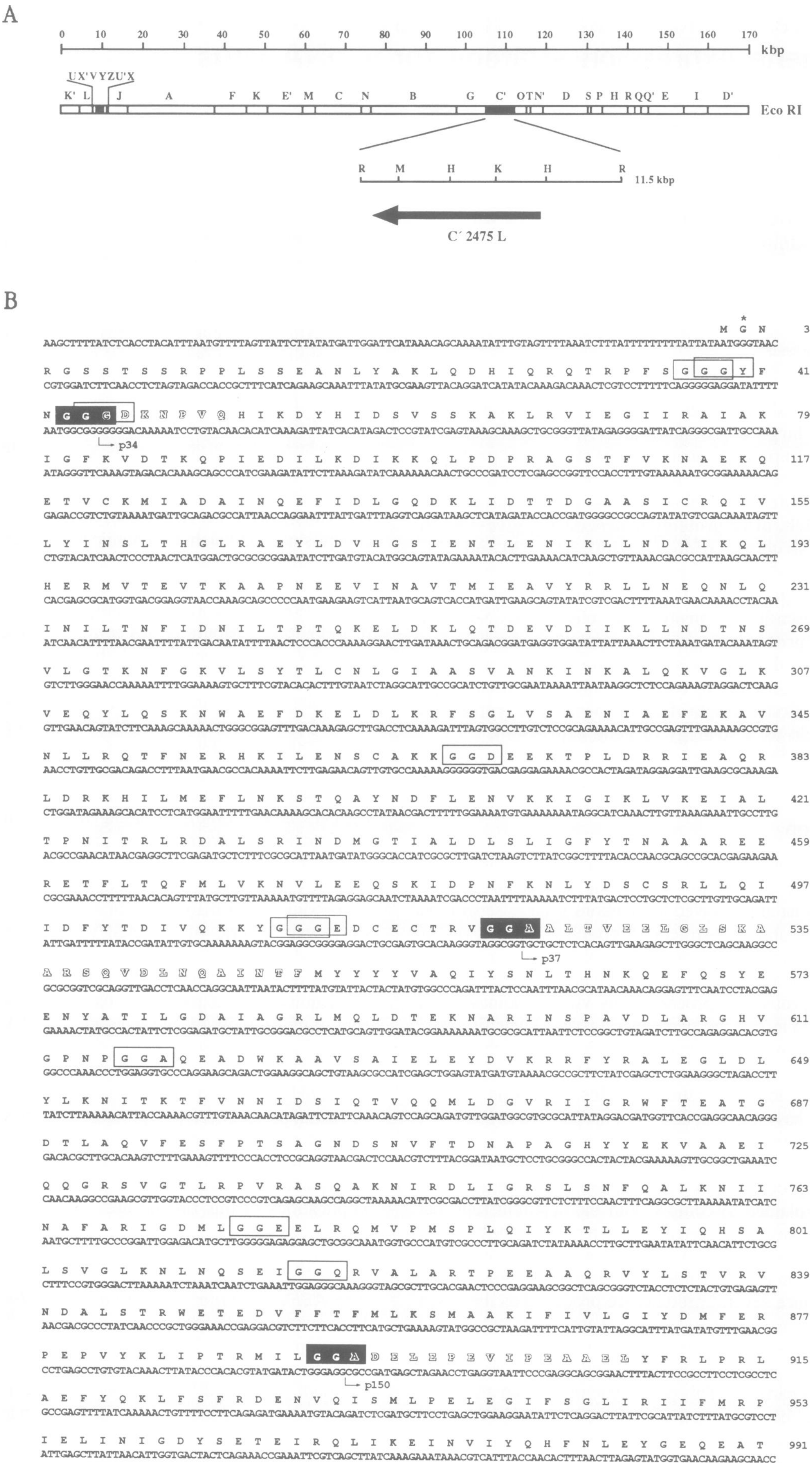
ASFV is a large DNA-containing virus responsible for a highly lethal disease of domestic pigs (for review see Viñuela, 1987; Costa, 1990); it also infects soft ticks (*Ornithodorus* sp.), which act as vectors for the virus. Animal viruses with very large DNA genomes include families of icosahedral viruses (Herpesvirus and Iridovirus) and brick-shaped viruses (Poxvirus). However, ASFV does not fit well into any of these groups since, although its genomic structure is similar to that of poxviruses (hairpin loop structures and terminal inverted repetitions at the DNA ends: Sogo *et al.*, 1984; González *et al.*, 1986), its icosahedral morphology is similar to that of iridoviruses (Moura Nunes *et al.*, 1975). The complexity of ASFV, with a genome 170 kb long, is evident from the fact that about 100 virus-induced polypeptides, with relative molecular masses ranging from 10 to 220 kDa, have been found by two-dimensional gel electrophoresis analysis of ASFV-infected cells (Santarén and Viñuela, 1986). In addition, the ASFV particle contains more than 50 proteins with relative molecular masses ranging from 10 to 150 kDa (Esteves *et al.*, 1986), including the enzymatic machinery required for the synthesis and processing of early mRNA (Salas *et al.*, 1981).

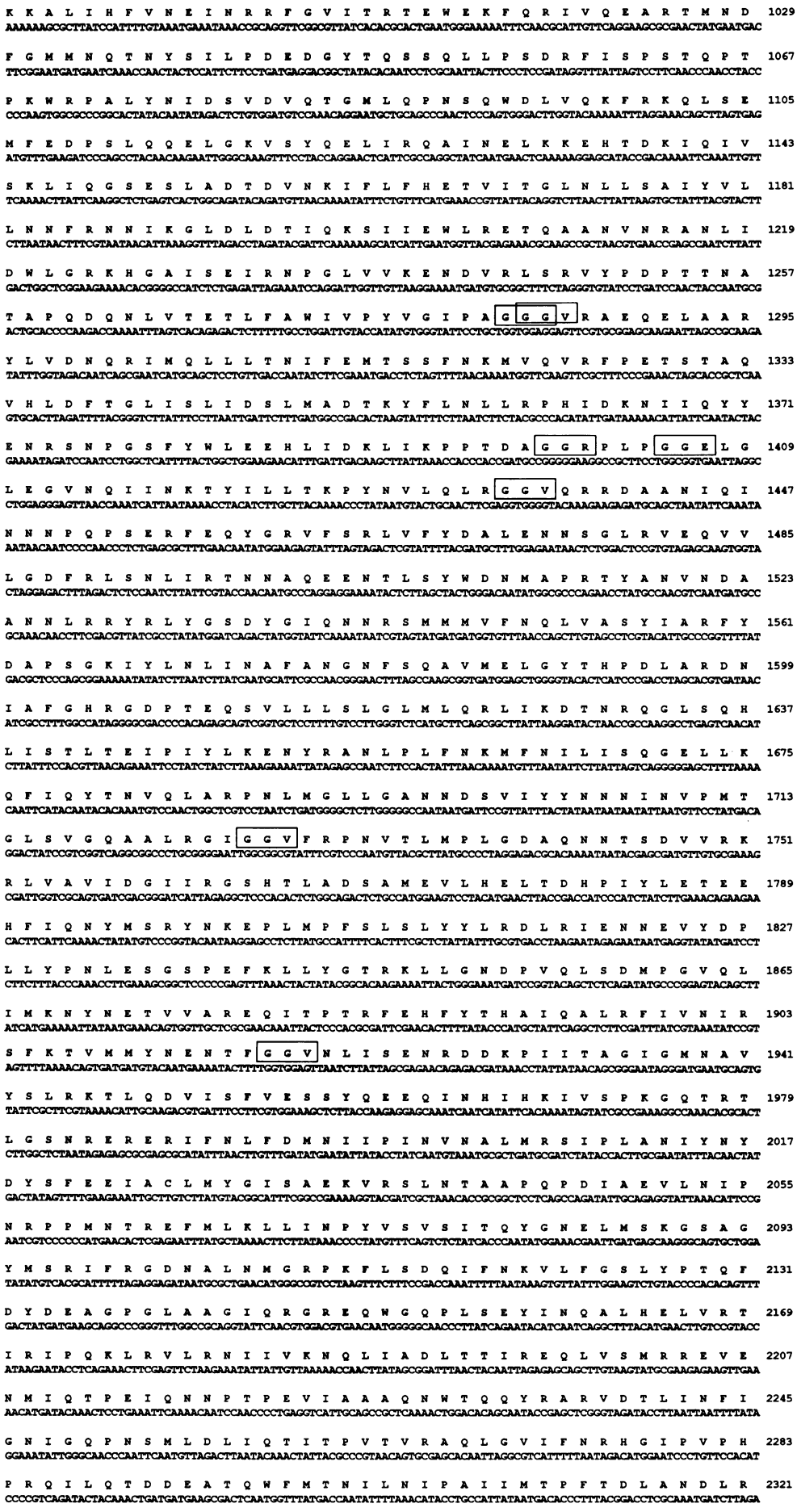
In this report we have extended the study of proteolytic processing in the maturation of some ASFV proteins. We show that several of the ASFV structural proteins are produced in the form of a polyprotein that is post-translationally processed. In complex DNA viruses, the regulation of gene expression is usually achieved by temporal control of transcription and translation. ASFV regulates gene expression by temporal control of transcription (Salas *et al.*, 1986; Almazán *et al.*, 1992) and, according to the data presented in this paper, also synthesizes some of the structural proteins as a polyprotein. Therefore, ASFV would be the first DNA-containing virus described that shares this strategy of gene expression with positive-strand RNA viruses and retroviruses. The possible functional and evolutionary implications of this finding are discussed.

Results

The ASFV genome encodes a polyprotein

Previous studies (López-Otín *et al.*, 1989) showed that some ASFV proteins (p150, p37 and p34) are synthesized as precursors of higher molecular weight. These proteins were purified from ASFV particles and their NH_2-terminal sequences were determined in order to map their genes in the ASFV genome. Mixtures of oligodeoxynucleotides derived from the amino acid sequences were used to probe


A
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170
kbp
UX'VYZU'X
K' L J A F K E' M C N B G C' OTN' D S P H R Q Q' E I D'
Eco RI
R M H K H R
11.5 kbp
C´2475 L
B
p34
p37
p150



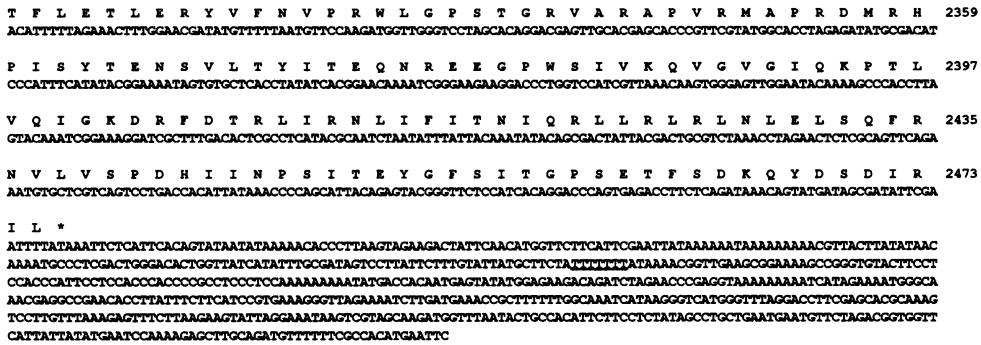


Fig. 1. Nucleotide and predicted amino acid sequence of ASFV polyprotein. (**A**) Localization in the ASFV genome of a 8148 bp region contained in the *Eco*RI C′ restriction fragment. The position and orientation of the ORF C′2475L are indicated by an arrow. (**B**) Sequence of the 8148 bp *Hin*dIII−*Eco*RI fragment containing the ORF C′2475L. The numbers on the right margin refer to amino acid position. The 7T motif after the stop codon is underlined. The asterisk indicates the amino acid covalently bound to the myristic acid. The NH_2-terminal sequences determined for proteins p150, p37 and p34 are indicated in shadowed letters. Potential cleavage sites (Gly-Gly-X) are indicated by open boxes and those involved in the polyprotein processing are indicated in black boxes.

the viral genome, and all the putative genes were found to map within the *Eco*RI C′ restriction fragment. After partial sequencing of this fragment, the NH_2-terminal sequences were found to be preceded by different stretches of additional amino acid residues, indicating that these proteins became mature after proteolytic processing. In the present study, we have determined the complete nucleotide sequence of the DNA fragment in which the NH_2-terminal sequences of proteins p150, p37 and p34 were mapped (Figure 1). There were some nucleotide differences between the nucleotide sequence shown here and that previously reported for the gene encoding protein p37 (López-Otín *et al.*, 1988), contained in this DNA fragment. These nucleotide modifications revealed the existence of a single open reading frame (ORF) encoding a polypeptide 2475 amino acids long. This ORF, named C′2475L (in this nomenclature, the letter indicates the *Eco*RI fragment in which translation is started, the number indicates the number of amino acids of the encoded protein and the L indicates leftward reading), contained the first AUG (nucleotide 106) in a favourable context for initiation of translation (Kozak, 1983) and was preceded by an in-frame stop codon (nucleotide 31). Moreover, a 7T motif recently suggested as a signal for transcription termination of ASFV genes (Almazán *et al.*, 1992), was found 173 nucleotides downstream of the C′2475L stop codon. ORF C′2475L contained the NH_2-terminal sequences determined for proteins p150, p37 and p34, indicating that all of them are synthesized as a polyprotein and then proteolytically processed. The peptide bonds cleaved to give rise to the mature proteins were those located after the second Gly residue of the previously reported consensus sequence Gly-Gly-X (López-Otín *et al.*, 1989), which is also recognized in adenovirus structural proteins and in some cellular proteins such as polyubiquitin. In addition to the three cleavage sites used for the NH_2-terminal maturation of proteins p150, p37 and p34, 16 other putative cleavage sites were present in the ASFV polyprotein. A computer search for ORF C′2475L was performed and no significant similarity to any entry in the databases was found.

The ASFV polyprotein is a myristoylated, 220 kDa polypeptide

To characterize the precursor polyprotein, polyclonal antisera against different regions of the C′2475L protein were obtained. The different polypeptides used as immunogens in rabbits were produced by expressing them in bacteria from the appropriate recombinant plasmids, as described in Materials and methods. A schematic representation showing the localization of the C′2475L regions used as immunogens is shown in Figure 5B.

Extracts of ASFV-infected cells were metabolically labelled with [^{35}S]methionine/cysteine for 30 min at 12 h.p.i., since previous results had shown that the precursor form of protein p150 is a late protein expressed after DNA replication (Andrés *et al.*, 1993). Immunoprecipitation with antisera against the different polyprotein regions allowed us to detect a protein with an apparent molecular mass of ~220 kDa (named pp220) (Figure 2, pulse lanes). The difference between the apparent (220 kDa) and the predicted (280 kDa) molecular masses of the polyprotein could be explained by either abnormal migration and/or deviations in the molecular weight estimation by SDS−PAGE in this size range (Hames, 1990).

Since the amino acid sequence of the ORF C′2475L showed the existence of several potential glycosylation sites and since an ASFV-induced 220 kDa glycoprotein has been described previously (Del Val and Viñuela, 1987), we investigated the possibility that polyprotein pp220 would be a glycoprotein. For this purpose, ASFV-infected cell extracts metabolically labelled with [^{14}C]glucosamine were immunoprecipitated with an antiserum specific for polyprotein pp220. Although a faint band of ~220 kDa was detected after long exposures of the fluorograph (data not shown), it proved to be the result of the metabolic conversion of the radioactive sugar to amino acids, since other non-glycosylated viral proteins were labelled in the same conditions as described by Alcamí *et al.* (1992). Moreover, polyprotein pp220 was resistant to treatment with O- and N-glycosidases and was not retained in lectin affinity columns (data not shown). Taken together, these results indicate that polyprotein pp220 is not glycosylated.

Since the proposed NH_2-terminal sequence of the polyprotein contained the consensus recognition sequence (MGXXXS/T/A) for cotranslational modification by N-myristoyltransferase (Figure 1), we also investigated whether the ASFV polyprotein could be covalently bound to myristic acid. In fact, a virus-induced myristoylated protein of 220 kDa has previously been reported (Aguado *et al.*, 1991). As shown in Figure 3, labelling of ASFV-infected cells with

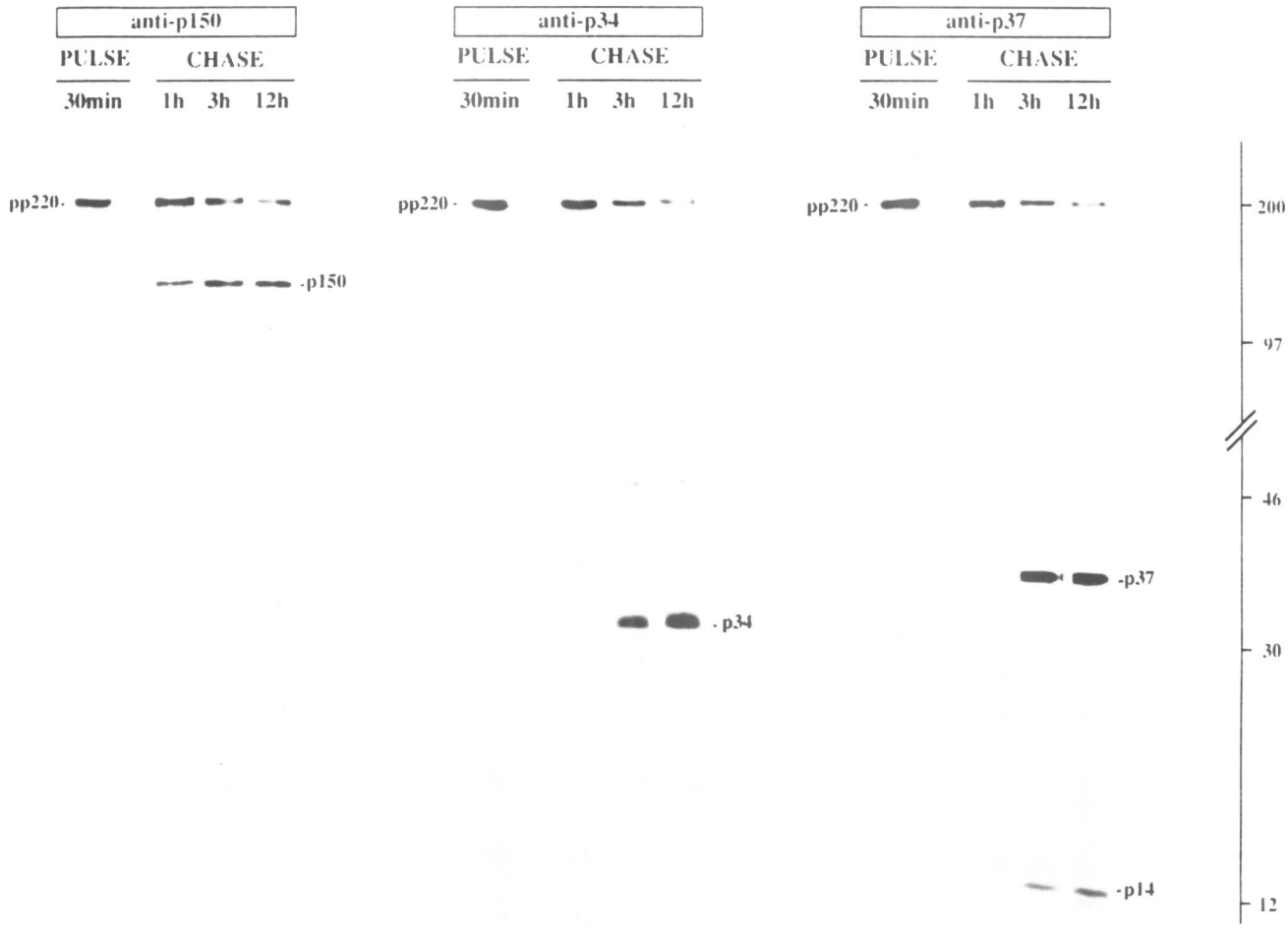


Fig. 2. Proteins derived from ASFV polyprotein processing. Infected Vero cells were pulse-labelled with [^{35}S]methionine/cysteine for 30 min (pulse) at 12 h.p.i. and chased for 1, 3 and 12 h (chase). Immunoprecipitation with p150 region (anti-p150), p34 region (anti-p34) and p37 region (anti-p37) antisera are shown. The proteins detected in both pulse and chases are indicated. Molecular masses in kilodaltons are indicated on the right. In the case of the immunoprecipitations with the anti-p34 and anti-p37 antisera the lower part of the fluorograph needed longer autoradiographic exposures (40-fold) to allow detection of the proteins with lower methionine and cysteine content than polyprotein pp220.

[^{3}H]myristic acid followed by immunoprecipitation analyses with the different antisera against ASFV polyprotein showed that the pp220 polyprotein was myristoylated. This finding implies that the initiator methionine is removed from the C′2475L protein and that the NH_2-terminal glycine residue (at position 1) is the acceptor residue of the myristic moiety.

Characterization of the mature proteins produced by proteolytic processing of the ASFV polyprotein

To identify the proteins derived from the ASFV polyprotein, infected cells were pulse-labelled with [^{35}S]methionine/cysteine for 30 min at 12 h.p.i. and chased for 1, 3 and 12 h in the presence of a 10-fold excess of methionine/cysteine. Figure 2 shows an immunoprecipitation analysis with the different antisera raised against polyprotein pp220. In the pulse, the antiserum specific for the p150 region recognized polyprotein pp220, the level of which decreased during the chases coinciding with the accumulation of a 150 kDa protein (named p150) (Figure 2, anti-p150). The cleavage to produce protein p150 was detected in the first hour of chase and continued in the following chase periods. During the pulse, the antiserum against the p34 region recognized polyprotein pp220; the level of this protein declined from 3 h of chase in parallel with an increase of a 34 kDa protein (named p34) (Figure 2, anti-p34). The antiserum against the p37 region also recognized polyprotein pp220 following the pulse-labelling (Figure 2, anti-p37) and, interestingly, during the chase periods, besides a 37 kDa protein (named p37), a 14 kDa protein (named p14) was also detected. Detection of proteins p37 and p14 took place at the same time as protein p34, 3 h after the pulse. The finding that a 14 kDa protein

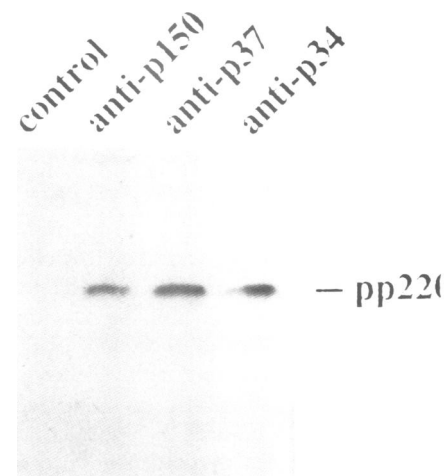


Fig. 3. Myristoylation of ASFV polyprotein. Immunoprecipitation analyses of ASFV-infected cells labelled with [^{3}H]myristic acid for 3 h at 12 h.p.i. were carried out using the control, p150 region (anti-p150), p37 region (anti-p37) and p34 region (anti-p34) antisera (see Figure 5).

was detected with this antiserum strongly suggests that the putative cleavage sequence Gly-Gly-X located at amino acid residues 368–369 (Figure 1) is also involved in polyprotein processing. In fact, this is the only putative cleavage site whose recognition might account for the detection of a 14 kDa protein with the antiserum against the p37 region (see Figure 5). The fact that the predicted and apparent molecular weights of the proteins derived from the polyprotein processing are similar, might indicate that the proposed proteolytic cleavages serve for the maturation of both the NH_2-termini of one protein and the COOH-termini of the preceding one. However, other proteolytic cleavages at the COOH-termini of the proteins releasing small peptides cannot be ruled out at present.

In summary, proteolytic processing of the ASFV polyprotein gives rise to four mature proteins: p150, p37, p34 and p14. It can also be concluded that at least four of the 19 putative cleavage sites are cleaved. A computer analysis of the primary sequence and secondary structure around the Gly-Gly-X sequences present in the polyprotein did not reveal any motif that could account for the restrictions in the processing of only some of the putative cleavage sites.

ASFV polyprotein processing occurs through an ordered cascade of cleavages

Interestingly, after overexposing the fluorograph corresponding to the immunoprecipitation analysis of a pulse–chase experiment, it was also possible to detect some intermediate precursors of the polyprotein processing. Thus, as shown in Figure 4A, two additional proteins were recognized by the antisera against the p34 and p37 regions. A 90 kDa protein (named pp90) was specifically detected with both antisera, suggesting that pp90 contains proteins p37, p34 and p14. The antiserum against the p37 region recognized, besides pp90, a 55 kDa protein (named pp55), suggesting that pp55 contains p37 and p14 (Figure 4A). It is interesting to note that pp90 appeared early after the pulse (Figure 4B, 30 min), at the same time as p150 (Figure 2, anti-p150), and was not processed until 3 h of chase. In contrast, pp55 was detected 3 h after the pulse, at the same time as p34 (Figure 2, anti-p34), and appeared to be rapidly processed to generate proteins p37 and p14 (Figure 4, anti-p37). The fact that anti-p37 and anti-p34 antisera did not recognize a 90 kDa protein as myristoylated (Figure 3), suggests that the cleavage site at positions 44–45 is recognized in the primary proteolytic event to give rise to preprotein pp90 lacking the myristic moiety. In addition, this cleavage to originate the NH_2-terminus of p34 should produce a myristoylated peptide 44 amino acids long (lacking methionine and cysteine); however, this product could not be detected in the labelling experiment with [^{3}H]myristic acid, probably due to either its small size or its degradation.

The results presented in Figures 2, 3 and 4 allow us to propose a scheme of the proteolytic cleavages involved in the generation of the mature proteins contained in polyprotein pp220 (Figure 5). The initial event in the proteolytic cascade is the recognition of the cleavage sites at positions 893–894 and 44–45. These cleavages take place early (30 min) after the synthesis of polyprotein pp220 to separate the mature protein p150 and preprotein pp90. The second cleavage takes place at positions 368–369 within preprotein pp90 and is later than the first one, occurring 2–3 h after the synthesis of polyprotein pp220. The result is the appearance of the

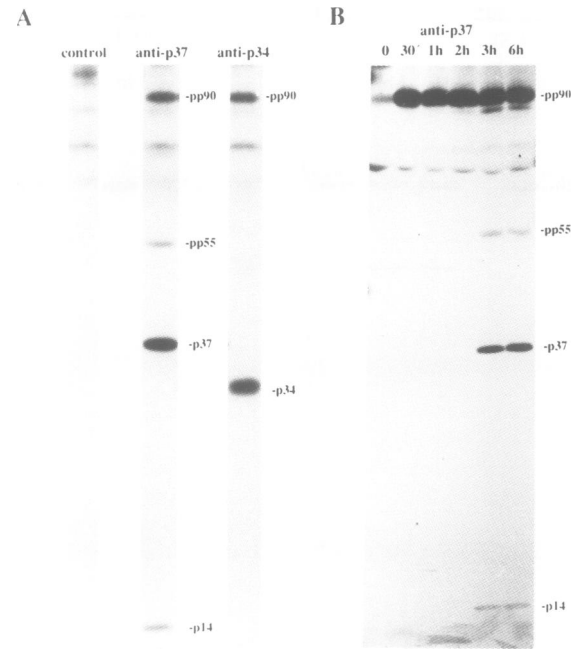


Fig. 4. Intermediate precursors in ASFV polyprotein processing. (**A**) Immunoprecipitation analysis of ASFV-infected cells labelled for 30 min at 12 h.p.i., and chased during 6 h using control, p37 region (anti-p37) and p34 region (anti-p34) antisera. (**B**) Immunoprecipitation analysis of ASFV-infected cells labelled for 30 min (0) at 12 h.p.i. and chased for 30 min, 1 h, 2 h, 3 h and 6 h using anti-p37 antiserum. The proteins detected with the antisera are indicated.

mature protein p34 and preprotein pp55. The final processing step is the cleavage at positions 522–523 within preprotein pp55 to originate p37 and p14. In summary, mature proteins are derived from the ASFV polyprotein as the result of an ordered recognition of the Gly-Gly-X sites to be cleaved.

Polyprotein processing gives rise to four major structural proteins

To determine whether the proteins generated from polyprotein pp220 are structural ones, immunoprecipitation analyses of ^{35}S-labelled, highly purified ASFV particles were performed with antisera against the different regions of polyprotein pp220. As shown in Figure 6A, the antisera recognized the mature proteins derived from the polyprotein in ASFV particles, indicating that all of them are incorporated in the viral particle. Analysis of highly purified virus particles by two-dimensional gel electrophoresis revealed that the proteins derived from the processing of the polyprotein are among the major structural proteins (Figure 6B). The positions of p150, p37, p34 and p14 were determined after analysing the immunoprecipitated proteins by two-dimensional gel electrophoresis (data not shown).

Next, we investigated the intracellular localization of polyprotein pp220 and its derived mature proteins. For this, indirect immunofluorescence analyses were carried out in ASFV-infected cells using antisera against polyprotein pp220. As shown in Figure 7, the immunofluorescence signal detected with all the antisera was mainly located in discrete cytoplasmic areas close to the nucleus corresponding to the viral factories where viral morphogenesis takes place (Sanz *et al.*, 1985). This finding, as well as the fact that all the mature proteins were found to be major structural

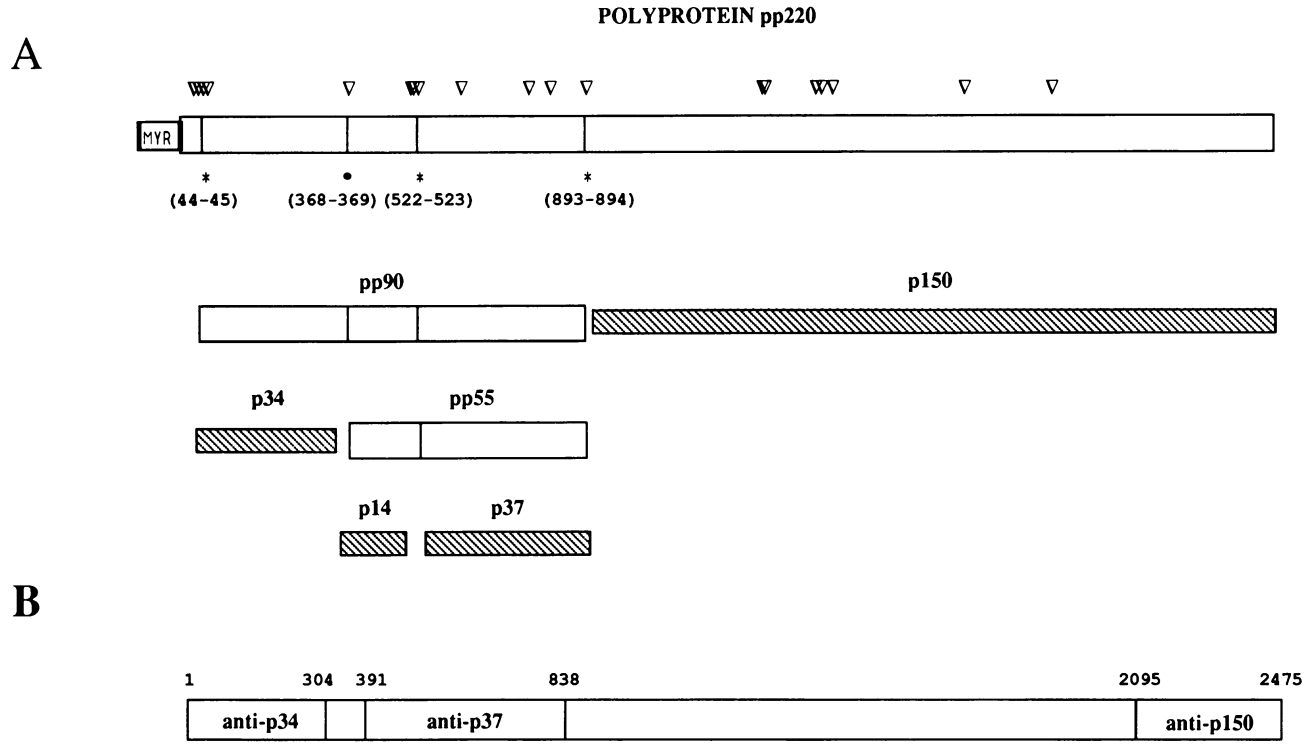


Fig. 5. Schematic representation of the cascade of proteolytic cleavages occurring during ASFV polyprotein processing. (**A**) Triangles indicate the position of the Gly-Gly-X sites in polyprotein pp220. The cleavage sites recognized are indicated with either * (those determined by NH_2-terminal microsequencing) or · (those deduced from immunoprecipitation of pulse−chase experiments). Myristoylation in polyprotein pp220 is indicated by 'MYR'. Precursors (pp220, pp90 and pp55) are represented by empty boxes. Mature proteins (p150, p37, p34 and p14) are represented by shadowed boxes and the name is indicated. (**B**) Scheme of the polyprotein with the position of the regions used to obtain the antisera against different regions of the polyprotein (anti-p150, anti-p37 and anti-p34). The numbers of the amino acids spanning the regions are indicated.

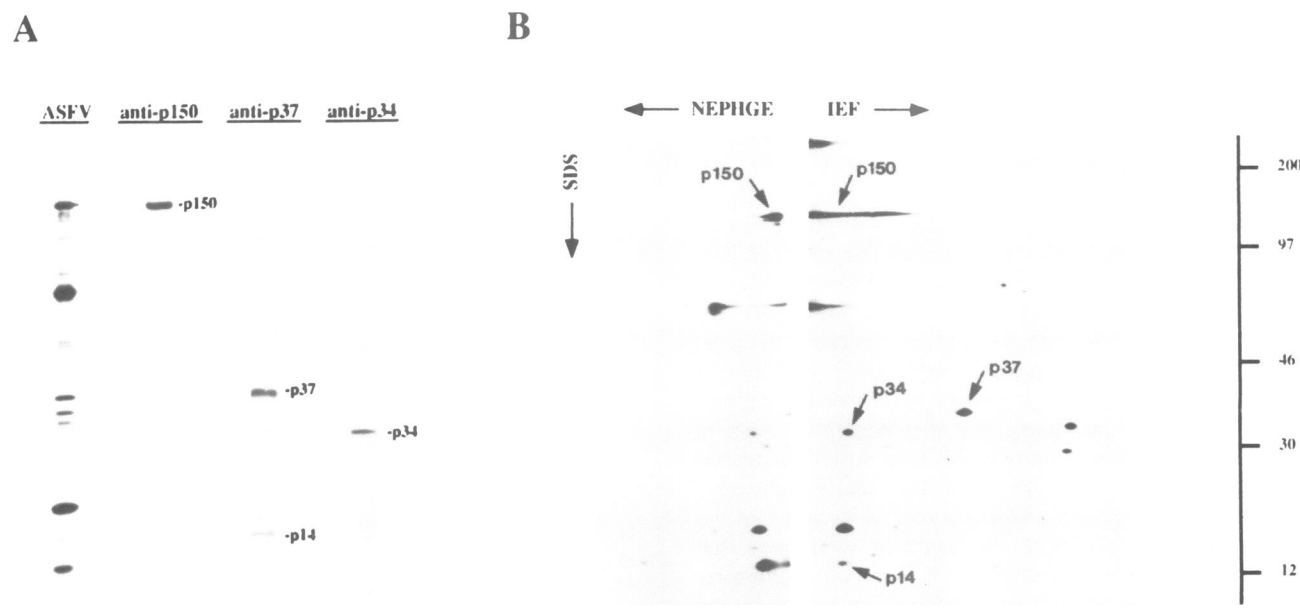


Fig. 6. Detection of proteins derived from ASFV polyprotein in viral particles. (**A**) Highly purified ASFV particles labelled with [^{35}S]methionine/cysteine were immunoprecipitated with antisera against the p150 region (anti-p150), p37 region (anti-p37) or p34 region (anti-p34). An electrophoretic profile of ASFV structural proteins is shown (ASFV). (**B**) Two-dimensional gel electrophoresis analysis of ^{35}S-labelled ASFV particles. The left part shows a NEPHGE (pH > 7) gel and the right part an IEF (pH 4.3−7.3) gel. The positions of proteins p150, p37, p34 and p14, determined by immunoprecipitation (not shown), are indicated.

components, suggests that the polyprotein processing in ASFV-infected cells might be linked with some step of the viral morphogenesis, as it occurs with the maturation of structural polyproteins in RNA viruses (Hellen and Wimmer, 1992) and of single proteins in some DNA viruses (Katz and Moss, 1970; Bhatti and Weber, 1979).

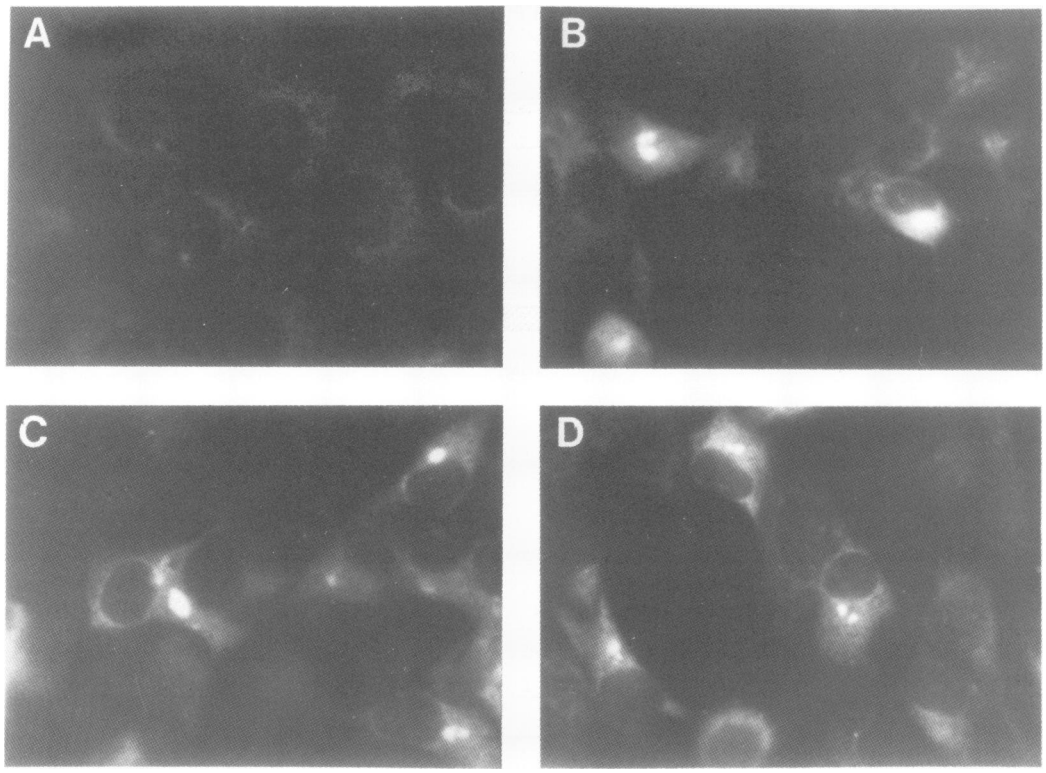


Fig. 7. Immunofluorescence detection of ASFV polyprotein in infected cells. Vero cells infected with ASFV were fixed/permeabilized at 12 h.p.i. and incubated in the presence of control (**A**), p150 region (**B**), p37 region (**C**) and p34 region (**D**) antisera. The cells were subsequently stained with a rhodamine-conjugated goat anti-rabbit secondary antibody.

Discussion

Until now, proteolytic processing in large DNA viruses (adenovirus, herpesvirus and poxvirus) was thought to be limited to the maturation of single proteins during viral assembly, since the regulation of gene expression is usually achieved by temporal control of transcription and translation. In this report, we have provided evidence that ASFV also uses proteolytic processing of a polyprotein as a mechanism for gene expression, sharing this strategy with RNA viruses (some positive-strand RNA viruses and all retroviruses). The synthesis of a polyprotein by ASFV could serve to cluster genes whose products (p150, p37, p34 and p14) are somehow related, allowing them to be synthesized as the result of a single transcription/translation event and, if necessary, allowing their transport to a specific cellular compartment. Furthermore, this mechanism of protein synthesis would determine that the four ASFV structural proteins are produced in equimolar amounts, which could be necessary to establish a certain ordered structure in the viral particle. In this sense, it should be mentioned that picornavirus capsids are composed of essentially equimolar amounts of four nonidentical polypeptides produced by the proteolytic processing of polyprotein P1 (Hellen and Wimmer, 1992).

Possible role of the ASFV polyprotein in morphogenesis of the virus

The involvement of proteolytic processing in viral morphogenesis has been described for the maturation of single proteins in DNA viruses such as adenovirus (Bhatti and Weber, 1979) and poxvirus (Katz and Moss, 1970), as well as for the maturation of the polyproteins that give rise to structural proteins in RNA viruses (Hellen and Wimmer, 1992). ASFV polyprotein processing could also be linked with some step of viral morphogenesis. In this context, mature proteins from the ASFV polyprotein are produced by proteolytic cleavage through an ordered recognition of the cleavage sites. Presumably, this order can be understood in terms of conformational changes in the intermediate precursors resulting from a cleavage event or from the interaction with other proteins during viral morphogenesis. In addition, this interpretation would be also in accordance with the finding that polyprotein processing takes place at late times of the infection, probably in viral factories to give rise to the mature structural proteins.

It is interesting that polyprotein pp220 is myristoylated. A biological function proposed for acylated viral proteins, which can also be attributed to polyprotein pp220 during ASFV morphogenesis, is a scaffolding role. According to this, the myristic acid could mediate protein–protein interactions within the virus, organizing (both temporally and spatially) some of the polypeptide chains during the later stages of infection (for review see Towler and Gordon, 1988; Grand, 1989). It is noteworthy that some polyproteins that give rise to structural proteins, gag in retroviruses (Henderson *et al.*, 1983), P1 in poliovirus (Rueckert and Wimmer, 1984) and ASFV polyprotein, are all myristoylated. This could reflect a unifying concept in polyprotein processing during viral morphogenesis.

In summary, the ordered cleavage of the polyprotein as well as the myristic moiety might have a chaperone-like activity in viral assembly, triggering unidirectional changes to ensure that only correctly folded protein precursors enter the assembly pathway. Taking into account the likely involvement of proteolytic processing in ASFV

morphogenesis, polyprotein pp220 could be an excellent starting point for understanding the last steps of viral replication, an unexplored topic in ASFV biology.

Selective recognition of the potential cleavage sites in the polyprotein

The finding that only some of the 19 putative cleavage sites (Gly-Gly-X) are recognized suggests that a conformational context, besides the conserved dipeptide, modulates the cleavage reaction catalysed by the proteinase. Interestingly, the NH_2-termini created in p150, p37 and p34 consist of Gly and Ala, which are among the protein-stabilizing amino acids (Varshavsky, 1992), in agreement with the predictions for viral polyprotein processing (Wellink and van Kammen, 1988). In contrast, the NH_2-terminus of p14 consists of an Asp, considered as destabilizing. Considering that this protein is the last one produced in the proteolytic cascade, it is possible that its interaction with other proteins in the virion could prevent its degradation.

With respect to the proteinase responsible for the ASFV polyprotein processing, it seems likely that only one proteolytic activity will be involved since only one type of cleavage site is found (Gly-Gly-X). Considering that proteolytic processing in either viral morphogenesis or polyprotein maturation is carried out by viral proteinases (Kräusslich and Wimmer, 1988), it could be predicted that the proteinase involved in the processing of ASFV polyprotein will be virus-encoded. An interesting possibility is that the proteinase responsible for the proteolytic processing in ASFV might be present in the polyprotein, as it occurs in many RNA virus polyproteins. This possibility remains to be investigated since a computer search of proteinase motifs in ASFV polyprotein, as well as a comparison with the known proteinases encoded by DNA viruses, adenovirus (Akusjarvi *et al.*, 1981) and HSV (Liu and Roizman, 1991), has not revealed any significant similarity.

On the other hand, it is particularly interesting that the sequence Gly-Gly-X is also recognized as a cleavage site in the maturation of adenovirus structural proteins, including the proteinase, and in some cellular proteins, including polyubiquitin (López-Otín *et al.*, 1989). Recently, a strikingly similar cleavage site (Ala-Gly-X) has also been described in the maturation of poxvirus (vaccinia virus) structural proteins (Vanslyke *et al.*, 1991). In addition, alphavirus (an RNA-containing virus) also exhibits similar amino acid sequences at the cleavage sites (Gly-Gly-X, Gly-Ala-X) that are used for the processing of its polyprotein (Shirako and Strauss, 1990). The fact that such different biological systems use a similar cleavage site could mean that the proteinases involved in the recognition have evolved from an ancestral common proteolytic activity.

Evolutionary implications of a polyprotein in a DNA-containing virus

It is well established that precursor processing by virus-encoded proteinases is a feature limited to positive-strand RNA viruses and DNA viruses, since proteolysis in negative-strand RNA viruses is usually confined to viral glycoprotein precursors catalysed by host proteinases (Kräusslich and Wimmer, 1988). Furthermore, while proteolytic processing in DNA viruses has only been described during morphogenesis (processing of single proteins), in positive-strand RNA viruses and retroviruses it is used, in addition, for gene expression (polyprotein processing) (Kräusslich and Wimmer, 1988; Hellen and Wimmer, 1992). In this context, it is tempting to speculate that polyprotein synthesis might have been an ancestral strategy for gene expression in both RNA and DNA viruses. However, while positive-strand RNA viruses have retained both competences (gene expression and virus maturation) imposed by having positive RNA as genetic material, in DNA viruses this mechanism of gene expression has evolved to be controlled at a different stage, and so only the ability of processing of single proteins to control virus assembly would remain. Therefore ASFV could be a DNA-containing virus in which evolutionary pressure would have caused polyprotein processing to be retained as a mechanism of gene expression.

With respect to the origin of a polyprotein in ASFV, one possibility is that this mechanism of regulating gene expression arose independently in RNA viruses and in DNA viruses. Another, more speculative, possibility is that the appearance of a polyprotein in ASFV is the result of a horizontal genetic transfer from an RNA virus. Some member of the Alphavirus family might be considered as a candidate since some of them infect ticks, like ASFV, and some cleavage sites recognized for proteolytic processing are similar to those cleaved in ASFV polyprotein. Transfer of the polyprotein gene from one virus to the other could perhaps have occurred in a dually infected tick cell. In this connection, genetic transfer in ticks has been already suggested to explain the similarity between telomeres of *Borrelia* (a spirocheta) and the hairpin loop termini of the ASFV genome (Hinnebusch and Barbour, 1991). Likewise, gene transfer in arthropods between a DNA virus (baculovirus) and an RNA virus (Thogoto virus) has been recently proposed (Morse *et al.*, 1992).

How and why a DNA virus has a polyprotein for gene expression remain open questions, but in any case, ASFV should be considered as an atypical virus whose study could shed light on viral biology.

Materials and methods

Cells and viruses

Vero cells (CCL 81), obtained from the American Type Culture Collection, were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% newborn calf serum (NCS), the concentration of which was reduced to 2% during viral infection. The ASFV strain BA71V, adapted to grow in Vero cells, has been previously described (Enjuanes *et al.*, 1976). Highly purified ASFV was obtained by Percoll equilibrium centrifugation as described by Carrascosa *et al.* (1985).

DNA sequencing and computer analysis

Sequencing was performed on both DNA strands using Sequenase from United States Biochem. Corp. and the standard dideoxy procedure (Sanger *et al.*, 1977). Templates were generated by subcloning the ASFV *Eco*RI C′ DNA fragment (Ley *et al.*, 1984) into M13mp18 and M13mp19 (Messing, 1983). Sequencing primers were the universal primer and synthetic oligodeoxynucleotides corresponding to specific internal sequences. Computer analysis of DNA and protein sequences was performed using the software package of the University of Wisconsin Genetics Computer Group (Devereux *et al.*, 1984).

Antisera

Recombinant plasmids were constructed by cloning the suitable fragments produced by PCR amplification (Saiki *et al.*, 1988) using the following oligonucleotides. p34 region: 5′-GCGCGCCATATGGGTAACCGTGG-ATCTTCAACC-3′ and 5′-GCGCGCGGATCCCTACTTTCTTGGAGA-GCC-3′; p37 region: 5′-GCGCCCATATGGAATTTTTGAACAAAAG-C-3′ and 5′-GCGCGGATCCACAGTAGAGAGGTAGACCC-3′; p150 region: 5′-GCGCGCATATGTCACGCATTTTTAGAGG-3′ and 5′-CG-CGCGGATCCTTATAAAATTCGAATATCGC-3′. These oligonucleotides

contained the *Nde*I and *Bam*HI restriction sites adequate for cloning into the pAR3038 vector (Rosenberg *et al.*, 1987). The p34 region spans amino acid residues 1–304, the p37 region includes amino acid residues 391–838 and the p150 region includes amino acid residues 2095–2475 of protein C'2475L.

The *Escherichia coli* strain BL21DE3 (Studier *et al.*, 1990), which contains in its genome the T7 RNA polymerase gene under an isopropyl-β-D-thiogalactopyranoside (IPTG) inducible promoter, was transformed with the recombinant plasmids. The proteins of 33 kDa (p34 region), 50 kDa (p37 region) and 43 kDa (p150 region) used as immunogens were obtained by inducing the cultures for 2 h with 1 mM IPTG. The cultures were further incubated for 2 h with rifampicin (0.2 mg/ml). The bacterial pellet (from 100 ml culture) was lysed in 5 ml of phosphate buffered saline (PBS) by freezing and thawing twice at −20°C and sonication twice for 2 min. Bacterial lysates were subjected to centrifugation (10 min at 12 000 *g*). In all cases the recombinant proteins were the main component present in the pellet. About 100 μg of each protein was injected subcutaneously into rabbits in complete Freund's adjuvant. At intervals of 2 weeks, rabbits were boosted three times, subcutaneously, once in incomplete Freund's adjuvant and twice in PBS. After the fourth injection the animals were bled and the sera were tested by immunoprecipitation.

Radiolabelling and immunoprecipitation

Preconfluent monolayers of Vero cells, cultured in 24-well plates, were infected with ASFV at 20 p.f.u./cell. At 12 h.p.i., cells were pulse-labelled for 30 min with 300 μCi/ml [^{35}S]methionine/cysteine (ICN Biomedicals) in methionine/cysteine-free DMEM. Before and after the pulse, medium was replaced with methionine/cysteine-free DMEM for 30 min to remove any residual methionine and cysteine. DMEM containing 2 mM methionine and cysteine was added and incubated for different chase periods. After the medium had been removed, cells were washed with PBS and lysed in dissociation buffer [2% SDS, 1% Nonidet P40 (NP-40), 1% sodium deoxycholate (DOC), 100 mM DTT , 1 mM PMSF, 0.1 mM TLCK, 0.1 mM TPCK in PBS]. The extracellular virus present in the medium was recovered by centrifugation in a Beckman Airfuge at 133 000 *g* for 20 min. Metabolic labelling with [^{3}H]myristic acid was performed as previously described (Aguado *et al.*, 1991). Metabolic labelling with [^{35}S]methionine/cysteine of highly purified ASFV particles was performed as described by Carrascosa *et al.* (1985).

For immunoprecipitation assays, cell extracts in dissociation buffer were diluted at least 10-fold in RIPA buffer (0.01 M Tris–HCl, pH 7.5, 0.15 M NaCl, 1% DOC, 1% NP-40, 0.1% SDS) and incubated with a 1:20 dilution of the specific antisera for 2 h at 4°C. Then a 25% suspension of protein A-coated Sepharose CL-4B (Sigma Chemical Co., St Louis, MO) was added and incubation continued for 30 min at 4°C. The immune complexes were isolated by centrifugation, washed four times with cold RIPA buffer and finally solubilized by boiling in Laemmli sample buffer (0.005 M Tris–HCl, pH 6.8, 2% SDS, 100 mM DTT, 10% glycerol) for 3 min. For immunoprecipitation with antisera obtained against p34 region, lysates were boiled for 3 min prior to the addition of the antisera, to avoid non-specific precipitation.

Gel electrophoresis

High resolution two-dimensional gel electrophoresis was performed essentially as described by Bravo (1984). Isoelectric focusing (IEF) gels were prepared using a combination of ampholytes in the pH ranges 5–7 (Serva, Heidelberg, Germany) and 3.5–10 (Pharmacia LKB, Uppsala, Sweden) in a 4:1 ratio. Before loading, gels were prerun at 700 V for 90 min. Then, ^{35}S-radiolabelled purified ASFV, solubilized in lysis buffer (9.8 M urea, 2% NP-40, 2% ampholytes pH 7–9, 0.1 M DTT), was applied and electrophoresed at 800 V for 18 h and 1000 V for 30 min. Non-equilibrium pH gradient electrophoresis (NEPHGE) was performed using a combination of ampholytes pH 7–9 and 8–9.5 (Pharmacia LKB) in a 1:1 ratio and running at 400 V for 4.5 h. SDS–polyacrylamide gradient (7–20%) gels (Laemmli, 1970) were performed for second dimension separation and for conventional SDS–PAGE. Radioactive proteins were detected by fluorography (Bonner and Laskey, 1974).

Indirect immunofluorescence

Vero cells were infected at a multiplicity of infection of 1 p.f.u./cell. At 12 h.p.i. cells were rinsed with PBS and fixed with methanol at −20°C for 5 min. For staining, fixed cells were incubated with a 1:100 dilution of the different antisera and then with a 1:200 dilution of rhodamine-conjugated goat anti-rabbit immunoglobulin G (TAGO Inc. Burlingame, CA). Cells were examined under a Zeiss Axiovert microscope.

Acknowledgements

We thank J.A.García and J.Ortín for helpful comments on the manuscript. This work was supported by grants from Comisión Interministerial de Ciencia y Tecnología, Junta de Extremadura and by an institutional grant from Fundación Ramón Areces. G.A. and C.S.-M. were recipients of fellowships from Ministerio de Educación y Ciencia.

References

Aguado,B., Viñuela,E. and Alcamí,A. (1991) *Virology*, **185**, 942–945.
Akusjarvi,G., Zabielski,J., Perricaudet,M. and Pettersson,U. (1981) *Nucleic Acids Res.*, **9**, 1–17.
Alcamí,A., Angulo,A., López-Otín,C., Muñoz,M., Freije,J.M.P., Carrascosa, A.L. and Viñuela,E. (1992) *J. Virol.*, **66**, 3860–3868.
Almazán,F., Rodríguez,J.M., Andrés,G., Pérez,R., Viñuela,E. and Rodriguez,J.F. (1992) *J. Virol.*, **66**, 6649–6654.
Andrés,G., Simón-Mateo,C. and Viñuela,E. (1993) *Virology*, **194**, 284–293.
Bhatti,A.R. and Weber,J. (1979) *J. Biol. Chem.*, **254**, 12265–12268.
Bonner,W.M. and Laskey,R.A. (1974) *Eur. J. Biochem.*, **46**, 83–88.
Bravo,R. (1984) In J.E.Celis and R.Bravo (eds), *Two Dimensional Gel Electrophoresis of Proteins. Methods and Applications.* Academic Press, New York, pp. 3–36.
Carrascosa,A.L., del Val,M., Santarén,J.F. and Viñuela,E. (1985) *J. Virol.*, **54**, 337–344.
Costa,J.V. (1990) In Darai,G. (ed.), *Molecular Biology of Iridoviruses.* Kluwer Academic Publishers, Boston, MA, pp. 247–270.
Del Val,M. and Viñuela,E. (1987) *Virus Res.*, **7**, 297–308.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
Enjuanes,L., Carrascosa,A.L., Moreno,M.A. and Viñuela,E. (1976) *J. Gen. Virol.*, **32**, 471–477.
Esteves,A., Marques,M.I. and Costa,J.V. (1986) *Virology*, **152**, 181–191.
González,A., Talavera,A., Almendral,J.M. and Viñuela,E. (1986) *Nucleic Acids Res.*, **14**, 6835–6844.
Grand,R.J.A. (1989) *Biochem. J.*, **258**, 625–638.
Hames,B.D. (1990) In Hames,B.D. and Rickwood,D. (eds), *Gel Electrophoresis of Proteins—A Practical Approach.* IRL Press, Oxford, pp. 16–21.
Hellen,C.U.T. and Wimmer,E. (1992) *Experientia*, **48**, 201–215.
Henderson,L.E., Krutzsch,H.C. and Oroszlan,S. (1983) *Proc. Natl Acad. Sci. USA*, **80**, 339–343.
Hinnebusch,J. and Barbour,A.G. (1991) *J. Bacteriol.*, **173**, 7233–7239.
Katz,E. and Moss,B. (1970) *Proc. Natl Acad. Sci. USA*, **66**, 677–684.
Kozak,M. (1983) *Microbiol. Rev.*, **47**, 1–45.
Kräusslich,H.G. and Wimmer,E. (1988) *Annu. Rev. Biochem.*, **57**, 701–754.
Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
Ley,V., Almendral,J.M., Carbonero,P., Beloso,A., Viñuela,E. and Talavera,A. (1984) *Virology*, **133**, 249–257.
Liu,F. and Roizman,B., (1991) *J. Virol.*, **65**, 5149–5156.
López-Otín,C., Simón,C., Méndez,E. and Viñuela,E. (1988) *Virus Genes*, **1**, 291–303.
López-Otín,C., Simón-Mateo,C., Martinez,L. and Viñuela,E. (1989) *J. Biol. Chem.*, **264**, 9107–9110.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
Morse,M.A., Marriott,A.C. and Nuttall,P.A. (1992) *Virology*, **186**, 640–646.
Moura Nunes,J.F., Vigario,J.D. and Terrinha,A.M. (1975) *Arch. Virol.*, **49**, 59–66.
Rosenberg,A.H., Lade,B.N., Chui,D.-S., Lin,S.-W., Dunn,J.J. and Studier,F.W. (1987) *Gene*, **56**, 125–135.
Rueckert,R.R. and Wimmer,E. (1984) *J. Virol.*, **50**, 957–959.
Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Ehrlich,H.A. (1988) *Science*, **239**, 487–491.
Salas,M.L., Kuznar,J. and Viñuela,E. (1981) *Virology*, **101**, 169–175.
Salas,M.L., Rey-Campos,J., Almendral,J.M., Talavera,A. and Viñuela,E. (1986) *Virology*, **152**, 228–240.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Santarén,J.F. and Viñuela,E. (1986) *Virus Res.*, **5**, 391–405.
Sanz,A., García-Barreno,B., Nogal,M.L., Viñuela,E. and Enjuanes,L. (1985) *J. Virol.*, **54**, 199–206.

Shirako,Y. and Strauss,J.H. (1990) *Virology*, **177**, 54–64.
Sogo,J.M., Almendral J.M., Talavera,A. and Viñuela,E. (1984) *Virology*, **133**, 271–275.
Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
Towler,D.A. and Gordon,J.I. (1988) *Annu. Rev. Biochem.*, **57**, 69–99.
Vanslyke,J.K., Whitehead,S.S., Wilson,E.M. and Hruby,D.E. (1991) *Virology*, **183**, 467–478.
Varshavsky,A. (1992) *Cell*, **69**, 725–735.
Viñuela,E. (1987) In Becker,Y. (ed.), *African Swine Fever*. Martinus Nijhoff Publishing, Boston, MA, pp. 31–49.
Wellink,J. and van Kammen,A. (1988) *Arch. Virol.*, **98**, 1–26.

Received on January 1, 1993; revised on March 3, 1993

Note added in proof

The nucleotide sequence data reported here have been deposited in the EMBL/DDBJ/GenBank databases under the accession number Z22777 'POLY'.