208–213 Nucleic Acids Research, 1994, Vol. 22, No. 2

Genetic identification and nucleotide sequence of the DNA polymerase gene of African swine fever virus

Alexandra Martins, Graça Ribeiro, M.Isabel Marques and João V.Costa*
Laboratory of Virology II, Gulbenkian Institute of Science, Apartado 14, P-2781 Oeiras Codex, Portugal

Received September 23, 1993; Revised and Accepted December 17, 1993 EMBL accession no. X73330

ABSTRACT

The DNA polymerase gene of African swine fever virus (ASFV) was mapped by marker rescue experiments using a phosphonoacetic acid-resistant mutant and hybridization with an oligonucleotide probe designed from the most conserved motif of family B DNA polymerases. Viral DNA fragments mapping in this region were cloned and sequenced. An open reading frame coding for a 1244 aminoacid long peptide with a molecular mass of 142.5 kDa was determined from the sequence. A unique feature of ASFV DNA polymerase is the presence of 13 tandem repeats of the sequence Ala-Gly-Asp-Pro near the carboxyl end of the molecule. Comparison with 30 sequences of alpha-like DNA polymerases of cellular and viral origin showed that ASFV DNA polymerase has all the conserved motifs of family B DNA polymerases. A 3.9 kb transcript was detected by Northern hybridization and the transcription initiation and termination sites were mapped by S1 analysis and primer extension. Late transcription was initiated at a site different from the early transcription initiation site. A 145 kDa protein, consistent with the size of the gene, was identified by an *in situ* enzyme assay after gel electrophoresis of infected cell extracts.

INTRODUCTION

Comparison of DNA polymerase aminoacid sequences from many organisms has enabled the classification of DNA polymerases into three families (1). The proposed family A includes the bacterial and bacteriophage enzymes that are homologous to *E.coli* DNA polymerase I. Family B comprises *E.coli* DNA polymerase II (2), eukaryotic alpha and delta DNA polymerases (3–12), DNA polymerases encoded by eukaryotic linear plasmids (13–17), several bacteriophage DNA polymerases (18–21), and all known DNA polymerases encoded by animal viruses: herpesviruses (22–25), poxviruses (26, 27), baculoviruses (28, 29), and adenoviruses (30). Family C polymerases are the bacterial replicative DNA polymerases, like *E.coli* DNA polymerase III α and ϵ subunits (1).

African swine fever virus (ASFV) is the causative agent of an economically important disease of swine. The virus is a large icosahedral deoxyvirus, formerly classified as an iridovirus. ASFV is now classified in an unnamed family of which it is the only member (31). However, ASFV is related to the iridoviruses in what concerns morphology and to the poxviruses in terms of its DNA structure and molecular biology (for a review see 32). ASFV replicates in the cytoplasm of infected cells and must therefore carry or code for the enzymes required for transcription and DNA replication, namely RNA polymerase (33), mRNA processing enzymes (34) and DNA polymerase (35). ASFV-induced DNA polymerase is similar to the alpha-like DNA polymerases in its sensitivity to cytosine arabinoside and phosphonoacetic acid (PAA) but, in contrast to all other known DNA polymerases of viral origin, ASFV-specific DNA polymerase is highly resistant to inhibition by aphidicolin (36). In this paper we report the entire nucleotide sequence of the ASFV DNA polymerase gene and we show that the enzyme has considerable homology to alpha-like DNA polymerases. The size of the deduced product is compatible with the size of a transcript detected by Northern analysis and with the dimension of a virus-induced protein detected by *in situ* enzymatic assays.

MATERIALS AND METHODS

DNA preparation

Monkey Vero cells and Vero cell-adapted ASFV were grown as described (37). Extracellular virus was concentrated by centrifugation at 6000 *g* for 16 h at 4°C and treated for 30 min at 37°C with 25 μg/ml DNase and 50 μg/ml RNase. EDTA and NaCl were added to give final concentrations of 25 mM and 1 M, respectively, and the virus was centrifuged at 85000 *g* for 30 min at 4°C through a cushion of 20% sucrose in 50 mM Tris–HCl, pH 7.5, 1 mM EDTA, and 1 M NaCl. The pellet was resuspended in TE (10 mM Tris-HCl, pH 8, 1 mM EDTA) and viral DNA was released by incubation for 1 h at 60°C with 500 μg/ml proteinase K and 0.5% SDS. DNA was then extracted twice with phenol and the aqueous phase was treated again with proteinase K and SDS for 1 h at 37°C. After extraction with phenol-chloroform-isoamyl alcohol and with chloroform-isoamyl alcohol, DNA was precipitated with ethanol and dissolved in TE.

*To whom correspondence should be addressed

DNA analysis and cloning

In vitro manipulation of DNA, agarose gel electrophoresis, Southern hybridization, and cloning in plasmid pBluescript II SK$^+$ (Stratagene) were performed following standard procedures (38).

Clones with nested deletions starting at either end of the initial insert were prepared by digestion with exonuclease III and S1 nuclease according to the method of Henikoff (39). Protected 3′ ends were generated by cleavage with *Sac*I or *Apa*I and protruding 5′ ends sensitive to exonuclease III digestion were generated by digestion with *Bam*HI and *Eco*RI, respectively.

Polymerase chain reaction (PCR)

DNA extracted from purified ASFV was amplified by PCR in 100 μl reactions containing 100 ng DNA, 50 pmol of each primer, 0.2 mM each of dATP, dCTP, dGTP and dTTP, 5 U Taq DNA polymerase (AmpliTaq, Perkin Elmer Cetus), in the buffer provided by the enzyme supplier (PCR buffer: 10 mM Tris−HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 0.001% gelatin). The samples were cycled 40 times with a temperature profile of 40 s at 94°C, 1 min at 55°C, and 1 min at 72°C. The individual strands of viral DNA were also amplified by asymmetrical PCR in reactions containing 1 pmol of one primer and 50 pmol of the other primer.

DNA sequencing

For sequencing, 1 pmol of double-stranded plasmid DNA was denatured in 0.2 N NaOH and 2 mM EDTA, neutralized and precipitated with ethanol. The denatured DNA was used as template for sequencing by the dideoxy chain termination method (40), using a Sequenase (version 2.0) kit from United States Biochemicals as recommended by the manufacturer.

Transcription analysis

Infected cells were harvested at 6 h or 12 h after infection at a m.o.i. of 30 p.f.u. per cell. RNA was extracted by the guanidinium thiocyanate method, subjected to agarose gel electrophoresis in the presence of formaldehyde, and hybridized with cloned viral DNA according to standard methods (38). The identification of transcription initiation and termination sites was performed by primer extension and by S1 nuclease protection (38).

In situ assay of DNA polymerase

Cytoplasmic extracts of mock-infected and infected cells were prepared at 6 h and 12 h after infection at a m.o.i. of 30 p.f.u. per cell and diluted in SDS-polyacrylamide gel electrophoresis sample buffer. Proteins in the extracts were partially denatured by incubation for 3 min at 37°C and separated in a 7.5% acrylamide gel containing 150 μg/ml activated calf thymus DNA. All subsequent steps of the assay were done essentially as described (41). The gel was washed three times for 20 min with 50 mM Tris−HCl pH 7.5 at 4°C and proteins were renatured for 3 h at 4°C in 500 ml reaction mixture (50 mM Tris−HCl pH 7.5, 7.5 mM magnesium acetate, 40 mM KCl, 1 mM dithiothreitol, 400 μg/ml bovine serum albumin, and 16% glycerol). They were then incubated for 12−18 h at 37°C in 40 ml of reaction mixture containing 5 μCi/ml α-[^{32}P]dTTP (3000 Ci/mmol, Amersham), 1 μM unlabelled dTTP, and 12.5 μM of the other deoxyribonucleoside triphosphates. To test the effect of phosphonoacetic acid, the inhibitor was added to the reaction mixture at a concentration of 100 μg/ml. After extensive washing with 5% TCA−1% sodium pyrophosphate, the gels were dried and radioautographed.

Marker rescue

Vero cells were infected at a m.o.i. of 0.5 p.f.u. per cell. Two hours after infection, cells were trypsinized, resuspended in 20 mM Hepes−HCl pH 7.05, 140 mM NaCl, 50 mM KCl, 0.7 mM Na_2HPO_4, 5 mM glucose, at a concentration of 5×10^6 cells/ml and electroporated with 25 μg DNA from mutant PAA100 or from wild type virus. Cells were plated in fresh medium and the virus was collected 48 h later, when total cytopathic effect was observed. Titration was performed in the presence of 300 μg/ml PAA.

Computer analysis

Routine analysis of DNA and protein sequences were performed with the programs of the University of Wisconsin Genetics Computer Group (42). The program PileUp was used for multiple alignments of homologous sequences and homology scores were determined with the program Distances.

RESULTS AND DISCUSSION

Mapping and cloning of the DNA polymerase gene

The aminoacid sequences of all known viral DNA polymerases as well as cellular alpha-like DNA polymerases show several conserved regions. The most highly conserved sequence, Tyr-Gly-Asp-Thr-Asp-Ser, was designated as motif I in the carboxyl-terminal polymerase domain (3) and is the proposed catalytic site (43, 44). This motif I is also conserved in ASFV DNA polymerase, as results will show.

To locate ASFV DNA polymerase gene, we deduced from that aminoacid sequence and synthesized the degenerate oligonucleotide probe (G/C)(A/T)(A/G)TCNGT(A/G)TCNCC(A/G)TA. The probe was 5′ end-labelled and hybridized to viral DNA and to DNA from a set of lambda EMBL3 recombinants covering the whole viral genome. As shown in Figure 1, the probe hybridized only to the 7.1 kb *Eco*RI fragment H of viral DNA and to the lambda recombinant lGR131, which maps in that region of viral DNA. Fragment *Eco*RI-H is located at position 90.3 to 97.4 (in kb) on the physical map of ASFV DNA (Figure 2, A). Further mapping was achieved by observing that the probe hybridized only to a 1.9 kb NcoI fragment that maps at the right end of fragment *Eco*RI-H.

To confirm the identification of the gene, we performed marker rescue experiments using DNA from a PAA-resistant mutant. Mutant PAA100 was selected after serial passages of ASFV in the presence of 100 μg/ml PAA and produces viral titers higher than 10^6 p.f.u./ml in the presence of 300 μg/ml PAA, whereas wild type virus titers are lower than 10 p.f.u./ml. Mutant viral DNA was cloned and fragments spanning different parts of the genome were transfected into cells infected with wild type virus in the presence of PAA. Production of PAA-resistant virus was observed only when infected cells were cotransfected with PAA100 DNA fragment *Eco*RI-H. The resistant virus had a titer of 3.7×104 p.f.u./ml. No viral plaque was observed when infected cells were transfected with control plasmid DNA or with all the other mutant DNA fragments. Further mapping was achieved by marker rescue using the 3.8 kb *Eco*RI/*Bam*HI fragment mapping at the right end of fragment *Eco*RI-H (Figure

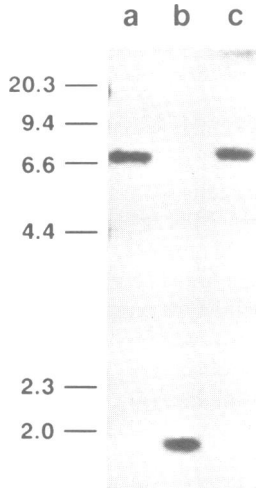

Figure 1. Mapping of the ASFV DNA polymerase gene. Viral DNA cleaved with *Eco*RI (**a**) or with *Nco*I (**b**) and DNA from recombinant lGR131 cleaved with *Eco*RI (**c**) were hybridized with an oligonucleotide probe designed from the most highly conserved motif of alpha-like DNA polymerases (YGDTDS). Dimension markers are fragments of lambda DNA digested with *Hin*dIII.

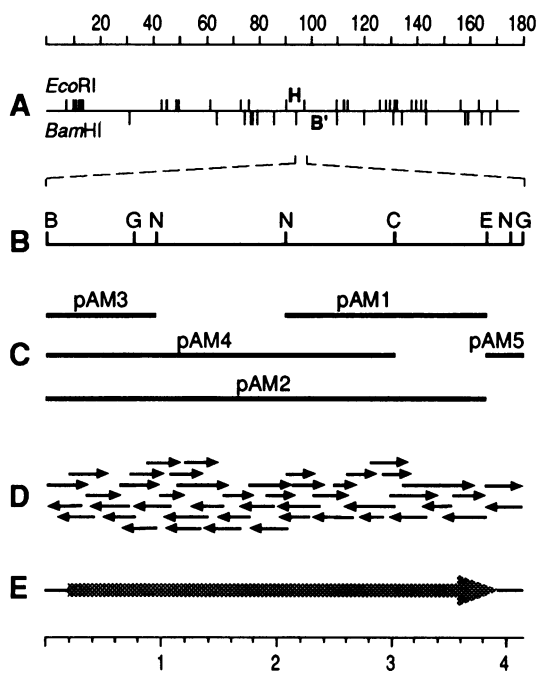

Figure 2. (**A**) Physical map of *Eco*RI and *Bam*HI fragments of ASFV DNA. (**B**) Expanded map of the region containing the DNA polymerase gene. (**C**) Location of the clones used for sequencing the gene. Abbreviations: B, *Bam*HI; C, *Cla*I; E, *Eco*RI; G, *Bgl*II; N, *Nco*I. (**D**) DNA sequencing strategy. The arrows indicate the direction and extent of the determined sequences. (**E**) Location and orientation of the gene. All the dimensions are indicated in kilobase pairs.

2, A). Infected cells transfected with this fragment produced PAA-resistant virus with a titer of 8.6×104 p.f.u./ml.

This *Eco*RI/*Bam*HI fragment was cloned in pBluescript II SK$^+$ and the clone, pAM2, was used to prepare two symmetrical sets of nested deletions by digestion with exonuclease III (39). Three other subclones were also constructed and used to confirm the sequences: pAM1, containing the region that extends from the second *Nco*I site to the *Eco*RI site; pAM3, containing the left part of pAM2, between the *Bam*HI site and the first *Nco*I site; and pAM4, containing the region between the *Bam*HI site and the *Cla*I site (Figure 2, C). A *Bgl*II site was mapped at 0.32 kb to the right of the *Eco*RI site. This small *Eco*RI/*Bgl*II fragment was also cloned (pAM5) and sequenced.

Nucleotide sequence of the DNA polymerase gene

The clones described above and the deletion clones prepared from pAM2 were sequenced by the dideoxy chain termination method using standard primers corresponding to both flanking regions of the vector DNA or specific primers synthesized on the basis of sequences already read. The strategy for sequencing both DNA strands is depicted schematically in Figure 2 (D).

The 4125 nucleotide-long complete sequence of pAM2 and of the adjacent clone pAM5 was entered in EMBL Data Library with accession number X73330. The sequence showed a single open reading frame (ORF) of 3732 nucleotides, starting at nucleotide 205 with an ATG codon that has similarity with Kozak's consensus (45). This long ORF can code for a basic polypeptide of 1244 aminoacids with a calculated molecular mass of 142.5 kDa and a predicted pI of 8.6. This predicted molecular mass makes ASFV DNA polymerase the largest viral DNA polymerase described so far.

A remarkable feature of this ORF is the presence near the 3′ end of a set of 13 tandem repeats of the 12 nucleotide-long sequence GCCGGCAATCCG, coding for alanine-glycine-asparagine-proline. This structure is not found in any other DNA polymerase. To confirm the presence of these repeats, we sequenced the corresponding region of two other independent clones of fragment *Eco*RI-H and two independent clones of fragment *Bam*HI-B′, which overlaps fragment *Eco*RI-H. All the clones showed the same set of repeated sequences. By using a forward primer that hybridizes with sequence 3152−3170 and a reverse primer that hybridizes with sequence 3716−3735, we amplified viral DNA by asymmetrical PCR and sequenced both viral DNA strands. This experiment confirmed that the repeats were indeed present in viral DNA and were not a cloning artifact. The sequence read from the PCR products did not show any additional termination codon that could interrupt translation before the repetitions. A sequencing error that might have made us miss a termination codon is very unlikely because the protein synthesized *in vitro* from a pAM1 T3 RNA polymerase transcript had the dimension expected if translation occurred through the repeat region (data not shown).

C-terminal repeats of a sequence of seven aminoacids were described in the large subunit of RNA polymerase II (46, 47). In contrast to the RNA polymerase II tail piece, the repeated sequence of ASFV DNA polymerase does not have phosphorylation sites and is not strongly hydrophilic. The role of the RNA polymerase tail piece is not known but it has been suggested that it might be involved in protein-protein interaction, in protein processing, or in competition with DNA for histone binding (46). Cellular DNA polymerases bind to DNA through zinc-finger motifs at the carboxyl end (3) and the carboxyl terminal region of herpesvirus simplex 1 DNA polymerase is essential for DNA binding and for interaction with protein UL42, an accessory protein that increases the enzyme's processivity (48). It is possible that the repeats in ASFV DNA polymerase are also involved in enzyme interactions.

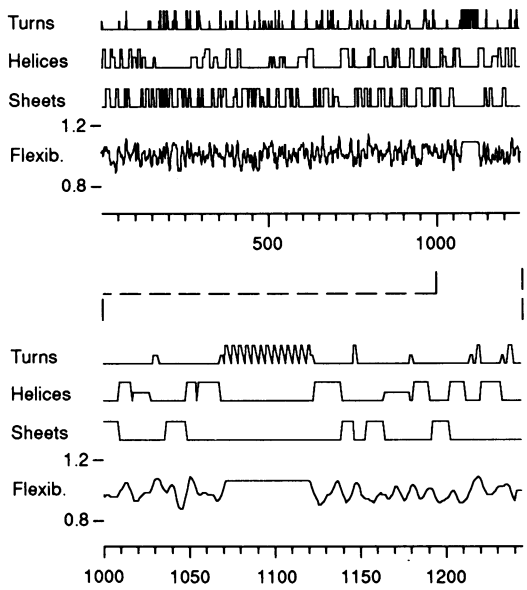

Figure 3. Prediction of the secondary structure of ASFV DNA polymerase. The complete sequence (upper panel) or the carboxyl-terminal 244 aminoacids (lower panel) were analyzed with the program PeptideStructure of the Wisconsin package, using the Chou-Fasman method to predict alpha helices, beta sheets, and turns. The lowest curve shows the flexibility index.

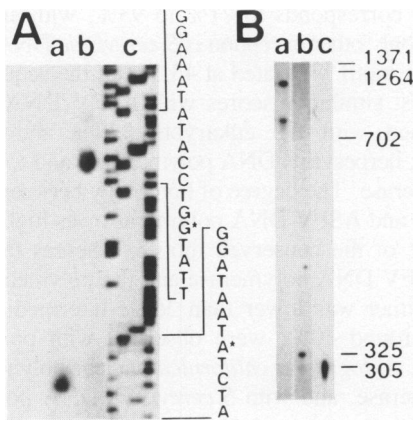

Figure 4. Transcription analysis of the DNA polymerase gene. **A.** primer extension of a synthetic primer complementary of sequence 226–255 on early (**a**) and late (**b**) RNA from infected cells. A sequencing reaction using the same primer is shown in panel (**c**). The written sequences represent the complement of the sequences actually shown in the gel. **B.** Nuclease S1 mapping of the 5′ end (**a**) and the 3′ end (**c**) of the mRNA for the DNA polymerase. pAM3 DNA labelled at the 5′ end and pAM5 DNA labelled at the 3′ end were used for mapping the 5′ end and the 3′ end, respectively. Lane (**b**) shows the intact pAM5 fragment. Dimension markers between 0.7 and 1.3 kb are *Bst*EII fragments of lambda DNA and the smaller dimensions were determined from a sequencing ladder run in the same gel.

Figure 3 depicts the predicted secondary structure of the ASFV DNA polymerase. The expanded region, containing the repeated sequences, shows a very flexible structure with high probability of turns, flanked by two probable alpha helices. The high degree of local flexibility may contribute to DNA affinity or to interactions with accessory proteins.

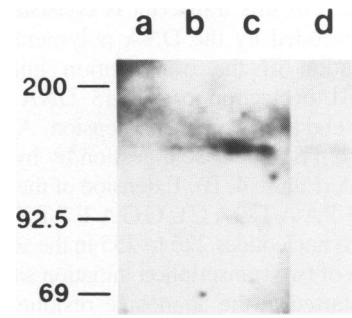

Figure 5. *In situ* assay of DNA polymerase after acrylamide gel electrophoresis of uninfected (**a**) and infected cell extracts, prepared at 6 (**b**) and 12 h (**c**) after infection. A parallel assay of late extracts was performed in the presence of 100 μg/ml phosphonoacetic acid (**d**).

```
Exo I
206        YDIET
6-702      FDIET
Exo II/Pol IV
261 (51)   TIVCSSEKNLLLSFAEQFSRWAPDICTGFNDSRYDWPFIVEK
(25-102)   V----sE--LL--F--------pdvitGfN---FDLpYi--R
Exo III
413 (109)  VAYYCIIDAQR
(58-177)   ---Y---D---
Pol II
491 (64)   YPGA..FVIDPVKGLEQDKPTTGLDFASLYPSLIMAYNFSP
(24-147)   --Ga---Vf-P--g-y----i--lDFnSLYPSI-------y
Pol VI
578 (47)   YPKVLIDLLNKRTALKQ
(19-112)   Lp--L--L---R---Kk
Pol III
624 (30)   DAKQKALKIFMN.TFYGEAGNNLSPFFL.LPLAGGVTSSG
(4-66)     d--Q-alKi-aN-S-YG-FG------------a--VT--G
Pol I
679 (13)   IKYGDTDSLYI
(13-136)   viYGDTDSv-v
Pol V
764 (76)   KKKYYGIAHVNTPNFNTKELF.IRGIDIIK
(30-85)    KKkY-g----------------KGld--R
```

Figure 6. Comparison of ASFV DNA polymerase sequence with the consensus sequences of the exonuclease and polymerase motifs of family B DNA polymerases. A multiple alignment of the 30 DNA polymerase sequences listed in the text was used to extract the consensus sequences. The consensus motifs of all the DNA polymerases start with exoI at the amino-terminal region and continue in the same order as indicated in the figure. Upper lines show the ASFV DNA polymerase sequences and lower lines show the consensus sequences. Letters in bold type indicate aminoacids in ASFV DNA polymerase that match the consensus. Aminoacids in uppercase in the consensus sequences are present in more than two thirds of the proteins and aminoacids in lowercase are present in at least one third. Numbers in upper lines indicate the position of the motif in the ASFV DNA polymerase sequence and the distance from the preceding motif (within parenthesis). Numbers in the consensus lines indicate the shortest and the longest distance from the preceding motif.

Transcription analysis of the DNA polymerase gene

ASFV genes are transcribed either early or late after infection, the two phases being separated by the onset of viral DNA replication. Some genes are expressed during both phases f infection (37). Transcription of the DNA polymerase gene was analyzed by Northern hybridization. A large 3.9 kb RNA was detected by hybridization of early and late RNA with pAM1. A transcript of the same size was also detected when we used as probes pAM3 and pAM5, which correspond to both ends of

the gene. The size of this transcript is consistent with the size of the peptide encoded by the DNA polymerase gene.

The identification of the transcription initiation site was performed by S1 protection of pAM3 DNA asymmetrically labelled at the 5′ end and by primer extension. A 0.8 kb fragment was protected from S1 nuclease digestion by hybridization with infected cell RNA (Figure 4, B). Extension of the synthetic primer CCGTTGGGTAATCACCGGATTTTCCCGTGC, complementary to nucleotides 226 to 255 in the sequence, enabled the identification of two transcription initiation sites: transcription of early RNA started at the adenosine residue at position 198 and transcription of late RNA was initiated at the guanosine residue at position 161 (Figure 4, A).

A consensus sequence for ASFV early promoters has not been identified up to now. Comparison of the region upstream from some early genes reveals a high A and T content, with frequent runs of consecutive A residues (49). The region immediately preceding the transcription initiation site of the DNA polymerase gene also shows a high A and T content, with stretches of four or six A residues. Multiple RNA species have been reported for other ASFV genes expressed at both early and late times after infection (50) but it was not shown whether this was due to alternative initiation or termination. Finer analysis of transcription of other genes expressed at both stages of infection will be required to see whether alternative initiation, as in the case of the DNA polymerase gene, is the general mechanism for alternative transcription in the early and late phases of infection.

Transcription termination was mapped by S1 nuclease protection of the 325 nucleotide-long *Eco*RI/*Bgl*II fragment cloned in pAM5, labelled at the 3′ end of the *Eco*RI site. A short ladder of protected bands around 305 nucleotides long was observed (Figure 4, B), indicating that transcription terminates close to a stretch of eight thymidylate residues starting at position 4099. Stretches of thymidylate residues were proposed as the termination signal for ASFV transcription (49).

Identification of DNA polymerase

Infected cell proteins extracted early and late after infection were assayed for DNA polymerase activity after electrophoresis in gels containing activated DNA (Figure 5). *In situ* assays of cytoplasmic extracts prepared at 6 h after infection showed a weak band of DNA polymerase activity corresponding to a protein of 145 kDa, a molecular mass that is compatible with the size of the product deduced from the gene sequence. A band of this size was much more intense in extracts prepared at 12 h after infection. This polymerase activity was sensitive to phosphonoacetic acid, an inhibitor of ASFV DNA polymerase. Cellular DNA polymerase alpha was detected in similar assays using total cell extracts from uninfected or infected cells but was not detected in cytoplasmic extracts.

Homology of ASFV DNA polymerase with the alpha-like (family B) DNA polymerases

The aminoacid sequence of ASFV DNA polymerase was compared with the aminoacid sequences of the following 30 DNA polymerases belonging to the B family: vaccinia virus (26), fowlpox virus (27), herpes simplex virus type 1 (23), Epstein–Barr virus (22), human cytomegalovirus (25), varicella-zoster virus (24), *Autographa californica* nuclear polyhedrosis virus (28), PBCV-1 *Chlorella* virus (51), human adenovirus type 2 (30), bacteriophage T4 (20), bacteriophage M2 (21), bacteriophage ϕ29 (18), bacteriophage PRD1 (19), *E.coli* DNA polymerase II (2), human DNA polymerase α (3), *Drosophila melanogaster* DNA polymerase α (7), *Trypanosoma brucei* DNA polymerase α (8), *S. cerevisiae* DNA polymerase I (α) (4), *Schizosaccharomyces pombe* DNA polymerase I (α) (9), human DNA polymerase δ (10), bovine DNA polymerase δ (11), *S.cerevisiae* DNA polymerase III (δ) (12), *S.cerevisiae* DNA polymerase II (ϵ) (6), *S.cerevisiae* DNA polymerase Rev3 (5), *Sulfolobus solfataricus* DNA polymerase (52), *Claviceps purpurea* plasmid (16), *Kluyveromyces lactis* plasmid 1 and 2 (13, 15), S-1 maize mitochondrial DNA (14), and *Ascobolus immersus* plasmid (17).

The comparison disclosed stretches of marked homology corresponding to the three conserved motifs in the amino-terminal 3′–5′ exonuclease domain (53) and to the five motifs in the carboxyl-terminal polymerase domain (3). We extracted from these alignments consensus sequences, which are compared to ASFV DNA polymerase sequence in Figure 6. The motifs present in the ASFV DNA polymerase sequence strongly matched the consensus of the motifs, defined as 20 or more identities in the 30 sequences. In most of the cases where a mismatch was found, the ASFV DNA polymerase sequence showed a conservative replacement.

The conserved polymerase motifs of ASFV DNA polymerase were located more centrally in the sequence than in most of the other DNA polymerases. The rightmost motif V was located at a position corresponding to 61% of the total length of the sequence, whereas the location of this motif in the other polymerases corresponds to 71% to 93%, with an average of 78.5%. The only other exception is *S.cerevisiae* DNA polymerase II, which has motif V located at 43.5% of the sequence length.

The highest similarity scores with ASFV DNA polymerase were obtained with the eukaryotic alpha and delta DNA polymerases, herpesvirus DNA polymerases, and *Chlorella* virus DNA polymerase. The degree of homology between these DNA polymerases and ASFV DNA polymerase was higher than 50% in the whole of the conserved motifs, whereas the homology between ASFV DNA polymerase and the polymerases that use a protein primer was lower than 35%. Intermediate levels of similarity, around 40%, were observed with poxvirus DNA polymerases, *Autographa californica* nuclear polyhedrosis virus DNA polymerase, and with *S.cerevisiae* DNA polymerases II and Rev3.

ACKNOWLEDGEMENTS

We are grateful to Jlia Fonseca and Fernando Marques for technical assistance and to Dr. Zilda Carvalho for critical comments on the manuscript. This work was supported in part by grant STRDA/C/BIO/412/92 from the Portuguese Junta Nacional de Investigação Cientfíca e Tecnológica.

REFERENCES

1. Ito,J. and Braithwaite,D.K. (1991) *Nucleic Acids Res.*, **19**, 4045–4057.
2. Iwasaki,H., Ishino,Y., Toh,H., Nakata,A. and Shinagawa,H. (1991) *Mol. Gen. Genet.*, **226**, 24–33.
3. Wong,S.W., Wahl,A.F., Yuan,P.-M., Arai,N., Pearson,B.E., Arai,K.-I., Korn,D., Hunkapiller,M.W. and Wang,T.S.-F. (1988) *EMBO J.*, **7**, 37–47.
4. Pizzagalli,A., Valsasnini,P., Plevani,P. and Lucchini,G. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3772–3776.
5. Morrison,A., Christensen,R.B., Alley,J., Beck,A.K., Bernstine,E.G., Lemontt,J.F. and Lawrence,C.W. (1989) *J. Bacteriol.*, **171**, 5659–5667.

6. Morrison,A., Araki,H., Clark,A.B., Hamatake,R.K. and Sugino,A. (1990) *Cell*, **62**, 1143–1151.
7. Hirose,F., Yamaguchi,M., Nishida,Y., Masutani,M., Miyazawa,H., Hanaoka,F. and Matsukage,A. (1991) *Nucleic Acids Res.*, **19**, 4991–4998.
8. Leegwater,P.A.J., Strating,M., Murphy,N.B., Kooy,R.F., van der Vliet,P.C. and Overdulve,J.P. (1991) *Nucleic Acids Res.*, **19**, 6441–6447.
9. Damagnez,V., Tillit,J., de Recondo,A.-M. and Baldacci,G. (1991) *Mol. Gen. Genet.*, **226**, 182–189.
10. Chung,D.W., Zhang,J., Tan,C.-K., Davie,E.W., So,A.G. and Downey,K.M. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 11197–11201.
11. Zhang,J., Chung,D.W., Tan,C.-K., Downey,K.M., Davie,E.W. and So,A.G. (1991) *Biochemistry*, **30**, 11742–11750.
12. Morrison,A. and Sugino,A. (1992) *Nucleic Acids Res.*, **20**, 375.
13. Stark,M.J.R., Mileham,A.J., Romanos,M.A. and Boyd,A. (1984) *Nucleic Acids Res.*, **12**, 6011–6030.
14. Paillard,M., Sederoff,R.R. and Levings,C.S.,III (1985) *EMBO J.*, **4**, 1125–1128.
15. Tommasino,M., Ricci,S. and Galeotti,C.L. (1988) *Nucleic Acids Res.*, **16**, 5863–5878.
16. Oeser,B. and Tudzynski,P. (1989) *Mol. Gen. Genet.*, **217**, 132–140.
17. Kempken,F., Meinhardt,F. and Esser,K. (1989) *Mol. Gen. Genet.*, **218**, 523–530.
18. Yoshikawa,H. and Ito,J. (1982) *Gene*, **17**, 323–335.
19. Jung,G., Leavitt,M.C., Hsieh,J.-C. and Ito,J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8287–8291.
20. Spicer,E.K., Rush,J., Fung,C., Reha-Krantz,L.J., Karam,J.D. and Konigsberg,W.H. (1988) *J. Biol. Chem.*, **263**, 7478–7486.
21. Matsumoto,K., Takano,H., Kim,C.I. and Hirokawa,H. (1989) *Gene*, **84**, 247–255.
22. Baer,R., Bankier,A.T., Biggin,M.D., Deininger,P.L., Farrell,P.J., Gibson,T.J., Hatfull,G., Hudson,G.S., Satchwell,S.C., Séguin,C., Tuffnell,P.S. and Barrell,B.G. (1984) *Nature*, **310**, 207–211.
23. Gibbs,J.S., Chiou,H.C., Hall,J.D., Mount,D.W., Retondo,M.J., Weller,S.K. and Cohen,D.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7969–7973.
24. Davison,A.J. and Scott,J.E. (1986) *J. Gen. Virol.*, **67**, 1759–1816.
25. Kouzarides,T., Bankier,A.T., Satchwell,S.C., Weston,K., Tomlinson,P. and Barrell,B.G. (1987) *J. Virol.*, **61**, 125–133.
26. Earl,P.L., Jones,E.V. and Moss,B. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3659–3663.
27. Binns,M.M., Stenzler,L., Tomley,F.M., Campbell,J. and Boursnell,M.E.G. (1987) *Nucleic Acids Res.*, **15**, 6563–6573.
28. Tomalski,M.D., Wu,J. and Miller,L.K. (1988) *Virology*, **167**, 591–600.
29. Bjornson,R.M., Glocker,B. and Rohrmann,G.F. (1992) *J. Gen. Virol.*, **73**, 3177–3183.
30. Gingeras,T.R., Sciaky,D., Gelinas,R.E., Bing-Dong,J., Yen,C.E., Kelly,M.M., Bullock,P.A., Parsons,B.L., O'Neill,K.E. and Roberts,R.J. (1982) *J. Biol. Chem.*, **257**, 13475–13491.
31. Brown,F. (1986) *Intervirology*, **25**, 141–143.
32. Costa,J.V. (1990) In Darai,G. (ed), Molecular Biology of Iridoviruses. Kluwer Academic, Boston, MA, pp. 247–270.
33. Kuznar,J., Salas,M.L. and Viñuela,E. (1980) *Virology*, **101**, 169–175.
34. Salas,M.L., Kuznar,J. and Viñuela,E. (1981) *Virology*, **113**, 484–491.
35. Polatnick,J. and Hess,W.R. (1972) *Arch. Gesamte Virusforsch.*, **38**, 383–385.
36. Marques,M.I. and Costa,J.V. (1992) *Virology*, **191**, 498–501.
37. Esteves,A., Marques,M I. and Costa,J.V. (1986) *Virology*, **152**, 192–206.
38. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Henikoff,S. (1984) *Gene*, **28**, 351–359.
40. Sanger,F., Coulson,A.R., Barrell,B.J., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161–178.
41. Spanos,A., Sedgwick,S.G., Yarranton,G.T., Hübscher,U. and Banks,G.R. (1981) *Nucleic Acids Res.*, **9**, 1825–1839.
42. Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
43. Marcy,A.I., Hwang,C.B.C., Ruffner,K.L. and Coen,D.M. (1990) *J. Virol.*, **64**, 5883–5890.
44. Bernad,A., Lázaro,J.M., Salas,M. and Blanco,L. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4610–4614.
45. Kozak,M. (1981) *Nucleic Acids Res.*, **9**, 5233–5252.
46. Corden,J.L., Cadena,D.L., Ahearn,J.M. and Dahmus,M.E. (1985) *Proc. Natl. Acad. Sci. U.S.A*, **82**, 7934–7938.
47. Allison,L.A., Moyle,M., Shales,M. and Ingles,C.J. (1985) *Cell*, **42**, 599–610.
48. Digard,P., Bebrin,W.R., Weisshart,K. and Coen,D.M. (1993) *J. Virol.*, **67**, 398–406.
49. Almazán,F., Rodríguez,J.M., Andrés,G., Pérez,R., Viñuela,E. and Rodríguez,J.F. (1992) *J. Virol.*, **66**, 6655–6667.
50. Rodríguez,J.M., Salas,M.L. and Viñuela,E. (1992) *Virology*, **186**, 40–52.
51. Grabherr,R., Strasser,P. and van Etten,J.L. (1992) *Virology*, **188**, 721–731.
52. Pisani,F.M., De Martino,C. and Rossi,M. (1992) *Nucleic Acids Res.*, **20**, 2711–2716.
53. Bernad,A., Blanco,L., Lázaro,J.M., Martn,G. and Salas,M. (1989) *Cell*, **59**, 219–228.