Acta Veterinaria Hungarica 48 (1), pp. 125–136 (2000)

SEQUENCE AND EXPRESSION ANALYSES OF THE UL37 AND UL38 GENES OF AUJESZKY'S DISEASE VIRUS

A. BRAUN[1], A. KALIMAN[2], ZS. BOLDOGKŐI[1], A. ASZÓDI[1] and I. FODOR[1,3*]

[1]Institute for Biochemistry and Protein Research, Agricultural Biotechnology Center, Gödöllő, Hungary; [2]Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Pushchino, Russia; [3]Center for Molecular Biology and Gene Therapy, Loma Linda University, Loma Linda, CA 92354, USA

(Received April 30, 1999; accepted June 24, 1999)

Previously, we sequenced the HSV-1 Ul39–Ul40 homologue genes of Aujeszky's disease virus (ADV), also designated as pseudorabies virus (Kaliman et al., 1994*a*, *b*). Now we report the nucleotide sequence of the adjacent DNA that encodes Ul38, the 5'-region (750 bp) of Ul37, and the promoter regions between these divergently arranged two genes. The ADV Ul38 gene encodes a protein of 368 amino acids. Amino acid sequence comparison of ADV Ul38 with that of other herpesviruses revealed significant structural homology. In a transcription study using RNase protection assay and Northern blot hybridization, we found that the Ul38 gene had one initiation site, but the Ul37 gene was initiated at two transcription sites with two potential initiator AUGs, one of which was dominant. Comparison of ADV Ul37, Ul38 and ribonucleotide reductase gene expression showed that these genes belong to the same temporal class with early kinetics. Data of structural and transcriptional studies suggest that regulation of the expression of these two ADV genes could differ from that of the HSV-1 virus.

Key words: Aujeszky's disease virus, DNA sequence, RNA protection assay, Northern blot, gene expression

Aujeszky's disease virus (ADV) is a herpesvirus which belongs to the subfamily of *Alphaherpesvirinae.* Aujeszky's disease is an important pig disease resulting in severe economic losses throughout the world. Although pigs are the natural host of the virus, many mammals can be infected by ADV. The genome of ADV is composed of a 142 kb linear double-stranded DNA molecule which contains two components, Ul (unique long) and Us (unique short), and two internal (IR) and terminal (TR) inverted repeats. ADV encodes at least 70 proteins, many of which have been sequenced. The gene arrangement of ADV and the prototype

*Correspondence: Istvan Fodor, Center for Molecular Biology and Gene Therapy, Loma Linda University, Mortensen Hall, Loma Linda, CA 92354, USA.
E-mail: ifodor@som.llu.edu; Fax: (909) 478-4177

herpesvirus, herpes simplex virus type 1 (HSV-1), is highly colinear. Amino acid sequence homology between ADV and HSV-1 proteins suggests that they share identical or similar functions in the biology of both viruses.

Analysis of the expression kinetics of HSV-1 mRNA species from each of the major temporal classes (α, β, γ) has indicated that gene expression is regulated primarily at the level of transcription (Honess and Roizman, 1974; Roizman and Sears, 1990). Most HSV-1 promoters have a recognizable TATA homology at approximately –25 to –30 relative to the cap site. For full levels of expression, promoters of the α, β and βγ (leaky-late) classes have demonstrated a requirement for upstream sequence elements (Coen et al., 1986; Blair and Snowden, 1991; Wagner, 1991). These cis-acting sites are typically binding sites for cellular transcription factors, e.g., Sp1 and the CAAT-binding protein. For expression of β and γ genes the HSV-1 α (immediate-early) genes are required (Wagner, 1991). The expression of Ul37 and Ul38 genes located near 0.55 map units on the HSV-1 genome is driven by two divergent promoters which belong to different kinetic classes (Flanagan et al., 1991). The promoter of the Ul 38 contains an important regulatory downstream activation sequence (DAS) (Guzowski et al., 1994). The Ul38 gene product is a DNA-binding protein which is involved in the packaging of the viral DNA into the viral particle (Braun et al., 1984), and in capsid assembly (Tatman et al., 1994; Thomsen et al., 1994). The Ul37 gene product is associated with the viral tegument (McLauchlan et al., 1994; Schmitz et al., 1995). Adjacent genes, rr1 and rr2, encode large (RR1) and small (RR2) subunits of the ribonucleotide reductase (Wagner, 1991).

The DNA sequences specifying the two subunits of the ADV ribonucleotide reductase have been determined and characterized earlier in our laboratory (Kaliman et al., 1994*a*, *b*). In the current study, we have extended the sequencing of this region of the ADV genome to the adjacent genes homologous to HSV Ul37 and Ul38, and have studied the expression of these genes at the transcriptional level.

Materials and methods

Cells and viruses

Strain Ka (Kaplan and Watter, 1959) of ADV was cultured in confluent monolayers of the porcine kidney (PK15) cell line. Cells were grown in Dulbecco's modified minimum essential medium (DMEM) supplemented with 5% fetal calf serum, 0.5 mg/l gentamicin, and 0.25 mg/l amphotericin-B at 37 °C with 5% CO_2. ADV from the medium of infected cells showing total cytopathic effect was purified by isopyknic centrifugation in a discontinuous gradient, as described previously (Kaliman et al., 1994*a*).

RNA isolation and Northern blot analysis

Total RNA was isolated from ADV-infected PK15 cells using guanidium isothiocyanate caesium chloride (Gilman, 1989). Cells were infected with ADV at a multiplicity of infection (MOI) of 1 or 10 plaque forming units (PFU) per cell. For Northern blot analysis, 10 µg of total RNA was fractionated by gel electrophoresis; blotting and hybridization were performed as described (Sambrook et al., 1989). Multiprime system (Amersham) and α-[^{32}P]dATP were used for DNA labelling. For the detection of mRNA transcribed from Ul38, the *Not*I-*Sal*I DNA fragment (Fig. 1) of plasmid pSP1was used as probe. The *Pst*I-*Xho*I fragment of p79P/28, containing the rr2 gene (Kaliman et al., 1994*a*), was used for detection and expression of *rr2.*

A

```
        HincII
   1  GTTGACCTCGCCCACGGGCACCAGCGCCGCCTCGTCCACGGCCCGCGCCGCCTCGGCGAA   60

  61  CTGCTCGCGCAGCGTCGGGTTCTGGCGGTCGTGGAAGAACTCGTCGGCGATCGTGAGGGC  120

 121  AAAGAGCGCGCGCGTCAGCGGCTTCAGGCCCTCGTTCTGCATCTGGCGCACCGCCAACGC  180

 181  CAGCGGCAGCGCCCAGTTGGTGTGCTGCGCCACGTAGCGCAGGCCGTCCTGCACAAAGTC  240

 241  CATGTCGTAGAGCGTCGCGAAGCGGGGCTGCACGAGCGCGCGCGCCGTCTGCTCCGGGGG  300

 301  CACGACGCGGGGCGCGGCCAGGACGTCGTCGGCGGAGCGCGCGCCCGCCAGGGCCTCGTC  360

 361  GAAGGCGGCGCACCACTCGCTCACGAGGCGGAACATGGCGTCGGGGCCGAACTCGGCCGA  420
                                                      HincII
 421  GGTGCGCAGCCCGGCCTCGGTGAGGGCGGCCGTGGTGTCGTCGCCGTTGACGACCCCCAG  480
       SacI
 481  GAGCTCCAGCAGCGGGGCCGTGCCCTCGTCCCAGTCGTGGCGCAGGCGCCACAGCGCCAG  540

 541  GCCGGCCAGGTTCTCGGCGAGCACGGCCGCCTCGCAGGTGCCCGTCTCGACGTAGGCGCG  600

 601  GCAGGCCAGGCGCAGCGCGCGCGCCCAGAGCCCGCGCACGCCCGCGGGCGTCTCGCGGCC  660

 661  CGCGGCCATGAAGAACTCCAGGATGCGCGACTGCACCACCGTGGCGACGGCGTGGTCGGC  720

 721  CTCCTCGAGCGCGCGCACGAGCGCCTCCATTATAAGAGGGTCCGCGCGCGGGCGCCGCCG  780

 781  CGCACTCGACCCCGCGCGGCGCGACGATGAGCGTGCAGATCGGCAACGGGCTGCTGATGG  840
                               M  S  V  Q  I  G  N  G  L  L  M  V

 841  TGGTCGCGCCGGGCACACTAACCGTGGGCTCGGCGCGCGCGCGCCTTATACGCCAGGTGA  900
        V  A  P  G  T  L  T  V  G  S  A  R  A  R  L  I  R  Q  V  T
                                          SmaI
 901  CGCTGGCGGACTTTTGCGAGCCCCAGGCCGAGCGCCCGGGGCTGGTGGTGCTCGCGCTGC  960
        L  A  D  F  C  E  P  Q  A  E  R  P  G  L  V  V  L  A  L  R

 961  GCCACCCCGCGGACCTGGCCGGCGCCGCCTACGCGGCCACGCCGCCCGGCAAGAACCACC 1020
        H  P  A  D  L  A  G  A  A  Y  A  A  T  P  P  G  K  N  H  R
1021  GCGACCTGGAGGAGGCGTGGCTCGCCCTCGACGAGGGCGGGCGCGGCCTCGGCGGCGACG 1080
        D  L  E  E  A  W  L  A  L  D  E  G  G  R  G  L  G  G  D  G
                                             NotI
1081  GCATCCGCGCCTCCGTCGTCTCGCTCAACTTCCTGGTGGCGGCCGCCGAGAACGCGGACG 1140
        I  R  A  S  V  V  S  L  N  F  L  V  A  A  A  E  N  A  D  D
```

```
1141  ACGCGCTGCGCGCGCACGTGACGACCAACTACCGCGACCGGCGCACGGCCGCGCGCCTCG 1200
        A  L  R  A  H  V  T  T  N  Y  R  D  R  R  T  A  A  R  L  E

1201  AGCGCTTCGCGACCGTGCTGCGCGCCATGATCCGGAGCCACGTGTTCCCGCACCGCGCGC 1260
        R  F  A  T  V  L  R  A  M  I  R  S  H  V  F  P  H  R  A  L

1261  TGCACGTCCTCGGGGGGCTCCTGGGCCACGTGACGCAAGACCGGCTGGCCAGCGTCACGT 1320
        H  V  L  G  G  L  L  G  H  V  T  Q  D  R  L  A  S  V  T  C

1321  GCGTGGCCCGCGGCGACCAGGAGGCGGCGCGCACCAACGACATGGCCGCGCGCCGCTCGC 1380
        V  A  R  G  D  Q  E  A  A  R  T  N  D  M  A  A  R  R  S  Q

1381  AGGTGCACGTGCCCGCGTGCGCGCTGATGGACGTGGACCGCGAGCTGCGCCTCGGCGGCG 1440
        V  H  V  P  A  C  A  L  M  D  V  D  R  E  L  R  L  G  G  D

1441  ACGACGGCCTCCGCTTCGCGTACCTGGTCTTTGTCTACACGCAGCGCCACCGCCGCGAGG 1500
        D  G  L  R  F  A  Y  L  V  F  V  Y  T  Q  R  H  R  R  E  A

1501  CGCTGCGCGTCCACGTGGCGGTGAGCCGCCTGCCCGAGCTCGGCGACGCCCTCAGCTTCC 1560
        L  R  V  H  V  A  V  S  R  L  P  E  L  G  D  A  L  S  F  L
                          SalI
1561  TCCTGGCCGGCACGCGCGTCGACAACGCGATCCACGGCACGGACGAGGCCGACGCCCCCG 1620
        L  A  G  T  R  V  D  N  A  I  H  G  T  D  E  A  D  A  P  A
1621  CCGCGCCCCGCCGCCGCCGCGGCCTTCCCCGCGTACCTGTTCAACGACCCGCGCAGCGCGC 1680
        A  P  A  A  A  A  A  F  P  A  Y  L  F  N  D  P  R  S  A  R

1681  GCTGCCCGACGGGCCGGCTCAACACGCCCGCCGCCGAGGCGCTGCCCGTGTGGGCCCCCG 1740
        C  P  T  G  R  L  N  T  P  A  A  E  A  L  P  V  W  A  P  D

1741  ACATGCGCGGCCGCGCCACCCGGAACTCGTGCATGTACGCCGCCTACGTGCGCCTCGGCA 1800
        M  R  G  R  A  T  R  N  S  C  M  Y  A  A  Y  V  R  L  G  T

1801  CCGTCGAGCGCGTCGTGCGCCGGGCCGAGCGCTGCGGCTCGGTGGACCTGCCGCTGGCCC 1860
        V  E  R  V  V  R  R  A  E  R  C  G  S  V  D  L  P  L  A  H

1861  ACATGGAGCGCTTCACCTGGGACGTGGGGGCGTGGGAGGAGTGTTTCTTCTGAAAAAACC 1920
        M  E  R  F  T  W  D  V  G  A  W  E  E  C  F  F  *

1921  GGGGGCGGGTGTGTGAGACGGATGTGATGTTGCTGACGAGGCTAATAAAAAGGCGGGCAC 1980

1981  ACGCGCGCGTGTCCCCCCGCTGACCCTCCGTGCCGCGTCCGCGTCGTGTGTGACTCACCC 2040

2041  CATCGTCTCTCCCGCCCGCGATCCCGGCCCGTCCCGGCTTGTCCCGCCCCGCCCAGACAC 2100

2101  ATCCCATCATG  2111
```

B

```
 900  TCACCTGGCG**TATAA**GGCGCGCGCGCGC**CGAGCCCACGGTTAGTGTGC**CCGGCGCGACCA  841

 840  CCATCAGCAGCCCGTTGCCGATCTGCACGCTCATCGTCGCGCCGCGCGGGGTCGAGTGCG  781

 780  CGGCGGCGCCCGCGCGCGGACCCTCT**TATA***ATG***GAGGCGCTCGTGCGCGCGCT**CGAGGAG  721
                                    **M**  E  A  L  V  R  A  L  E  E

 720  GCCGACCACGCCGTCGCCACGGTGGTGCAGTCGCGCATCCTGGAGTTCTTC*ATG*GCCGCG  661
      A  D  H  A  V  A  T  V  V  Q  S  R  I  L  E  F  F  M  A  A
 660  GGCCGCGAGACGCCCGCGGGCGTGCGCGGGCTCTGGGCGCGCGCGCTGCGCCTGGCCTGC  601
      G  R  E  T  P  A  G  V  R  G  L  W  A  R  A  L  R  L  A  C

 600  CGCGCCTACGTCGAGACGGGCACCTGCGAGGCGGCCGTGCTCGCCGAGAACCTGGCCGGC  541
      R  A  Y  V  E  T  G  T  C  E  A  A  V  L  A  E  N  L  A  G
                                                          SacI
```

```
540  CTGGCGCTGTGGCGCCTGCGCCACGACTGGGACGAGGGCACGGCCCCGCTGCTGGAGCTC  481
     L  A  L  W  R  L  R  H  D  W  D  E  G  T  A  P  L  L  E  L
              HincII
480  CTGGGGGTCGTCAACGGCGACGACACCACGGCCGCCCTCACCGAGGCCGGGCTGCGCACC  421
     L  G  V  V  N  G  D  D  T  T  A  A  L  T  E  A  G  L  R  T

420  TCGGCCGAGTTCGGCCCCGACGCCATGTTCCGCCTCGTGAGCGAGTGGTGCGCCGCCTTC  361
     S  A  E  F  G  P  D  A  M  F  R  L  V  S  E  W  C  A  A  F

360  GACGAGGCCCTGGCGGGCGCGCGCTCCGCCGACGACGTCCTGGCCGCGCCCCGCGTCGTG  301
     D  E  A  L  A  G  A  R  S  A  D  D  V  L  A  A  P  R  V  V

300  CCCCCGGAGCAGACGGCGCGCGCGCTCGTGCAGCCCCGCTTCGCGACGCTCTACGACATG  241
     P  P  E  Q  T  A  R  A  L  V  Q  P  R  F  A  T  L  Y  D  M

240  GACTTTGTGCAGGACGGCCTGCGCTACGTGGCGCAGCACACCAACTGGGCGCTGCCGCTG  181
     D  F  V  Q  D  G  L  R  Y  V  A  Q  H  T  N  W  A  L  P  L

180  GCGTTGGCGGTGCGCCAGATGCAGAACGAGGGCCTGAAGCCGCTGACGCGCGCGCTCTTT  121
     A  L  A  V  R  Q  M  Q  N  E  G  L  K  P  L  T  R  A  L  F

120  GCCCTCACGATCGCCGACGAGTTCTTCCACGACCGCCAGAACCCGACGCTGCGCGAGCAG   61
     A  L  T  I  A  D  E  F  F  H  D  R  Q  N  P  T  L  R  E  Q
                                                        HincII
 60  TTCGCCGAGGCGGCGCGGGCCGTGGACGAGGCGGCGCTGGTGCCCGTGGGCGAGGTCAAC    1
     F  A  E  A  A  R  A  V  D  E  A  A  L  V  P  V  G  E  V  N
```

Fig. 1. Nucleotide and deduced amino acid sequences of divergent Ul37 and Ul38 genes of ADV. **A**. DNA and complete amino acid sequences of the UL38 gene. **B.** The DNA sequences from nucleotide 1 to 900 shown in **A**, encoding the N-terminal portion of the open reading frames homologous to Ul37. The DNA sequence is shown as the rightward 5' to 3' strand and its nucleotides are numbered. Encoded protein sequences are shown below the DNA sequence. The Ul38 ORF extends from nucleotide 807 to 1910. The coding sequence of Ul37 extends from nucleotide 750 to 1 on the complementary DNA strand. The transcription start sites of the Ul38 and Ul37 homologues identified in this study are in bold and underlined in A and B**,** respectively. Putative TATA boxes upstream from each coding sequence are in bold. Relevant restriction sites are designated above the DNA sequence. The nucleotide sequence data have been submitted to the EMBL nucleotide sequence database and have been assigned Accession No. X80797

Sequencing and computer analysis

DNA sequences were isolated from pRL425 carrying the *Kpn*I-A fragment of ADV using subcloning strategy (Kaliman et al., 1994*a*). Both DNA strands of subclones were sequenced by the dideoxynucleotide chain termination method using T7 and Taq DNA polymerases, as described earlier (Kaliman et al., 1994*a*). In order to resolve band compression due to the high G+C content of the ADV genome (74%), deaza C-7 dGTP or dITP were also used as substrates. DNA sequences were analyzed using the sequence analysis software package GCG version 7.1 of the University of Wisconsin Genetics Computer Group (Devereux et al., 1985).

RNase protection assays

USB Ribonuclease Protection Kit was used for the RNase protection assays according to the protocol provided by the manufacturer. Labelled ^{32}P-RNA probes of Ul38 and Ul37 were prepared *in vitro* using the pCPrib vector. The vector con-

tained the ADV *Sac*I-*Sma*I fragment carrying the 5'- end fragments of both genes (Fig. 1). After hybridization with total RNA, excess single-stranded RNA probe and unhybridized sample RNA were removed by digestion with a mixture of RNases, and the products were separated on a denaturing 6% polyacrylamide gel, then visualized by autoradiography.

Results and discussion

Previously, we described the restriction map of the *Kpn*I-A fragment (29 kb) of ADV and determined the nucleotide sequence of the *rr1* (Ul39) and *rr2* (Ul40) genes (Kaliman et al., 1994*a*). In this report, we present the structure of the adjacent 2,118 bp sequence encoding the Ul37 and Ul38 homologue genes of ADV. Similar to other alphaherpesviruses, these two genes have a divergent arrangement. Computer analysis identified the Ul38 ORF extending from ATG_{807} to TGA_{1913} and encoding a polypeptide of 368 amino acids (Fig. 1A). The 250 N-terminal amino acids of a Ul37 ORF (1–750 bp) have also been determined (Fig. 1B). The intergenic region located between 750 (or 669) and 807 bp contains putative promoters for both genes. Multiple sequence alignment of the ADV Ul38 ORF and various herpesvirus homologues revealed high similarity in the 121–368 aa part of the protein, and high variability in the N-terminal part (Fig. 2). Similarly, alignment of the ADV Ul37 ORF and various herpesvirus homologues (data not shown) revealed high similarity in the corresponding proteins of alphaherpesviruses. Recent studies indicate that HSV-1 Ul38 and Ul37 are associated with virions: the Ul38 protein is involved in capsid assembly (Tatman et al., 1994; Thomsen et al., 1994) and the Ul37 protein is associated with the viral tegument (Schmitz et al., 1995). The functions of the Ul37 and Ul38 homologues of ADV are not known; however, due to high similarity between the proteins of both viruses, we anticipate that both ADV gene products are also associated with virions.

To characterize the expression of the sequenced genes in cells infected with ADV, Northern blot hybridization was applied for analysis of total RNA isolated from cells infected with the virus at a MOI of 1 PFU/cell. We found that both Ul38 and Ul37 genes were templates for monocistronic mRNAs (Fig. 3). We were not able to find a large readthrough mRNA encoding Ul38 and Ul39, as described for the HSV-1 (Flanagan et al., 1991), using either Ul38- (Fig. 3A, B) or Ul39-specific probe (data not shown), which implied that the expression regulation of the respective genes of the two viruses could be different. The determined size of the mRNA, 1.1 kb (Fig. 3A), is consistent with the one predicted from the DNA sequence of the Ul38 gene of ADV (see Fig. 1). The mRNAs of the homologous gene of HSV-1 and bovine herpesvirus 1 are longer (1.6–1.7 kb, Flanagan et al., 1991; Simard et al., 1995). The exact size of the Ul37 mRNA was not determined, but our estimate is 3.5–4.0, which is in accordance with the 3.6 kb found for HSV-1 Ul37 (Flanagan et al., 1991).

Temporal accumulation of the Ul38 and Ul37 gene products of ADV was analyzed by Northern blot hybridization of RNA isolated from cells infected with virus at a high MOI: 10 PFU/cell. The Ul38 and Ul37 mRNAs were detectable as soon as 2 h postinfection (Fig. 3B, C), indicating that they were expressed similarly, early in infection. In contrast, both the Ul37 and Ul38 gene products of HSV-1 are expressed with significantly different kinetics in productive infection: the HSV-1 Ul37 gene has been classified as a gene of the early kinetic class (Flanagan et al., 1991), while the HSV-1 Ul38 gene is regulated with late kinetics (Flanagan et al., 1991; Guzowski et al., 1994), suggesting that tegument proteins of HSV-1 function both in early and late events (packing and cleavage of the viral DNA) during virus development.

In view of the discrepancy observed in kinetics of ADV and HSV-1 Ul38 transcriptions, we compared the transcription kinetics of the ADV Ul38 with that of the *rr* genes (Ul39 and Ul40) regulated with early kinetics in all herpesviruses. Total RNA isolated from cells infected with low (1) and high (10) MOI of ADV was analyzed in Northern blot experiments using rr2-specific probe. Our experiments indicate two RNA transcripts of 3.6 and 1.1 kb, encoding both large (RR1) and small (RR2) subunits of RR, and only the small subunit (RR2), respectively (Fig. 4A). Both transcripts (RR1+RR2 and RR2) appear at the early stage of infection and accumulate during virus replication. Interestingly, at higher MOI the RR2 transcript appears at the very early stage of infection (Fig. 4B), and then the intensity of the RNA band decreases sharply. At lower MOI, the accumulation of the RR2 transcript is similar to that of RR1. Comparing the kinetics of RR2, Ul38 and Ul37 gene expressions, we conclude that they belong to the early viral genes of ADV. Additional studies will be performed to confirm this.

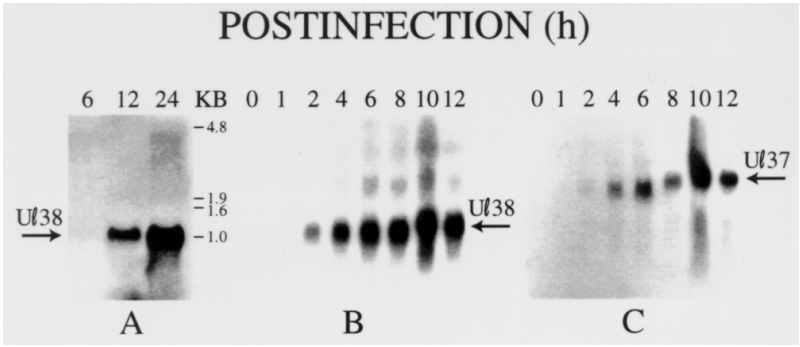
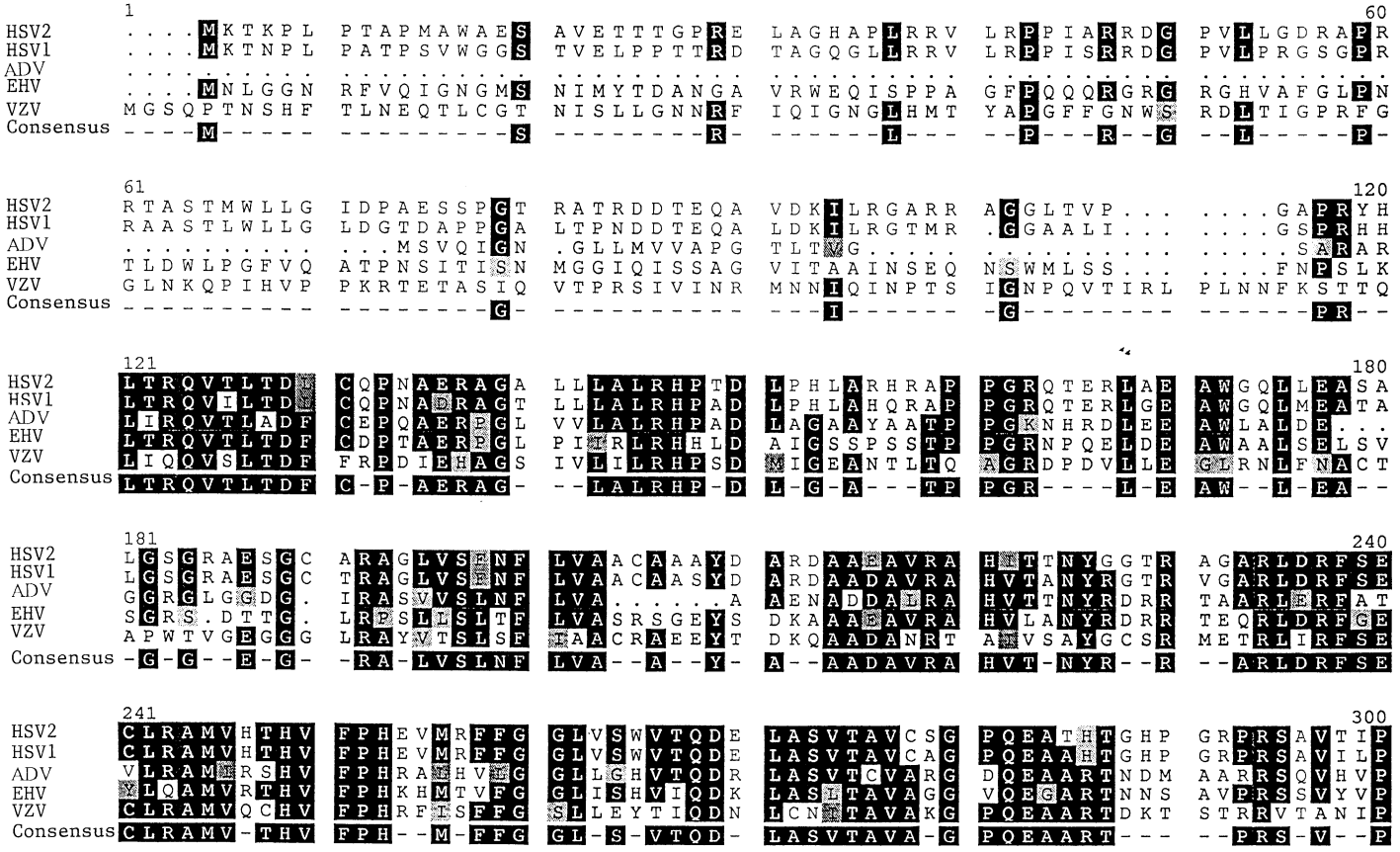


Fig. 3. Northern blot analysis of Ul38- and Ul37-specific RNAs of ADV. Positions of molecular weight markers are indicated on the right (A). RNA was purified from the virus-infected cells at the times postinfection (hours, h) indicated on the top. The multiplicity of infection (MOI) was 1 PFU/cell (A) or 10 PFU/cell (B, C). For further details see the text


1 60
HSV2 ....MKTKPL PTAPMAWAES AVETTTGPRE LAGHAPLRRV LRPPIARRDG PVLLGDRAPR
HSV1 ....MKTNPL PATPSVWGGS TVELPPTTRD TAGQGLLRRV LRPPISRRDG PVLPRGSGPR
ADV ....
EHV ....MNLGGN RFVQIGNGMS NIMYTDANGA VRWEQISPPA GFPQQQRGRG RGHVAFGLPN
VZV MGSQPTNSHF TLNEQTLCGT NISLLGNNRF IQIGNGLHMT YAPGFFGNWS RDLTIGPRFG
Consensus ----M----- ---------S --------R- ------L--- --P---R--G --L-----P-
61 120
HSV2 RTASTMWLLG IDPAESSPGT RATRDDTEQA VDKILRGARR AGGLTVP... ....GAPRYH
HSV1 RAASTLWLLG LDGTDAPPGA LTPNDDTEQA LDKILRGTMR .GGAALI... ....GSPRHH
ADV ....MSVQIGN .GLLMVVAPG TLTVG..... ....SARAR
EHV TLDWLPGFVQ ATPNSITISN MGGIQISSAG VITAAINSEQ NSWMLSS... ....FNPSLK
VZV GLNKQPIHVP PKRTETASIQ VTPRSIVINR MNNIQINPTS IGNPQVTIRL PLNNFKSTTQ
Consensus ---------- --------G- ---------- ---I------ -G-------- ------PR--
121 180
HSV2 LTRQVTLTDV CQPNAERAGA LLLALRHPTD LPHLARHRAP PGRQTERLAE AWGQLLEASA
HSV1 LTRQVILTDL CQPNADRAGT LLLALRHPAD LPHLAHQRAP PGRQTERLGE AWGQLMEATA
ADV LIRQVTLADF CEPQAERPGL VVLALRHPAD LAGAAYAATP PGKNHRDLEE AWLALDE...
EHV LTRQVTLTDF CDPTAERPGL PITRLRHHLD AIGSSPSSTP PGRNPQELDE AWAALSELSV
VZV LIQQVSLTDF FRPDIEHAGS IVLILRHPSD MIGEANTLTQ AGRDPDVLLE GLRNLFNACT
Consensus LTRQVTLTDF C-P-AERAG- --LALRHP-D L-G-A---TP PGR----L-E AW--L-EA--
181 240
HSV2 LGSGRAESGC ARAGLVSENF LVAACAAAYD ARDAAEAVRA HVTTNYGGTR AGARLDRFSE
HSV1 LGSGRAESGC TRAGLVSENF LVAACAASYD ARDAADAVRA HVTANYRGTR VGARLDRFSE
ADV GGRGLGGDG. IRASVVSLNF LVA..... AENADDALRA HVTTNYRDRR TAARLERFAT
EHV SGRS.DTTG. LRPSLLSLTF LVASRSGEYS DKAAAEAVRA HVLANYRDRR TEQRLDRFGE
VZV APWTVGEGGG LRAYVTSLSF IAACRAEEYT DKQAADANRT AIVSAYGCSR METRLIRFSE
Consensus -G-G--E-G- -RA-LVSLNF LVA--A--Y- A--AADAVRA HVT-NYR--R --ARLDRFSE
241 300
HSV2 CLRAMVHTHV FPHEVMRFFG GLVSWVTQDE LASVTAVCSG PQEATHTGHP GRPRSAVTIP
HSV1 CLRAMVHTHV FPHEVMRFFG GLVSWVTQDE LASVTAVCAG PQEAAHTGHP GRPRSAVILP
ADV VLRAMLRSHV FPHRAEHVEG GLLGHVTQDR LASVTCVARG DQEAARTNDM AARRSQVHVP
EHV YLQAMVRTHV FPHKHMTVFG GLISHVIQDK LASLTAVAGG VQEGARTNNS AVPRSSVYVP
VZV CLRAMVQCHV FPHRFTSFFG SLLEYTIQDN LCNITAVAKG PQEAARTDKT STRRVTANIP
Consensus CLRAMV-THV FPH--M-FFG GL-S-VTQD- LASVTAVA-G PQEAART--- --PRS-V--P

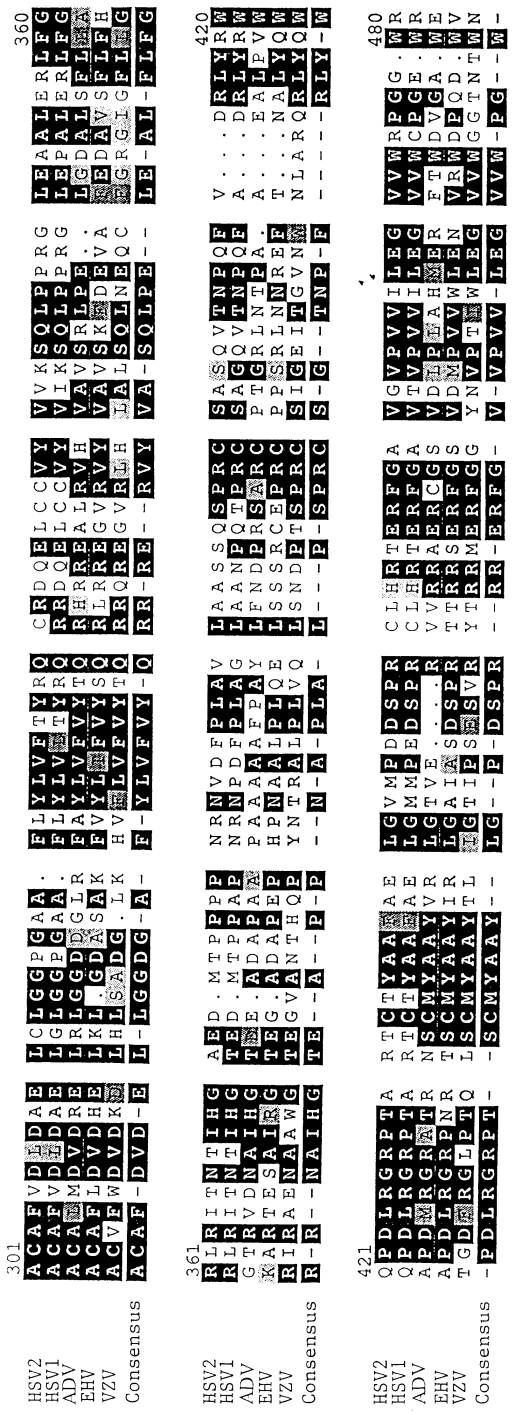


Fig. 2. Comparison of the deduced amino acid sequences of the UL38 homologues of ADV, herpes simplex virus type 1 (HSV1) and type 2 (HSV2), ADV, equine herpesvirus 1 (EHV) and varicella-zoster virus (VZV). Identical and similar residues between the sequences are shaded in black and grey, respectively

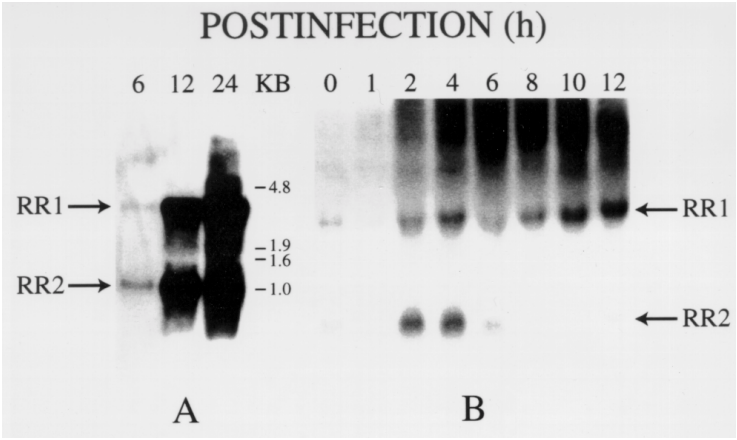


Fig. 4. Northern blot analysis of Ul39 (RR1)- and Ul40 (RR2)-specific RNAs of ADV. Positions of molecular weight markers are indicated on the right (A). RNA was purified at the times posti n-fection (h) indicated on the top. The MOI was 1 (A) or 10 (B) PFU/cell

To characterize the transcription of the Ul37 and Ul38 in more detail, RNase protection assays were used which allowed us to localize the 5'-end of the mRNAs. Labelled Ul38- and Ul37-specific probes were separately hybridized with the total RNA of ADV-infected cells followed by RNase treatment. The transcription site was deduced from the size of RNase-protected RNA fragment. For the Ul38, we obtained a 150–155-base RNase protected fragment, which allowed us to identify the transcription initiation site within the 5'-CCGCGCACTCGAC-3' sequence of the gene, at 16–29 base upstream from the closest ATG (Fig. 1A and 5A). Using a Ul37-specific probe, we obtained two truncated RNase protected fragments of 370–380 (minor) and 260 (major) bases (Fig. 5B). These fragments indicated that the Ul37 gene has two transcription initiation sites: a minor site within the sequence 5'-CGAGCCCACGGTTAGTGTGC-3', and a major site within the sequence 5'-GAGGCGCTCGTGCGCGCGCT-3' (Fig. 1B). Sequence analysis showed that the intergenic region contains two TATA-boxes, each of which could be linked to the separate translation (ATG) and transcription initiation sites of Ul37 (Fig. 1). Similarly to HSV-1 Ul38 promoter (Flanagan et al., 1991), the major start of ADV Ul37 is located only several bases downstream of the TATA-box. The distance between the two ATGs is 27 amino acid residues, which allows us to hypothesize that two polypeptides are translated from two overlapping alternate mRNAs. It is worthwhile to note that analysing the transcription site of the HSV-1 Ul37 the authors also observed a dominant (shorter) and a minor (larger) mRNA (Flanagan et al., 1991). We believe that these mRNAs of both viruses could be derived from alternate start sites and could play an important role in the regulation of the Ul37 gene.

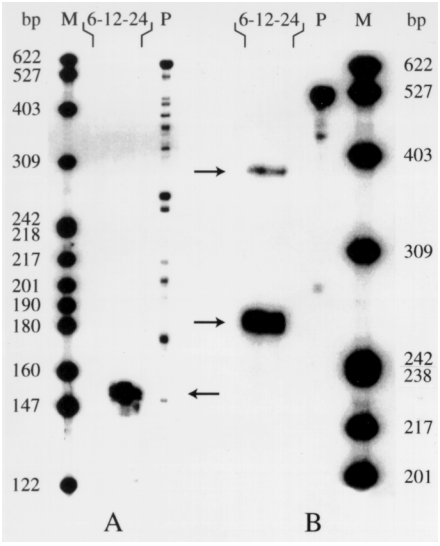


Fig. 5. Mapping of the 5' end transcripts encoding the Ul38 (A) and Ul37 (B) analogues of ADV by RNase protection assay. RNase protected RNA fragments are shown by arrows. In lane M, molecular weight standards are presented with the sizes in base pairs (bp). In lane P, the initial size of the labelled probe is shown. RNase protection assay was performed with the total RNA isolated from infected cells at 6, 12 and 24 h postinfection. The analyzed products in the corresponding lanes are designated as 6-12-24. Due to the close distance and identical patterns of the probes, the radioactive bands in these three lanes look confluent. For further details see the text

Acknowledgements

We thank Edit Erdélyi and Mariann Ihász for excellent technical help, and Sandra Hilliker for her critical reading of the manuscript. The work was partly supported by a Hungarian OTKA grant (T17095) and by the Hungarian Credit Bank Foundation for Hungarian Science.

References

Blair, E. D. and Snowden, B. W. (1991): Comparative analysis of the parameters that regulate expression from promoters of two late HSV-1 gene products. In: Wagner, E. K. (ed.) Herpesvirus Transcription and Its Regulation. CRC Press, Boca Raton, FL. pp. 181–206.

Braun, D. K., Batterson, W. and Roizman, B. (1984): Identification and genetic mapping of a HSV capsid protein that binds DNA. J. Virol. **58,** 645–648.

Coen, D. M., Weinheimer, S. P. and McKnight, S. L. (1986): A genetic approach to promoter recognition during trans induction of viral gene expression. Science **234,** 53–59.

Devereux, J., Haeberli, P. and Smithies, O. (1985): A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. **12,** 387–395.

Flanagan, W. M., Papavassilion, A. G., Rice, M. K., Hecht, L. B., Silverstein, S. and Wagner, E. K. (1991): Analysis of the herpes simplex virus type 1 promoter controlling the expression of UL38, a true late gene involved in capsid assembly. J. Virol. **65,** 769–786.

Gilman, M. (1989): Preparation and analysis of RNA. In: Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (eds) Current Protocols in Molecular Biology. Wiley-Interscience, New York, pp. 4.7.1–4.7.8.

Guzowski, J. F., Singh, J. and Wagner, E. K. (1994): Transcriptional activation of herpes simplex type 1 UL38 promoter conferred by the cis-acting downstream activation sequence is mediated by a cellular transcription factor. J. Virol. **68,** 7774–7789.

Honess, R. W. and Roizman, B. (1974): Regulation of herpesvirus macromolecular synthesis I. Cascaded regulation of the synthesis of three groups of viral proteins. J. Virol. **14,** 8–19.

Kaliman, A. V., Boldogkői, Zs. and Fodor, I. (1994*a*): Large and small subunit of the Aujeszky's disease virus ribonucleotide reductase: nucleotide sequence and putative structure. Biochim. Biophys. Acta **1219,** 151–156.

Kaliman, A. V., Boldogkői, Zs. and Fodor, I. (1994*b*): Structural features of the ribonucleotide reductase of Aujeszky's disease virus. Acta Vet. Hung. **42,** 251–161.

Kaplan, A. S. and Watter, A. E. (1959): A comparison of herpes simplex and pseudorabies virus. Virology **7,** 394–407.

McLauchlan, J., Liefkens, K. and Stow, N. D. (1994): The herpes simplex virus type 1 UL37 gene product is a component of virus particles. J. Gen. Virol. **75,** 2047–2052.

Roizman, B. and Sears, A. (1990): Herpes simplex viruses and their replication. In: Fields, B. Knipe, D. et al. (eds) Virology, 2nd ed. Raven Press, New York. pp. 1795–1841.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989): Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory, NY.

Schmitz, J. B., Albricht, A. G., Kinchington, P. R. and Jenkins, F. J. (1995): The UL37 protein of herpes simplex virus type 1 is associated with the tegument of purified virions. Virology **206,** 1055–1065.

Simard, C., Langlois, I., Styger, D., Vogt, B., Vlcek, C., Chalifour, A., Trudel, M. and Schwyzer, M. (1995): Sequence analysis of the UL39, UL38 and UL37 homologues of bovine herpesvirus 1 and expression studies of UL40 and UL39, the subunits of ribonucleotide reductase. Virology **212,** 734–740.

Tatman, J. D., Preston, V. G., Nicholson, P., Elliott, R. M. and Rixon, F. J. (1994): Assembly of herpes simplex virus type 1 capsids using a panel of recombinant baculoviruses. J. Gen. Virol. **75,** 1101–1113.

Thomsen, D. R., Roof, L. L. and Homa, F. L. (1994): Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. J. Virol. **68,** 2442–2457.

Wagner, E. K. (1991): Herpesvirus transcription – general aspects. In: Wagner, E. K. (ed.) Herpesvirus Transcription and Its Regulation. CRC Press, Boca Raton, FL. pp. 1–15.