Arch Virol (1997) 142: 17–35

Archives of Virology

© Springer-Verlag 1997
Printed in Austria

Identification and transcriptional analysis of pseudorabies virus UL6 to UL12 genes

J. M. Dijkstra, W. Fuchs, T. C. Mettenleiter, and **B. G. Klupp**

Institute of Molecular and Cellular Virology, Friedrich-Loeffler-Institutes,
Federal Research Centre for Virus Diseases of Animals, Insel Riems,
Federal Republic of Germany

Accepted August 28, 1996

Summary. We determined the nucleotide sequence of an 11 059 bp fragment of the pseudorabies virus genome located in the right part of genomic BamHI fragment 3 and the adjacent part of BamHI fragment 6. Within this region eight open reading frames were identified whose deduced amino acid sequences exhibited homology to the UL6, UL7, UL8, UL8.5, UL9, UL10, UL11, and UL12 protein products of herpes simplex virus type 1. Transcriptional analyses indicated presence of 3′-coterminal mRNAs for genes UL8, UL8.5, and UL9 as well as for genes UL6 and UL7, respectively, while UL10 was represented by a very abundant unique transcript. Both gene arrangement and transcriptional organization within this region of the pseudorabies virus genome thus parallels the situation found in other alphaherpesviruses.

Introduction

Pseudorabies virus (PrV) is an alphaherpesvirus [46] with great economic importance causing Aujeszky's disease in pigs [54]. The PrV genome consists of a linear double-stranded DNA of approx. 150 kbp. It is classified as a class II herpesviral genome since it is divided into a unique long (U_L) and a unique short (U_S) region by repeated sequence elements which bracket the U_S portion [7]. Within the alphaherpesvirus genomes gene arrangement appears to be largely collinear with the exception of an inversion in the U_L region [7] as well as differences found at one end of the U_L region [6, 15, 36, 52]. For the herpes simplex virus type 1 (HSV-1) [36], varicella-zoster virus (VZV) [15] and equine herpesvirus 1 (EHV-1) [52] genomes, complete sequence information is available, and the genomic sequence of bovine herpesvirus 1 (BHV-1) is currently being assembled [49, 53]. In contrast, the genomic sequence for PrV is not complete yet and it is estimated that approx. 70% of the viral genome have been sequenced [38]. One of the remaining gaps in sequence information is located in

genomic PrV BamHI fragment 3 between the UL5 [16] and UL13 [19] genes, and thus expected to contain genes encoding homologs of the HSV-1 UL6 to UL12 proteins.

HSV-1 UL6 is found associated with virus capsids [41], but it is not required for capsid assembly. Mutants defective in UL6 failed to process and encapsidate viral DNA and only immature capsids accumulated in the nucleus [42]. A protein product with unclear function is encoded by HSV-1 UL7. The homologous BHV-1 UL7 was identified as a nonessential gene, whose protein product is present in the cytoplasm of infected cells, but absent from mature virions [48].

Genetic and biochemical studies revealed that seven HSV-1 genes, UL5, UL8, UL9, UL29, UL30, UL42 and UL52 are required for viral DNA synthesis [40]. The UL8 protein is part of the helicase/primase complex formed by UL5 and UL52 gene products [14] and appears to be necessary for efficient primer utilization [51]. UL8 protein interacts specifically with the UL9 protein [37], the viral origin-binding protein [22]. A similar mechanism of DNA replication can be postulated for PrV since sequence determination of the PrV genomic fragment BamHI-4 revealed UL9 consensus recognition sequences located between the genes coding for gH (UL22) and UL21 [29]. A collinear location of this consensus sequence indicating presence of an origin of replication has also been described for EHV-1 [45], whereas in HSV-1, ori_L is found between the UL29 and UL30 genes [46]. Besides its origin-binding activity UL9 also functions as a DNA helicase on partially double stranded templates [11]. The protein appears to be organized into at least two separate functional domains. The C-terminal part is necessary for sequence-specific DNA binding [17] whereas several motifs characteristic for the helicase superfamily have been identified in the N-terminal part of the protein [34].

The HSV-1 UL10 gene encodes the virion glycoprotein gM [3] which has also been identified in PrV [21]. Gene UL11 of HSV-1 encodes a virion protein which is myristylated [31]. Functional studies indicate a role in virus egress [2]. The alkaline exonuclease, a DNase with a high pH optimum is encoded by HSV-1 UL12. Studies using a deletion mutant in UL12, which is able to replicate on noncomplementing cells, albeit with reduced titers, indicate a role of the UL12 protein in processing and packaging of viral DNA, probably by resolution of branched genomic structures [35, 50]. A PrV mutant defective in UL12 exhibited reduced virulence in mice [20].

For a detailed understanding of various aspects of herpesviral replication, as well as for efficient mutagenesis and vaccine development, molecular analysis of alphaherpesviral genomes is important. As a major step towards our goal to gain complete sequence information for PrV, an economically important pathogen in pig husbandry, we closed a gap between two already sequenced regions of the PrV genome [16, 19]. In this paper we describe the nucleotide sequence of PrV genes UL6, UL7, UL8, UL8.5, UL9, UL10, UL11, and UL12 as well as the transcriptional pattern of genes UL6 to UL10.

Materials and methods

Plasmids and sequence determination

Plasmids containing genomic BamHI fragments 3 and 6 (pBam3, pBam6) of PrV strain Kaplan (Ka) [27] were kindly provided by T. Ben-Porat, Nashville, Tenn. Genomic KpnI-fragment L which encompasses the junction between BamHI fragments 3 and 6 was cloned into vector pFBI-14 (Pharmacia, Freiburg, Germany). The approx. 17kb BamHI fragment 3 was subcloned after cleavage with KpnI and/or SalI into respective cleavage sites of plasmid TN-77, a pBR derivative containing the multiple cloning site of phage M13mp18 [39]. Fragments located in the right portion of BamHI fragment 3 (see Fig. 1) were sequenced as were the termini of the parental plasmids pBam3 and pBam6 and the part of KpnI-fragment L comprising the BamHI 3/BamHI 6 junction. For sequencing a set of overlapping subclones was prepared by exonuclease III/ S1 digestion [25] using a commercially available kit (Pharmacia, Freiburg, Germany). Sequence was determined by the dideoxy chain termination method [47] using double stranded plasmid DNA as template [24] and pBR322-specific oligonucleotides as primers (New England Biolabs, Eggenstein, Germany). Both DNA strands were sequenced at least twice on different nested deletion subclones. Sequences were assembled using the program *Assemble* of the Wisconsin Genetics Computer Group software package (GCG UNIX version 8) [18]. Open reading frames were predicted by programs *Frames* and *Condonpreference* and comparison of amino acid sequences was performed with *Gap*. Homology values as e.g. given in Table 2 represent percent identical amino acids. Multiple sequence alignments were performed with programs *Pileup* and *Pretty*. The nucleotide sequence described here has been submitted to the EMBL Data Library and assigned the accession number X97257.

RNA isolation and Northern blot hybridizations

For RNA analysis pig kidney cells (EFN-R) [44] were infected with PrV Ka at a multiplicity of infection (MOI) of 20. Cells were harvested 1, 2, 3, 4, and 5 h p.i., lysed, and whole cell RNA was isolated as described [28]. Control RNA was isolated from mock-infected cells. For hybridization double-stranded DNA fragments or single-stranded RNA probes were used as shown in Fig. 1. DNA probes (probes 1–3, 5, 6) were labeled with ^{32}P-dCTP using the Mega Prime labeling system (Amersham, Braunschweig, Germany). For probe 4 the UL10 open reading frame was cloned into vector pRc/CMV (InVitrogen, Leek, The Netherlands) and transcribed antisense in vitro using SP6 polymerase (Boehringer Mannheim, Germany) and ^{32}P-CTP.

Results

Nucleotide sequence and gene arrangement

We determined the nucleotide sequence of the right portion of BamHI fragment 3 and the adjacent part of BamHI fragment 6 of PrV strain Ka thus closing a gap between two regions of the PrV genome which have already been sequenced [16, 19]. Our sequence comprises 11 059bp with an overall G + C content of 75%. It is numbered from left to right with respect to the orientation of the PrV U_L region (Fig. 1). Numbering starts 324 bp upstream of the UL12 open reading frame which has partially been sequenced before [19] and ends 108bp upstream of UL6 in BamHI fragment 6 overlapping with the sequence published by Dean and Cheung [16].

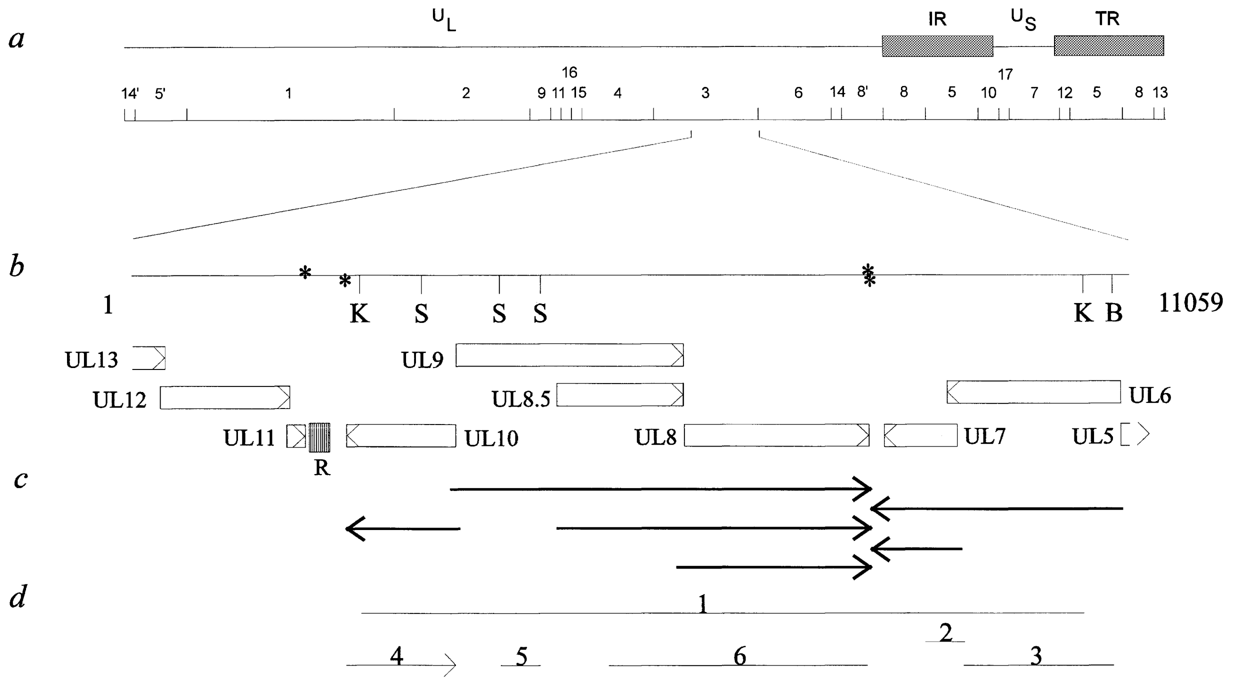

Fig. 1. Location of the PrV UL6 to UL12 gene cluster. **a** Schematic diagram of the PrV genome. Shaded boxes represent inverted repeat regions (TR = terminal repeat, IR = internal repeat), which bracket the unique short (U_S) region and separate it from the unique long region (U_L). A BamHI restriction fragment map is also presented. In **b** the region analyzed in this study is enlarged. Relevant restriction sites are given (*B* BamHI, *K* KpnI, *S* SalI). Positions of polyadenylation sites on both DNA strands are marked by asterisks. The location and orientation of the identified open reading frames is indicated as is the position of the repeat cluster (*R*). Arrows in **c** indicate location of the identified transcripts using hybridization probes shown in **d**

Computer analysis revealed several open reading frames (ORFs) whose deduced amino acid (aa) sequences showed homology to predicted translation products of HSV-1 genes UL6 to UL12 and corresponding homologs in other alphaherpesviruses [15, 36, 52]. We designated these genes according to the nomenclature in HSV-1 [5, 36]. Location of the predicted ORFs is depicted in Fig. 1. Properties of the deduced polypeptides are compiled in Table 1 and amino acid identities to homologous proteins of other selected herpesviruses are listed in Table 2.

An open reading frame whose deduced protein product shares 44% identical amino acids with the HSV-1 UL12 protein and 39% with the gene 48 product of VZV starts at nucleotide (nt) 325 and ends at nt 1 774. This gene, designated PrV UL12, has previously been partially sequenced in PrV strain NIA-3 [19]. No consensus TATA box [10, 13] or poly A-addition signal [9, 23] was found

Table 1. Properties of identified ORFs

Designation	TATA[a]	ATG[a]	Stop[a]	Poly A[a]	Amino acids	Molecular mass of deduced protein (kDa)
PrV-UL12	–[b]	325	1774	1930	483	51
PrV-UL11	1367	1734	1923	1930	63	7
PrV-UL10	3672	3585[c]	2406	2408	393	42
PrV-UL9	3485	3584	6113	8166	843	91
PrV-UL8.5	–[b]	4703[c]	6113	8166	470	51
PrV-UL8	5972	6112	8161	8166	683	71
PrV-UL7	9217	9129[c]	8331	8182	266	29
PrV-UL6	11010	10951	9022	8182	643	70

[a]First nucleotide of TATA-box, initiation codon, stop codon and poly A signal is indicated relative to the presented sequence
[b]For UL12 and UL8.5 no putative TATA-boxes have been detected
[c]Alternative initiation codons are present. The most likely (see text) has been indicated

Table 2. Amino acid identity of deduced PrV proteins with homologous herpesviral gene products

Designation	Percentage identity				
	HSV-1	VZV	EHV-1	HCMV	EBV
PrV-UL12	UL12: 44%	Gene 48: 39%	Gene 50: 51%	UL98: 29%	BGLF5: 29%
PrV-UL11	UL11: 31%	Gene 49: 25%	Gene 51: 40%	UL99: 25%	BBLF1: 23%
PrV-UL10	UL10: 32%	Gene 50: 36%	Gene 52: 40%	UL100: 24%	BBRF3: 24%
PrV-UL9	UL9: 50%	Gene 51: 47%	Gene 53: 57%	–[a]	–
PrV-UL8.5	UL8.5: 47%	–	–	–	–
PrV-UL8	UL8: 37%	Gene 52: 37%	Gene 54: 42%	UL102: 24%	BBLF2/BBLF3: 23%
PrV-UL7	UL7: 37%	Gene 53: 40%	Gene 55: 43%	UL103: 28%	BBRF2: 22%
PrV-UL6	UL6: 50%	Gene 54: 52%	Gene 56: 63%	UL104: 27%	BBRF1: 27%

[a]No homologous gene described

immediately upstream or downstream of this ORF, respectively. The predicted PrV strain Ka UL12 protein comprises 483 aa with a molecular mass of 51kDa. The gene is conserved throughout the herpesviruses and multiple sequence analysis of the deduced amino acid sequences shows several domains with high conservation as indicated in Fig. 2. The HSV-1 homolog is characterized by an amino terminal extension of approx. 100 residues with a high proline content.

```
             1                                                     50
PrV UL12     .......... .......... .......... .......... ..........
HSV-1 UL12   mestvgpacp pgrtvtkrpw alaedtprgp dsppkrprpn slpltttfrp
VZV gene 48  .......... .......... .......... .......... ..........
EBV BGLF5    .......... .......... .......... .......... ..........
HCMV UL98    .......... .......... .......... .......... ..........
Consensus    ---------- ---------- ---------- ---------- ----------
             51                                                   100
PrV UL12     .......... .......... .......... .......... ..........
HSV-1 UL12   lppppqttsa vdpsshspvn pprdqhatdt adekpraasp alsdasgppt
VZV gene 48  .......... .......... .......... .......... ..........
EBV BGLF5    .......... .......... .......... .......... ......madv
HCMV UL98    .......... .......... .......... .......... ......mwgv
Consensus    ---------- ---------- ---------- ---------- ----------
             101                                                  150
PrV UL12     .......... .......... .......maa LVlpn..... ....sLpeel
HSV-1 UL12   pdiplspggt hardpdadpd spdldsMwsa sVipn..... ....aLpshi
VZV gene 48  ..MarsglDr idispqpakk iarvggLqhp FVktd..... ....intinv
EBV BGLF5    deLe.dpmEe mtsy...... .......... .......... ..........
HCMV UL98    ssLdydddEe ltrllavwdd eplslfLmnt Fllhqegfrn lpftvLrlsy
Consensus    --L-----E- ---------- ------L--- FV-------- -----L----
             151                                                  200
PrV UL12     aarTFlRfLR gapr...... ..paaggaAp layRLaYvhd llveLarh..
HSV-1 UL12     laeTFeRhLR gllrgv.... ..Raplaigp lWaRLdYLCs lavvLeea..
VZV gene 48  .ehhFidtLq ktspnm.... ..dcrgMtAg iFiRLshMYk ilttLespnd
EBV BGLF5    ...TFaRfLR speteafvrn ldRppqMpAm rFvyLyCLCk qiqeFsgetg
HCMV UL98    ayriFakmLR ahgtpvaedf mtRvaaL.Ar deglrdiLgq rhaaeasrae
Consensus    ---TF-R-LR ---------- --R---M-A- -F-RL-YLC- ----L-----
             201                                                  250
PrV UL12     .glaapdaaa aaFggarppP apagvpaaAa .......raa iltVEaaTRa
HSV-1 UL12   .gmvdrglgr hlWrltrrgP .paaadavAp .......rpL MgFyEaaTqn
VZV gene 48  vtyttpgstn alFfktstqP qepRpeelAs kltqddikri LltIEseTRG
EBV BGLF5    fcdfvsslvq endskdgpsl ksiYwg..lq eatdeqrtvL csYVEsmTRG
HCMV UL98    iaealervae rcddrhggsd dyvWlsrlld lapnyrqveL FqLlEkesRG
Consensus    ---------- --F------P --------A- ---------L ---VE--TRG
             251                                                  300
PrV UL12     QSEsdLWtLL RRGLaTASTv rWgadGPrFP PtWceastar cg.......t
HSV-1 UL12   QaDcqLWaLL RRGLtTASTL rWgpqGPcFs PqWlkhnasl rp.......d
VZV gene 48  QgDNaiWtLL RRnLITASTL kWsvsGPviP PqWfyhhntt dt.......y
EBV BGLF5    QSENlMWdiL RnGiIssSkL lstikngptk vfepapistn ...hyfgg..
HCMV UL98    QSrNsvWhLL RmdtVsAtkF yeafvsgcLP gaaaadgsgg ggshytgsra
Consensus    QS-N-LW-LL RRGLITASTL -W---GP-FP P-W------- ----------
             301                                                  350
PrV UL12     pdnaAliFGr vNEsvARaaV aalYaeaptp dlpgaiaggD gGgdgAkeeM
HSV-1 UL12   vqssAVmFGr vNEptARsll fryCVg.... ......radD gGeagAdtrr
VZV gene 48  gdaaAmaFGk tNEpaARaiV ealFId.... ..padirtpD hltpeAttkF
EBV BGLF5    ....pVaFGl rcEdtvkdiV .cklI..... .......... cGdasAnrqF
HCMV UL98    gvspgIqFGi khEglvktlV ecyvm..... .......... hGrepvrdgL
Consensus    ----AV-FG- -NE--AR--V ----I----- ---------D -G---A---F
             351                                                  400
PrV UL12     FtFdetgapp pghdl..... ..fsCGLLlD PrTGMVGASL DLLVCdRDai
HSV-1 UL12   FiFhepsd.l aeenv..... ..htCGvLmD ghTGMVGASL DiLVCpRDih
VZV gene 48  FnFdmlntks psllvgtpri gtyeCGLLID vrTGLIGASL DvLVCdRDpl
EBV BGLF5    .......... .......... .....GFMIs PtdGifGvSL DLcVnvesqg
HCMV UL98    .......... .......... .....GLLID PtsGLlGASM DLcfgvlkqg
Consensus    F-F------- ---------- ----CGLLID P-TG-VGASL DLLVC-RD--
```

```
              401                                                     450
PrV UL12      .GrLaPhrtq TemrFFEIKC RAKYLFsaDD as.PtaraYa rLLeRPdadt
HSV-1 UL12    .GyLaP.vpk TplaFYEVKC RAKYaFdpmD psdPtasaYe dLMahrspea
VZV gene 48   tGtLnPhpaE TdisFFEIKC RAKYLFdpDD knnPlgrtYt tLinRPtman
EBV BGLF5     .dfi..lftD rsc.iYEIKC RfKYLFsksE .fdPiypsYt aLykRPckrs
HCMV UL98     sGrt..llvE pcarvYEIKC RyKYLrkkED ......pfvq nvLrRhdaaa
Consensus     -G-L-P---E T---FYEIKC RAKYLF--DD ---P----Y- -LL-RP----
              451                                                     500
PrV UL12      LRgFLySIar PgVEFFega. .PgpgEALaT aDpAWrRgga edApptrrrc
HSV-1 UL12    FRaFirSIpk PsVrYFapgr vPgpeEALvT qDqAW...se ahAsgeKrrc
VZV gene 48   LRdFLytIkn PcVsFFgpsa nPStrEALiT dhveWkRlgF KggRal....
EBV BGLF5     FirFinSIar PtVEYvpdgr lPSegDyLlT qDeAW...nL KdvRkrK.lg
HCMV UL98     vasLLqShpv PgVEFrgere tPSarEfLls hDaAlfRatL KrARplKppe
Consensus     -R-FL-SI-- P-VEFF---- -PS--EAL-T -D-AW-R--L K-AR--K---
              501                                                     550
PrV UL12      gafD..grhv aaNahaqSEV WlFsdPvdgr qdI..VpWas GEra......
HSV-1 UL12    saaD..ralv eLNsgVvSEV llFgaPdlgr htIspVsWss GDlv......
VZV gene 48   telD..AhhL gLNrtIsSrV WVFndPdiqk gtIttIaWat GDta......
EBV BGLF5     pghDlvAdsL aaNrgVeSml YVm......t dpsenagRig ik........
HCMV UL98     plrEylAdlL yLNkaecSEV iVFdakhlsd dnsdgdatit inaslglaag
Consensus     ---D--A--L -LN--V-SEV WVF--P---- --I--V-W-- GD--------
              551                                                     600
PrV UL12      .......... .......... .......... .......... ...LrVPVFA
HSV-1 UL12    .......... .......... .......... .......... ...rrePVFA
VZV gene 48   .......... .......... .......... .......... ...LqIPVFA
EBV BGLF5     .......... .......... .......... .......... .drvpVnIFi
HCMV UL98     dgagggadhh lrgspgdspp pipfedentp ellgrlnvye varFslPaFv
Consensus     ---------- ---------- ---------- ---------- ---L-VPVFA
              601                                                     650
PrV UL12      NPRHaNFrQI LVQsYVvagv FP........ ...DrpaRPh LaTFfGRrRr
HSV-1 UL12    NPRHpNFkQI LVQgYVLdsh FP........ ...DcpPhPh LVTFIGRhRt
VZV gene 48   NPRHaNFkQI aVQtYVLsgY FP........ ...alklRPf LVTFIGRvRr
EBV BGLF5     NPRHnyFyQV LlQykIvgdY v..rhsgggk pgrDcsPRvn iVTaffRkRs
HCMV UL98     NPRHqyYfQm LIQqYVLsqY Yikkhpdper idfrdlPtvy LVsaIfReR.
Consensus     NPRH-NF-QI LVQ-YVL--Y FP-------- ---D--PRP- LVTFIGR-R-
              651                                                     700
PrV UL12      PcEqnrtLdL .......... .aslcdVpPa cAVPVlLIVT PVsVc.eeaf
HSV-1 UL12    saEegvtFrL EdgagaLGaa gpskasIlPn qAVPIaLIIT PVrID.peiy
VZV gene 48   PhEvgvpLrv Dtqaaaiyey ..nwptIpPh cAVPViaVlT PIeVDvprvt
EBV BGLF5     P........L DpatctLGsd llldasV... .eIPVavlVT PVvlpd....
HCMV UL98     .......... ..eeseLGce llaggrVfhc dhIPllLIVT PVvfDp....
Consensus     P-E----L-L D-----LG-- ------V-P- -AVPV-LIVT PV-VD-----
              701                                                     750
PrV UL12      edlRaraeeA Frvt...... ....AsRtwd SvaAdspaTa s.........
HSV-1 UL12    kaiqrssrlA FddtlaeLw. ....AsRsPg pgpAaaetTs ssptttgrssr
VZV gene 48   qilkdtgnnA itsalrsLrw dnlhpaveee SvdcangtTs llratekpll
EBV BGLF5     sviRktlstA agswkayadn tfdtApWvPs glfAddesTp ..........
HCMV UL98     qftRhavstv LdrwsrdLsr ktnlpiWvPn Saneyvvssv prpvsp....
Consensus     ---R-----A F------L-- ----A---P- S--A----T- ----------
```

Fig. 2. Amino acid comparison of UL12 homologous proteins. Multiple sequence analysis was performed using the program *Pileup*. A consensus sequence created by using the program *Pretty* is given below. Deduced amino acid sequences of HSV-1 UL12 [36], PrV UL12 (this study), VZV gene 48 [15], EBV BGLF5 [1], and HCMV UL98 [12] were compared. Identical or similar amino acids in at least three of five compared sequences are shown as upper case letters and residues conserved in all sequences are marked with shaded boxes. Conserved regions are boxed. Gaps were introduced for best alignment

```
              1                                                      50
PrV UL9       ....MAADAG GRRGG....A EEDGGAFASS VSLARMLYGC DLPAVVRSRW
HSV-1 UL9     MPFVGGAESG DPLGAGRPIG DDECEQYTSS VSLARMLYGG DLAEWVPRVH
VZV gene 51   .......... ....MSPNTG ESNAAVYASS TQLARALYGG DLVSWIKHTH

                                           Motif I
              51                                                    100
PrV UL9       FGVSLDLQRD APVELPSPHD TACRRVLVAR APMGSGKTTA LLKWLSAALA
HSV-1 UL9     PKTTIERQQH GPVTFPNASA PTARCVTVVR APMGSGKTTA LIRWLREAIH
VZV gene 51   PGISLELQLD VPVKLIKPGM SQTRPVTVVR APMGSGKTTA LLEWLQHALK

              101                                                   150
PrV UL9       ATDMSALVLS CRRSFTRTLA RRMDDAGL.G FVTYFDSDAY VMTGRPYRRL
HSV-1 UL9     SPDTSVLVVS CRRSFTQTLA TRFAESGLVD FVTYFSSTNY IMNDRPFHRL
VZV gene 51   A.DISVLVVS CRRSFTQTLI QRFNDAGLSG FVTYLTSETY IMG...FKRL
                                                                      *

                                Motif II
              151                                                   200
PrV UL9       LVQIESLHRV DEHLINNYDV IVVDEVMSTL GQLYSPTMAR LARVDALLAR
HSV-1 UL9     IVQVESLHRV GPNLLNNYDV LVLDEVMSTL GQLYSPTMQQ LGRVDALMLR
VZV gene 51   IVQLESLHRV SSEAIDSYDV LILDEVMSVI GQLYSPTMRR LSAVDSLLYR
                   *         *        *

                         Motif III
              201                                                   250
PrV UL9       LIRGCPRVLV MDATINAQLV ELLVELRGEP SVHVVVSDYA TTAFASRRCL
HSV-1 UL9     LLRICPRIIA MDATANAQLV DFLCGLRGEK NVHVVVGEYA MPGFSARRCL
VZV gene 51   LLNRCSQIIA MDATVNSQFI DLISGLRGDE NIHTIVCTYA GVGFSGRTCT

              251                                                   300
PrV UL9       VLRHLGAEVA AGAAGAREDG GGDGSEDAAR AGSPAPTTAA ATTAVEAAGA
HSV-1 UL9     FLPRLGTELL QAALRPPGPP SGPSPD.... .......... .....ASPEA
VZV gene 51   ILRDMGIDTL VRVIKRSPEH EDVRTIHQLR .......... ..........

                                 Motif IV
              301                                                   350
PrV UL9       AGDSFFGLLG ARLAAGDNVC VFSSTLAFSE LVARFCARFT PSVLVLNSQR
HSV-1 UL9     RGATFFGELE ARLGGGDNIC IFSSTVSFAE IVARFCRQFT DRVLLLHSLT
VZV gene 51   ..GTFFDELA LRLQCGHNIC IFSSTLSFSE LVAQFCAIFT DSILILNSTR

                              Motif V
              351                                                   400
PrV UL9       PPEDVGRWA. VRALVYTTVV TVGLSFDAPH FHSMFAYVKP MAHGPDMASV
HSV-1 UL9     PLGDVTTWGQ YRVVIYTTVV TVGLSFDPLH FDGMFAYVKP MNYGPDMVSV
VZV gene 51   PLCNVNEWKH FRVLVYTTVV TVGLSFDMAH FHSMFAYIKP MSYGPDMVSV

              401 Motif VI                                          450
PrV UL9       YQSTGRVRRL LRDELFVYVD GSGARGEPIF TPVLLNHVVG S..GWPARLS
HSV-1 UL9     YQSLGRVRTL RKGELLIYMD GSGARSEPVF TPMLLNHVVS SCGQWPAQFS
VZV gene 51   YQSLGRVRLL LLNEVLMYVD GSRTRCGPLF SPMLLNFTIA NKFQWFPTHT

              451                                                   500
PrV UL9       QVTNLVCAQF QRRCRPAFAA ...ARGMRLF SRFKFKHLFE RCTLTSVNDS
HSV-1 UL9     QVTNLLCRRF KGRCDASACD TSLGRGSRIY NKFRYKHYFE RCTLACLSDS
VZV gene 51   QITNKLCCAF RQRCANAFTR S....NTHLF SRFKYKHLFE RCSLWSLADS

              501                                                   550
PrV UL9       LNILHALLEN NRLRVALEGC EP..PLTARA FCDFLRDARL DAFASQQVLR
HSV-1 UL9     LNILHMLLTL NCIRVRFWGH DD..TLTPKD FCLFLRGVHF DALRAQRDLR
VZV gene 51   INILQTLLAS NQILVVLDGM GPITDVSPVQ FCAFIHDLRH SANAVASCMR
```

```
              551                                                    600
PrV UL9       QL..RPPDRP VAADIADSG. ...EVATFVE KYLVADVPED ELQELLRALA
HSV-1 UL9     ELRCRDPEAS LPAQAAETE. ...EVGLFVE KYLRSDVAPA EIVALMRNLN
VZV gene 51   SL..RQDNDS CLTDFGPSGF MADNITAFME KYLMESINTE EQIKVFKALA

              601                                                    650
PrV UL9       NPVTREQFVG LAVLGACARV PEALRSERVF GAVYGHYASG AVPVV.ADGR
HSV-1 UL9     SLMGRTRFIY LALLEACLRV PMATRSSAIF RRIYDHYATG VIPTINVTGE
VZV gene 51   CPIEQPRLVN TAILGACIRI PEALEAFDVF QKIYTHYASG WFPVLDKTGE

              651                                                    700
PrV UL9       LELAALAPDF NVPARWALTR RCARVAEAAG LFEGASPEVD SAAVAAAAAD
HSV-1 UL9     LELVALPPTL NVTPVWELLC LCSTMA.ARL HWDSAAGGSG RTFGPDDVLD
VZV gene 51   FSIATITTAP NLTTHWELFR RCAYIAKT.L KWNPSTEGCV TQVLDTDINT

              701                                                    750
PrV UL9       ......AELA PLLLEVLRCH VLDATTAARR PVRAALSALG AGGG...AGP
HSV-1 UL9     LLTPHYDRYM QLVFELGHCN VTDGLLLSEE AVKRVADALS GCPP...RGS
VZV gene 51   LFNQHGDSLA QLIFEVMRCN VTDAKIILNR PVWRTTGFLD GCHNQCFRPI

              751                                                    800
PrV UL9       LSRGRHAALV FKVMWEEAFG VRVGRSRQTF PGPTRVKNLR KAEIAALLRD
HSV-1 UL9     VSETDHAVAL FKIIWGELFG VQMAKSTQTF PGAGRVKNLT KQTIVGLLDA
VZV gene 51   PTKHEYNIAL FRLIWEQLFG ARVTKSTQTF PGSTRVKNLK KKDLETLLDS

              801                                                    850
PrV UL9       AGLDPPAGAT HRQLYALLME RRGDFAGERY KLRLPAWSRL MYL....TQ.
HSV-1 UL9     HHIDHSACRT HRQLYALLMA HKREFAGARF KLRVPAWGRC LRTHSSSAN.
VZV gene 51   INVDRSACRT YRQLYNLLMS QRHSFSQQRY KITAPAWARH VYFQAHQMHL

              851                                  882
PrV UL9       .GGFDAPLDA ALSLVPAEAW PRTEGAVDFA AL
HSV-1 UL9     .PNADIILEA ALSELPTEAW PMMQGAVNFS TL
VZV gene 51   APHAEAMLQL ALSELSPGSW PRINGAVNFE SL
```

Fig. 3. Amino acid sequence comparison of origin binding proteins. Comparison of the predicted amino acid sequences of PrV UL9, HSV-1 UL9 [36] and the VZV gene 51 product [15] is shown. Conserved amino acids in all sequences compared are boxed. The six highly conserved motifs of DNA and RNA helicases are numbered. Leucine residues which may form the leucine zipper motif are indicated by asterisks. Gaps were introduced for best alignment. Arrows point to possible start methionines for PrV UL8.5. The predicted start for HSV-1 UL8.5 [5] is marked by black arrow

A small ORF located in the same orientation and partially overlapping UL12 is predicted to encode a protein with a molecular mass of 7kDa comprising 63 aa. A putative TATA box is located at nt 1 367–1 372 (5′-TATTTA-3′). A poly-adenylation signal (5′-AATAAA-3′) is found immediately downstream of the gene at nt 1 930–1 935. Based on the overall collinearity of the alphaherpesvirus genomes this ORF was named UL11, although amino acid identity to the

deduced HSV-1 UL11 protein is limited (31%) and restricted to the N-terminal part (data not shown). The HSV-1 UL11 protein has been shown to be myristylated and deduced UL11 homologous proteins of VZV (gene 49) [15], EHV-1 (gene 51) [52], HCMV (UL99) [12], and EBV (BBLF1) [1], including PrV (this study) all contain a conserved glycine residue which is predicted to carry the myristyl moiety [31].

Downstream of the UL11 poly A signal a cluster of repeated elements is located in which the sequence 5′-GGGGGAGAGGAT-3′ is repeated 18 times (R; Fig. 1).

Sequence analysis of the PrV UL10 gene which encodes the abundant virion glycoprotein gM has been described recently [21]. The UL10 open reading frame starts at nt 3 585 and ends at nt 2 406.

In opposite orientation to UL10 with overlapping start codons a large open reading frame is located whose deduced translation product shows strong homology to the HSV-1 UL9 protein. The open reading frame starts at nt 3 584 and ends at nt 6 113. A potential TATA box, 5′-ATAA-3′, is located 99bp upstream, but a consensus sequence for polyadenylation was not detected immediately downstream of the ORF. The PrV UL9 protein product comprises 843 aa with a predicted molecular mass of 91 kDa. Homology to respective proteins in other alphaherpesviruses is high and amounts to 50% identical amino acids with the HSV-1 UL9 polypeptide and 47% with the VZV gene 51 product. Similar to the HSV-1 UL9, the VZV gene 51 and the EHV-1 gene 53 products which constitute the origin-binding protein [11, 22, 33], the deduced PrV UL9 protein specifies helicase motifs as well as a leucine zipper motif as shown in Fig. 3. No UL9 homolog has been found in Beta-and Gammaherpesviruses, with the exception of human herpesvirus 6 (HHV-6B), a betaherpesvirus [26].

In HSV-1 an open reading frame designated as UL8.5 that overlaps and is in-frame with the 3’-terminal half of the UL9 gene has been described [5]. In PrV, possible start codons at positions 4 703, 4 724, and 4 742 are located within and in-frame with UL9 (see Fig. 3). Thus, the PrV UL8.5 protein could consist of the carboxyterminal 470, 463, or 457 aa of the UL9 product, respectively. As judged from the sequence, the third initiation codon is in the most favourable translation initiation context [30].

Downstream of UL9 in the same transcriptional orientation an ORF starts at nt 6112, overlapping by one base with the UL9 gene, and ends at nt 8161. A putative TATA box, 5′-TATAAA-3′, is situated 140 nt upstream and a poly A signal is found at nt 8166–8171 immediately downstream from the ORF. The deduced 683 aa protein shows 37% amino acid identity to the HSV-1 UL8 protein and to the VZV gene 52 product.

PrV UL7 is transcribed in opposite orientation to UL8. Two in-frame start codons were detected at nt 9 129 and nt 9 117. Both are in favourable translation initiation context according to the rules of Kozak [30]. Assuming that translation starts at the first ATG (nt 9129) the gene is predicted to encode a protein of

266 amino acids with a molecular mass of 29kDa. The sequence 5′-TTTAA-3′ located at nt 9 217–9 213 might function as TATA box while a poly A signal is located at nt 8 182–8 177. Comparison with the UL7 homologs in the other alphaherpesviruses shows amino acid sequence identities between 37% (HSV-1) and 43% (EHV-1).

In the same transcriptional orientation, and partially overlapping the UL7 gene, resides the UL6 ORF. The sequence 5′-TAAA-3′ located at nt 11 010–11 007 could represent a TATA box. With a start codon at nt 10 951 and stop codon at nt 9022 the gene is predicted to specify a 643 aa protein with a molecular mass of 70kDa. Homology to the HSV-1 UL6 and VZV gene 54 proteins amounts to 50% and 52% identical amino acids, respectively. Multiple sequence analysis shows several motifs conserved throughout the herpesviruses (Fig. 4).

Transcriptional analysis

To determine the transcriptional pattern in the sequenced region, Northern blot analyses were performed. For that purpose whole-cell RNA of PrV infected EFN-R cells was isolated at 1, 2, 3, 4, and 5 h after infection (Fig. 5, lanes 1–5). As negative control, mock-infected cell RNA was also assayed (Fig. 5, lanes 0). Location of the hybridization probes is shown in Fig. 1.

Probe 1, comprising most of the sequenced region, hybridized to six RNAs of 1.3, 1.6, 2.4, 3.1, 3.6 and 5.1 kb (Fig. 5A). The most abundant transcript, 1.6 kb in size, was first detectable at 2 h p.i. and increased in amount until 5 h p.i. The 1.3 kb transcript could be detected from 3 h p.i. until 5 h p.i. The 2.4 kb mRNA was already visible at 1 h p.i., increased in amount until 3 h p.i., and decreased thereafter. The 3.1kb RNA showed similar kinetics as the 1.6kb transcript but appeared to be less abundant. The 3.6kb transcript was first detected at 3 h p.i. and increased until 5 h p.i. The 5.1 kb transcript gave only a very faint signal being barely visible at 1 h p.i. and increasing in amount until 3 h p.i.

To assign transcripts to the predicted genes more specific hybridization probes were used as indicated in Fig. 1. Probe 2 containing only UL7 sequences hybridized to the 3.1 kb and 1.3 kb transcripts (Fig. 5B) while probe 3, specific for UL6, detected the 3.1kb transcript only (Fig. 5C). This indicates that the 1.3kb mRNA represents the UL7 transcript, and the 3.1 kb transcript is structurally bicistronic encompassing UL6 and UL7. Based on Northern blot and sequencing data, both transcripts appear to be 3′-coterminal, which is also indicated by presence of a polyadenylation signal downstream from UL7. Probe 4, a single stranded RNA antisense to the UL10 gene (Fig. 5D), detected the very abundant 1.6kb transcript. No signal was detectable with a probe in sense orientation (data not shown). Additional Northern blots identified the 5.1kb RNA as the UL9 transcript, the 2.4kb RNA as the UL8 mRNA, and a 3.6kb RNA as the putative UL8.5 transcript (data not shown). A summary of the transcript map is shown in Fig. 1.

```
             1                                                50
PrV UL6      .......... .......msA ataaaadgLc pggaaEeann lLg.......
HSV-1 UL6    .......... .......mtA prsrapttra rgdtealcs. ..........
VZV gene 54  maeitslfnn ssgseekriA ssvsidqgLn gsnpnDqykn mFdiywneya
EBV BBRF1    .......... .......... ........Mf nmnvdEsasg aLg.......
HCMV UL104   .........m ernhwnekss gakrsrerdl tlstirsila aderlrikas
Consensus    ---------- ---------A --------L- -----E---- -L--------
             51                                              100
PrV UL6      .......... apsRIlIHPT PrTMvFkEIL mGnLGYTEGQ GIYnsVRStE
HSV-1 UL6    .........p EDgWVkVHPs PgTMlFrEIL hGqLGYTEGQ GVYnvVRSsE
VZV gene 54  pdigfctfpe EDgWmlIHPT tqsMlFrkIL aGdFGYTDGQ GIYsaVRStE
EBV BBRF1    .......... .ssaIpVHPT PasvrLfEIL qGkYaYvqGQ tIYanlRnpg
HCMV UL104   sylgvgrgvd DEavIdIfPT gqTMsFlrlL hGfLGtcrGQ smhqvlRdpc
Consensus    ---------- ED-WI-IHPT P-TM-F-EIL -G-LGYTEGQ GIY--VRS-E
             101                                             150
PrV UL6      aavRQIQstl LtrtLnAarY EDVArDWmaH lrargLgaeA LarRFg....
HSV-1 UL6    attRQlQaaI FhalLnAtTY rDleaDWlgH vaargLqpqr LvrRYrnarE
VZV gene 54  tviRQVQatV LmnaLdAtrY EDlAaDWehH iqqcnLhagA LaeRYglcgE
EBV BBRF1    vfsRQVfthl FkraishcTY DDVlhDW... ....nkfeac iqkRWp.sdD
HCMV UL104   vlrkQllygV cktlFdtiTv rrVAeEWklH .........A alfpYraldE
Consensus    ---RQVQ--V ----L-A-TY EDVA-DW--H -----L---A L--RY----E
             151                                             200
PrV UL6      dEApavAErl FDTWyrTLQm aLLDFvRGIa aCFsasesnG taSFaKYIDW
HSV-1 UL6    aDiagvAErv FDTWRnTLrt tLLDFahGlv aCFapggpsG psSFpKYIDW
VZV gene 54  sEAvrLAhqv FETWRqTLQs sLLEFLRGIt gCLytsglnG rvgFaKYVDW
EBV BBRF1    scAsrFrEst FEsWstTMkl tvrDLL..tt niYrvlhsrs vlSYerYVDW
HCMV UL104   eDleqYllvw saslRqsvQt gvLggLRdIl yqYadnd... ..dYglYVDW
Consensus    --A---AE-- F-TWR-TLQ- -LLDFLRGI- -C-------G --SF-KYVDW
             201                                             250
PrV UL6      IvClGvVPVr Rar.pggkrr rercvde.h. ....Dlaghl rVAgsvLgqg
HSV-1 UL6    ltClGLVPIl RKr.qeg..g vtqgLra.Fl kqhpltrqLa tVAeaa.era
VZV gene 54  IaCvGiVPVv RKv.rseqng tpapLnt.Ym gqaaElsqMl kVAdatLarg
EBV BBRF1    IcatGMVPav kKpitqelhs kiksLrdrCv crelgherti rsigteLyea
HCMV UL104   cvtvGLVPll dvktkpseaa eraqFvraav qratEthpL. ..AqdlLqan
Consensus    I-C-GLVPV- RK-------- ----L----- ----E---L- -VA---L---
             251                                             300
PrV UL6      ldevaELAEa MrgVtImDYD RVqlYYepRh rrvLArDalt GeRGECLVLW
HSV-1 UL6    gpgffELAla FdstrVaDYD RVyIYYnhRr gdwLvrDpis GqRGECLVLW
VZV gene 54  aavvtsLvEc MqnVaImDYD RtrlYYnynr rliMAkDdvt GmkGECLVvW
EBV BBRF1    tk...EiiEs LnstfIpqFt eVtIeYlpRs deyvA..yyc GrRirlhVLF
HCMV UL104   lalllqvAEr LgaVrVanap eVrVFkkvRs erleAq..lr GkhirlyVaa
Consensus    -----ELAE- ---V-I-DYD RV-IYY--R- ---LA-D--- G-RGECLVLW
             301                                             350
PrV UL6      qPLwrd.geL lFDSPaQRvh gEVLAChaLR EHArlCQLLN TAPVKVLVGR
HSV-1 UL6    PPLwtg.DrL vFDSPVQRLf pEIvAChsLR EHAhVCrLrN TAsVKVLlGR
VZV gene 54  PPvvCg.Egv vFDSPlQRLs gEVLACyaLR EHArVCQvLN TAPlrVLIGR
EBV BBRF1    PPaiFa.gtv tFDSPVQRLy qnIFmCyrtl EHAkICQLLN TAPlKaiVG.
HCMV UL104   ePLaYerDkL lFttPVahLh eEILrYdgLc rHqkICQLLN TfPVKVvtas
Consensus    PPL----D-L -FDSPVQRL- -EILAC--LR EHA--CQLLN TAPVKVLVGR
             351                                             400
PrV UL6      kpaeg..Hpa gA..VEkmLG ED...pAGSA AaRLVrLIVN MKgMRHIGDI
HSV-1 UL6    ksDsergvag aArvVnkVLG EDdetkAGSA AsRLVrLIIN MKgMRHVGDI
VZV gene 54  rnEddrsHst rA..VDrIMG EndttrAGSA AsRLVkLIVN LKnMRHVGDI
EBV BBRF1    .......Hgg rdmykDilah lEqnsqrkdp kkeLlnLlVk LsenktIsgV
HCMV UL104   rhEln..... ...ckklVem mEqhdrgsdA kksimkFllN vsdsksrigI
Consensus    --E----H-- -A--VD-VLG E-----AGSA A-RLV-LIVN -K-MRH-GDI
```

Fig. 4 (*continued*)

```
               401                                                 450
PrV UL6        TETVRSYLDE tGarilD... sVDtsqPG.. ..FGhhG... .....ag...
HSV-1 UL6      nDTVRSYLDE aGghLIDap. aVDgtlPG.. ..FGkgGnsr gsagqdqggr
VZV gene 54    TETVRSYLEE tGnhilEgsg sVDtsqPG.. ..FGkanqsf nggamsg...
EBV BBRF1      TDvVeeFitD asnnLVD... ....rnrl.. ..FGqpGet. ..........
HCMV UL104     eDsVeSFLqD ltpsLVDqnr llpargPGgp gvvGpgGavv ggpaghvgll
Consensus      TDTVRSYLDE -G--LVD--- -VD---PG-- --FG--G--- ----------
               451                                                 500
PrV UL6        aqpvqdA... ..FRTsVVNs INGmLEGYVN NLFkTIEsLk adNggLreqV
HSV-1 UL6      apqlrqA... ..FRTaVVNn INGvLEGYIN NLFgTIErLr etNagLatql
VZV gene 54    ttnvqsA... ..FkTsVVNs INGmLEGYVN NLFkTIEgLk dvNsdLterl
EBV BBRF1      ...aaqg... ..LkkkVsNt VvkcLtdqIN eqFdqIngLe kerelylkkI
HCMV UL104     ppppgpAape rdiRdlfkkq VikcLEeqIq sqvdeIqdLr tlNqtwenrV
Consensus      ------A--- --FRT-VVN- ING-LEGYIN NLF-TIE-L- --N--L---V
               501                                                 550
PrV UL6        RErEqEvrRl reqalraaqa gAdgtgaagp aggrapgaAd gqpprDLghE
HSV-1 UL6      qErDrELrRa tagalerqqr aAdlAaesvt ggcgsrpaga DllraD..yD
VZV gene 54    qfkEgELkRl reervkik.. .........p skgshitmAe EtriaDLnhE
EBV BBRF1      RsmEsqLqas lgpggnnpaa sApaAvaaea asvdiltgst asaiekLfns
HCMV UL104     RElrdlLtRy asrredsmsl gArdAelyhl pvleavrkAr Daap......
Consensus      RE-E-EL-R- ---------- -A--A----- --------A- D----DL--E
               551                                                 600
PrV UL6        VID..itram GDDaYVANSF QSrYVEPYes DveRISrLWE qELIRCFKLt
HSV-1 UL6      IID..vsksm dDDtYVANSF QhpYIPsYaq DleRISrLWE hELVRCFKil
VZV gene 54    VID..ltgii GDDaYIANSF QSrYIPPYgd DikRISeLWk qELVRCFKLh
EBV BBRF1      psaslgarvs GhnesIlNSF vSqYIPPsre mtkdLteLWE sELfntFKLt
HCMV UL104     ....frplav eDnrlVANSF fSqFVPgtes lerfLtqLWE nEyfRtFrLr
Consensus      VID------- GDD-YVANSF QS-YIPPY-- D--RIS-LWE -ELVRCFKL-
               601                                                 650
PrV UL6        RVaNNQGQEv SVSYSNssIS lilAPYFFsI LRvrhLGfLI TsqEVyrSEE
HSV-1 UL6      chrNNQGQEt SISYSsgaIa aFvAPYFesV LRaprvGapI TgsDViLgEE
VZV gene 54    RVnNNQGQEi SVSYSNasIS lLvAPYFsfI LRatrLGfLV TqsEVhrSEE
EBV BBRF1      pVvdNQGQrl yVrYSsdtIS iLlgPFtYlV aelspve.LV TdvyatLgiv
HCMV UL104     RlvthQGaEe aIvYSNytVe rvtlPYLChI LalgtLd.pV peaylqLSfg
Consensus      RV-NNQGQE- SVSYSN--IS -L-APYF--I LR---LG-LV T--EV-LSEE
               651                                                 700
PrV UL6        DLcgvVFKKt RLeaYLtEIa aLF...aAdV R.....rAte rlrDggR...
HSV-1 UL6      ELwdAVFKKt RLqtYLtDIa aLF...vAdV qhaalpppPs pvgadfRpga
VZV gene 54    ELcqAIFKKa RtesYLsqIr iLY...emqV RaevikrgPr rtpspsWglp
EBV BBRF1      EiidelYrss RLaiYieDlg rkYcpasAtg gdhgirqAPs argDtepdha
HCMV UL104     EivaAaYdds kFcrYve... .Licsrekar RrqmsreAag gvpErgtass
Consensus      EL--AVFKK- RL--YL-DI- -L-----A-V R------AP- ---D--R---
               701                                                 750
PrV UL6        dgradlaRpP aaG....... .......... ........vR gargaradpg
HSV-1 UL6      sprgr.sRsr spGrta.... .......... .........R gapdqgggig
VZV gene 54    dpteddERiP epnkinnqym hvgyknlshf mkghpperlR vhkvnaadst
EBV BBRF1      kskparDppP gaGs...... .......... .......... ..........
HCMV UL104     ggpgtlERsa prrlitadee rrgpervgrf rnggpddprR aggpygfh..
Consensus      ------ER-P --G------- ---------- ---------R ----------
               751                                                 800
PrV UL6        lhErlgd... .......... .......... .......... ..........
HSV-1 UL6      hrDgrrdgrr .......... .......... .......... ..........
VZV gene 54    llDkiranrr rgdgrwdvrn kytqhfrlqr ndrqltntsr rgvgcerrdr
EBV BBRF1      .......... .......... .......... .......... ..........
HCMV UL104     .......... .......... .......... .......... ..........
Consensus      --D------- ---------- ---------- ---------- ----------
```

Fig. 4 (*continued*)

```
                801
PrV UL6         ..
HSV-1 UL6       ..
VZV gene 54     rs
EBV BBRF1       ..
HCMV UL104      ..
Consensus       --
```

Fig. 4. Multiple sequence analysis of UL6 homologous proteins. Sequence alignment of deduced PrV UL6, VZV gene 54 [15], HSV-1 UL6 [36], EBV BBRF1 [1] and HCMV UL104 [12] proteins was performed by the program *Pileup* and a consensus sequence was created with the program *Pretty*. Conserved amino acid residues in at least three of five sequences compared are shown in upper case letters. Amino acids conserved in all sequences compared are shown in grey shaded boxes. Regions of high homology are boxed

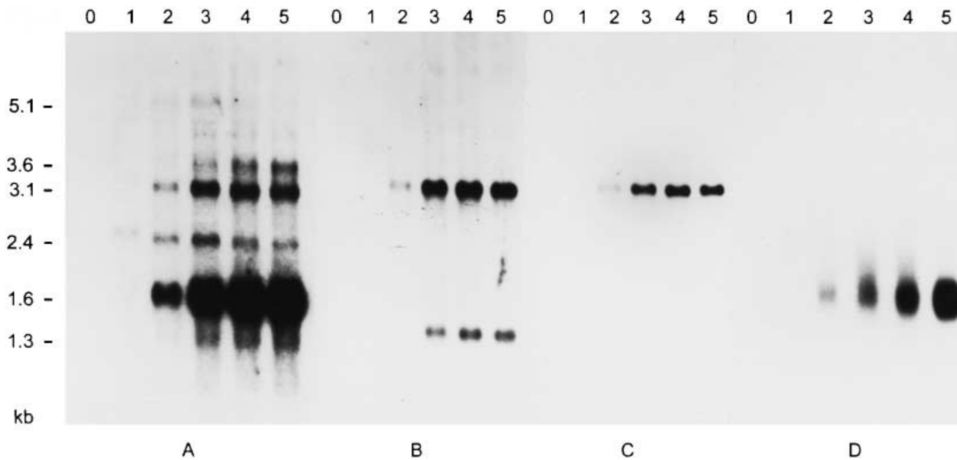

Fig. 5. Mapping of transcripts in the UL6 to UL10 region by Northern Blot analysis. Whole cell RNA from PrV infected pig kidney cells was isolated 1, 2, 3, 4, and 5 h p.i. and analyzed in Northern blot hybridizations with probes 1 (**A**), 2 (**B**), 3 (**C**), and 4 (**D**). For location of hybridization probes see Fig. 1. Lanes 0 contain RNA from mock-infected cells. Sizes of transcripts are indicated

Discussion

In this study an 11 059bp fragment of the PrV genome located in the U_L region has been sequenced and analyzed. This result represents another important step toward our goal to gain complete sequence information of the PrV genome. Within this region eight open reading frames were identified which were named PrV UL6, UL7, UL8, UL8.5, UL9, UL10, UL11 and UL12 based on their homology to respective HSV-1 genes [5, 36]. In addition, we analyzed transcription from the open reading frames UL6, UL7, UL8, UL8.5, UL9 and UL10.

The PrV UL12 gene has partially been sequenced previously [19]. These authors also analyzed transcription in this region and showed the presence of 3′-coterminal transcripts for UL14, UL13, UL12 and UL11 [20]. HSV-1 UL12

encodes an alkaline exonuclease which appears to be involved in efficient egress of capsids from the nucleus [50]. However, its precise role at the molecular level is still unclear, although a function in resolution of DNA recombination intermediates has been discussed recently [35]. Since amino acid sequences of the UL12 homologous proteins are well conserved, it is likely that the PrV UL12 gene product executes a similar function. Inactivation of PrV UL12 resulted in a strong reduction of viral virulence for mice [20], which indicates an important role in vivo.

The predicted PrV UL11 protein comprises only 63 aa with a calculated molecular mass of 7kDa. The HSV-1 UL11 protein of 96 aa is associated with cytoplasmic and nuclear membranes [4], and facilitates nucleocapsid envelopment and egress from cells [2]. It is modified by myristic acid moieties [31]. Although the UL11 homologs exhibit only limited sequence homology, and sizes of the deduced protein products differ, the glycine residue (aa 2 in the PrV protein) which is predicted to carry the myristic acid is conserved throughout all subfamilies of herpesviruses. However, the precise function of this modification is not clear at present.

With the elucidation of the sequences for the UL9 and UL8 genes of PrV, all homologs to genes whose products are known to be required for origin-dependent DNA replication in HSV-1 [40] have now also been identified in PrV [6, 8, 16, 43]. Generally, proteins involved in DNA replication are well conserved. This is also true for PrV UL9 and UL8 with 50% and 37% amino acid identity to corresponding HSV-1 proteins, respectively. In addition, the deduced PrV UL9 product contains six motifs indicative for the helicase superfamily 2 and a leucine zipper which may play a role in dimerization (Fig. 3) [32].

Two additional transcripts have been described in HSV-1 representing ORFs UL9.5 and UL8.5 [5]. While no UL9.5 protein product has been described yet, the UL8.5 polypeptide consists of the carboxy terminal 486 aa of the origin binding protein, which is expressed as an independent protein during viral infection. Sequence analysis shows presence of a corresponding ORF in PrV, with three possible start methionines. The first initiation codon is predicted as start codon for HSV-1 UL8.5 [5], while for PrV the third is the more probable according to the rules of Kozak [30]. Northern blot analysis revealed a 3.6kb transcript which could represent the PrV UL8.5 mRNA. It is detectable from 3 h p.i. until 5 h p.i. indicating delayed-early or late expression. As regards a possible UL9.5 gene, an ORF in opposite orientation of UL10, and overlapping UL10 and the 5′-terminal part of UL9 starts at nt 2 197 and ends at nt 3 865. However, no putative transcriptional control elements were detected upstream of this ORF, which is preceded by the repeat cluster, and no third position G/C bias was predicted by the program *Codonpreference*. In addition, we were unable to detect a possible UL9.5 transcript in PrV, although RNAs of very low abundance might have escaped detection. Therefore, the presence of an UL9.5 homolog in PrV appears unlikely.

No function of any UL7 homologous protein is known at present. In BHV-1, UL7 is not essential for viral replication in cell culture and the protein product

was detected in infected cells but not in purified virions [48]. PrV and BHV-1 UL7 proteins share 50% identical amino acids. This high conservation probably reflects common important functions which might become clearer after infection of the respective natural host.

HSV-1 UL6 codes for a capsid-associated protein [41]. Since amino acid identity of HSV-1 and PrV UL6 proteins amounts to 50%, it is reasonable to assume that the PrV protein also represents a capsid constituent. Transcripts for PrV UL6 and UL7 are 3′-coterminal sharing a common polyadenylation site which parallels the situation in HSV-1 and BHV-1 [41, 48]. No functional analysis of PrV UL6 or UL7 has been performed so far.

In summary, we established the sequence of an 11 059bp region of the PrV genome and showed conservation of gene arrangement and transcriptional pattern between PrV and other alphaherpesviruses.

Acknowledgement

Part of this work was supported by the European Union, Grant No. ERB-CHRX-CT92-0029.

References

1. Baer R, Bankier AT, Biggin MD, Deininger PL, Farrell PJ, Gibson TJ, Hatfull T, Hudson GS, Satchwell SC, Seguin C, Tuffnell PS, Barrell BG (1984) DNA sequence and expression of the B95–8 Epstein-Barr virus genome. Nature 310: 207–211
2. Baines JD, Roizman B (1992) The UL11 gene of herpes simplex virus 1 encodes a function that facilitates nucleocapsid envelopment and egress from cells. J Virol 66: 5 168–5 174
3. Baines JD, Roizman B (1993) The UL10 gene of herpes simplex virus 1 encodes a novel viral glycoprotein, gM, which is present in the virion and in the plasma membrane of infected cells. J Virol 67: 1 441–1 452
4. Baines JD, Jacob RJ, Simmerman L, Roizman B (1995) The herpes simplex virus 1 UL11 proteins are associated with cytoplasmic and nuclear membranes and with nuclear bodies of infected cells. J Virol 69: 825–833
5. Baradaran K, Dabrowski CE, Schaffer PA (1994) Transcriptional analysis of the region of the herpes simplex virus type 1 genome containing UL8, UL9, and UL10 genes and identification of a novel delayed-early gene product, OBPC. J Virol 68: 4 251–4 261
6. Baumeister J, Klupp BG, Mettenleiter TC (1995) Pseudorabies virus and equine herpesvirus 1 share a nonessential gene which is absent in other herpesviruses and located adjacent to a highly conserved gene cluster. J Virol 69: 5 560–5 567
7. Ben-Porat T, Veach RA, Ihara S (1983) Localization of the regions of homology between the genomes of herpes simplex virus type 1 and pseudorabies virus. Virology 127: 194–204
8. Berthommé H, Monahan S, Parris D, Jacquemont B, Epstein A (1995) Cloning, sequencing and functional characterization of the two subunits of the pseudorabies virus DNA polymerase holoenzyme: evidence for specificity of interaction. J Virol 69: 2 811–2 818
9. Birnstiel LM, Busslinger M, Strub K (1985) Transcription termination and 3'-processing: the end is in site! Cell 41: 349–359

10. Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu Rev Biochem 50: 349–383
11. Bruckner RG, Crute JJ, Dodson MS, Lehman IR (1991) The herpes simplex virus type 1 origin binding protein: a DNA helicase. J Biol Chem 266: 2 669–2 674
12. Chee A, Bankier T, Beck S, Bohni R, Brown C, Cerny R, Horsnell T, Hutchinson C, Kouzarides T, Martignetti J, Preddie E, Satchwell S, Tomlinson P, Weston K, Barrel B (1990) Analysis of the protein content of human cytomegalovirus strain AD169. Curr Top Microbiol Immunol 154: 125–185
13. Cordon JB, Wasylyk B, Buckwalder A, Sassome-Corsi P, Kedinger C, Chambon P (1980) Promotor sequences of eucaryotic protein coding genes. Science 209: 1 406–1 414
14. Crute JJ, Tsurumi T, Zhu L, Weller SK, Olivo PD, Challberg MD, Mocarski ES, Lehman IR (1989) Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products. Proc Natl Acad Sci USA 86: 2 186–2 189
15. Davison AJ, Scott JE (1986) The complete DNA sequence of varicella zoster virus. J Gen Virol 67: 1 759–1 816
16. Dean HJ, Cheung AK (1994) Identification of the pseudorabies UL4 and UL5 (helicase) genes. Virology 202: 962–967
17. Deb S, Deb SP (1991) A 269-amino acid segment with a pseudo leucine zipper and a helix-turn-helix motif codes for a sequence specific DNA-binding domain of herpes simplex virus type 1 origin binding protein. J Virol 65: 2 829–2 838
18. Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis for the VAX. Nucleic Acids Res 12: 387–395
19. DeWind N, Domen J, Berns A (1992) Herpesviruses encode an unusual protein-serine/threonine kinase which is nonessential for growth in cultured cells. J Virol 66: 5 200–5 209
20. DeWind N, Peeters BP, Zijderveld A, Gielkens ALJ, Berns A, Kimman TG (1994) Mutagenesis and characterization of a 41-kilobase-pair region of the pseudorabies virus genome: transcript map, search for virulence genes, and comparison with homologs of herpes simplex virus type 1. Virology 200: 784–790
21. Dijkstra JM, Visser N, Mettenleiter TC, Klupp BG (1996) Identification and characterization of pseudorabies virus glycoprotein gM as a nonessential virion component. J Virol 70: 5 684–5 688
22. Fierer DS, Challberg MD (1992) Purification and characterization of UL9, the herpes simplex virus type 1 origin-binding protein. J Virol 66: 3 986–3 995
23. Fitzgerald M, Shenk T (1981) The sequence 5′-AAUAAA-3′ forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell 24: 251–260
24. Hattori M, Sakaki Y (1986) Dideoxy sequencing method using denatured plasmid templates. Anal Biochem 153: 232–467
25. Henikoff G (1993) Unidirectional digestion with exonuclease III creates targeted break-points for DNA sequencing. Gene 28: 351–359
26. Inoue N, Dambaugh TR, Rapp JC, Pellett PE (1994) Alphaherpesvirus origin-binding protein homolog encoded by human herpesvirus 6B, a betaherpesvirus, binds to nucleotide sequences that are similar to *ori* regions of alphaherpesviruses. J Virol 68: 4 126–4 136
27. Kaplan AS, Vatter AE (1959) A comparison of herpes simplex and pseudorabies viruses. Virology 7: 394–407
28. Klupp BG, Mettenleiter TC (1991) Sequence and expression of the glycoprotein gH gene of pseudorabies virus. Virology 182: 732–741
29. Klupp BG, Kern H, Mettenleiter TC (1992) The virulence-determining genomic BamHI fragment 4 of pseudorabies virus contains genes corresponding to the U_L15 (partial),

U_L18, U_L19, U_L20, and U_L21 genes of herpes simplex virus and a putative origin of replication. Virology 191: 900–908
30. Kozak M (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44: 283–292
31. MacLean CA, Clark B, McGeoch DJ (1989) Gene UL11 of herpes simplex virus type 1 encodes a virion protein which is myristylated. J Gen Virol 70: 3 147–3 157
32. Martin DW, Munoz RM, Oliver D, Subler MA, Deb S (1994) Analysis of the DNA-binding domain of the HSV-1 origin-binding protein. Virology 198: 71–80
33. Martin DW, Deb S (1994) Cloning and expression of an equine herpesvirus 1 origin-binding protein. J Virol 68: 3 674–3 681
34. Martinez R, Shao L, Weller SK (1992) The conserved helicase motifs of the herpes simplex virus type 1 origin-binding protein UL9 are important for function. J Virol 66: 6735–6746
35. Martinez R, Sarisky RT, Weber PC, Weller SK (1996) Herpes simplex virus type 1 alkaline exonuclease is required for efficient processing of viral DNA replication intermediates. J Virol 70: 2 075–2 085
36. McGeoch DJ, Dalrymple MA, Davison AJ, Dolan A, Frame MC, McNab D, Perry LJ, Scott JE, Taylor P (1988) The complete DNA sequence of the unique long region in the genome of herpes simplex virus type 1. J Gen Virol 69: 1 531–1 574
37. McLean GW, Abbotts AP, Parry ME, Marsden HS, Stow ND (1994) The herpes simplex virus type 1 origin-binding protein interacts specifically with the viral UL8 protein. J Gen Virol 75: 2 699–2 706
38. Mettenleiter TC (1993) Pseudorabies (Aujeszky's disease) virus: State of the art. Acta Vet Hung 42: 153–177
39. Mettenleiter TC (1996) Unpublished
40. Olivo PD, Nelson NJ, Challberg MD (1989) Herpes simplex virus type 1 gene products required for DNA replication: identification and overexpression. J Virol 63: 196–204
41. Patel AH, MacLean J (1995) The product of the UL6 gene of herpes simplex virus type 1 is associated with virus capsids. Virology 206: 465–478
42. Patel AH, Rixon FJ, Cunningham C, Davison AJ (1996) Isolation and characterization of herpes simplex virus type 1 mutants defective in the UL6 gene. Virology 217: 111–123
43. Pederson N, Enquist LW (1989) The nucleotide sequence of a pseudorabies virus gene similar to ICP18.5 of herpes simplex virus type 1. Nucleic Acids Res 17: 3 597
44. Riebe R (1996) Unpublished
45. Robertson GR, Scott NA, Miller JM, Sabine M, Zheng M, Bell CW, Whalley JM (1991) Sequence characteristics of a gene in equine herpesvirus 1 homologous to glycoprotein H of herpes simplex virus. DNA Sequence 1: 241–249
46. Roizman B, Derosiers R, Fleckenstein B, Lopez C, Minson AC, Studdert MJ (1992) The family Herpesviridae: an update. Arch Virol 123: 425–449
47. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5 463–5 467
48. Schmitt J, Keil GM (1996) Identification and characterization of bovine herpesvirus 1 UL7 gene and gene product which are not essential for virus replication in cell culture. J Virol 70: 1 091–1 099
49. Schwyzer M (1995) Sequence analysis of bovine herpesvirus 1 genome: An exercise in international cooperation. In: Immunobiology of viral infections, Proc. 3rd Congress Europ. Soc. Vet. Virol., Interlaken
50. Shao L, Rapp LM, Weller SK (1993) Herpes simplex virus 1 alkaline nuclease is required for efficient egress of capsids from the nucleus. Virology 196: 146–162

51. Sherman G, Gottlieb J, Challberg MD (1992) The UL8 subunit of the herpes simplex virus helicase-primase complex is required for efficient primer utilization. J Virol 66: 4 884–4 892
52. Telford EA, Watson MS, McBride K, Davison AJ (1992) The DNA sequence of equine herpesvirus-1. Virology 189: 304–316
53. Vlcek C, Benes V, Lu Z, Kutish G, Paces V, Rock D, Letchworth G, Schwyzer M (1995) Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome which exhibits a colinear gene arrangement with the UL21 to UL4 genes of herpes simplex virus. Virology 210: 100–108
54. Wittmann G, Rziha HJ (1989) Aujeszky's disease (pseudorabies) in pigs. In: Wittmann G (ed) Herpes diseases of cattle, horses and pigs. Kluwer, Boston, pp 230–325

Authors' address: Dr. B. G. Klupp, Federal Research Centre for Virus Diseases of Animals, D-17498 Insel Riems, Federal Republic of Germany.

Received June 28, 1996