GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

# Molecular characterization of the pseudorabies virus *UL2* gene

**M.S. Cai[1,2], B.Y. Wang[2], W. Cui[1], Z.Y. Zhao[1], J.H. Chen[2], X.M. Wen[1], Z. Li[1] and M.L. Li[1]**

[1]Department of Pathogenic Biology and Immunology, School of Basic Science, Guangzhou Medical University, Guangzhou, Guangdong, China
[2]Department of Veterinary Medicine, School of Life Science, Foshan Science and Technology University, Foshan, Guangdong, China

Corresponding author: M.L. Li
E-mail: meili_2011@hotmail.com

Genet. Mol. Res. 12 (4): 4147-4161 (2013)
Received September 13, 2012
Accepted February 20, 2013
Published October 7, 2013
DOI http://dx.doi.org/10.4238/2013.October.7.1

**ABSTRACT.** A 948-bp sequence of the *UL2* gene was amplified from the pseudorabies virus (PRV) Becker strain genome using polymerase chain reaction, and the gene identity was confirmed through further cloning and sequencing. Bioinformatic analysis indicated that the PRV *UL2* gene encodes a putative polypeptide with 315-amino acid residues. Its encoding protein, designated UL2, has a conserved uracil-DNA glycosylase (UDG)_F1 domain, which is closely related to the herpesvirus UDG family and is highly conserved among its counterparts encoded by UDG genes. Multiple nucleic acid and amino acid sequence alignments suggested that the product of PRV *UL2* has a relatively higher homology with UL2-like proteins of Alphaherpesvirinae than that of other subfamilies of Herpesviridae. In addition, phylogenetic analysis showed that PRV UL2 had a close evolutionary relationship with members of Alphaherpesvirinae, especially members of the genus *Varicellovirus* of bovine herpesvirus 1 and bovine herpesvirus 5. Antigen prediction indicated the presence of several potential B-cell epitopes in PRV UL2. In addition, secondary structure and 3-dimensional structure prediction revealed that PRV UL2 consisted predominantly of

 

an α-helix. Taken together, these results provide molecular biological insight for the further study of the function and mechanism of *UL2* during PRV infection.

**Key words:** Pseudorabies virus; Cloning; Bioinformatic analysis; *UL2*; UDG; Molecular characterization

## INTRODUCTION

Pseudorabies virus (PRV) is the agent of Aujeszky's disease, a frequently fatal disease that has a global distribution and primarily affects swine and incidentally affects other domestic and wild animals. Owing to its neurotropic nature, PRV has been used as a tool to trace circuits in the neuronal system. In addition, PRV has served as a useful model organism for the study of herpesvirus pathogenesis. PRV causes considerable economic losses in the pig industry worldwide, and although efforts to eradicate it have shown great progress, it is still an endemic problem in many countries (Kramer et al., 2011).

Comprehensive investigation of the function of each PRV gene during viral replication is particularly important for the elucidation of the fundamental mechanisms underlying the spread and pathogenesis of PRV. PRV UL2, a *UL2*-encoded nonstructural protein predicted to be a uracil-DNA glycosylase (UDG), is less well understood thus far. However, the homologs herpes simplex virus 1 (HSV-1) UL2 (Mullaney et al., 1989), HSV-2 UL2 (Winters and Williams, 1993), varicella zoster virus ORF59 (Reddy et al., 1998), and Epstein-Barr virus BKRF3 (Lu et al., 2007) have been extensively studied. UDG is reported to be involved in the DNA excision repair pathway that specifically removes an inappropriate uracil from DNA (Prichard et al., 2005; Bogani et al., 2009). UDG is also associated with the viral replisome via interaction with DNA polymerase (Prichard et al., 2005; Bogani et al., 2009). Moreover, UDG may be indispensable for viral replication in culture cells, because the UL2 mutant virus exhibits reduced neurovirulence, decreased frequency of reactivation from latency (Pyles and Thompson, 1994) and impairment of efficient viral gene expression and DNA synthesis as well as efficient production of virus *in vivo* (Lu et al., 2007; Ward et al., 2009; Strang and Coen, 2010).

In this study, *UL2* was amplified from the PRV Becker strain genome using polymerase chain reaction (PCR) followed by cloning and sequencing. Subsequently, a comprehensive bioinformatic analysis was carried out to determine the molecular characteristics of *UL2* and provide molecular biological insight for further study of the function and mechanism of *UL2* during PRV infection. The study used several bioinformatic tools, including open reading frame (ORF) Finder, Conserved Domains, DNAstar 7.0, Bioedit 7.0, SignalP-4.0, NetPhos 2.0, PSIpred, and CPHmodels 3.2.

## MATERIAL AND METHODS

### Cloning of PRV *UL2*

PCR amplification primers for *UL2* (accession No. U02512) were designed using Oligo 6.0 and Primer 5.0 and were synthesized by TaKaRa (Dalian, China). The upstream primer

 

(synthesized by Sangon Biotech, Shanghai, China) 5'-CGGAATTCATGGAGGGCCCCCCG CCGAGT-3' annealed with the first 21 nucleotides of *UL2* and was introduced with *Eco*RI (underlined) to facilitate subsequent cloning. The downstream primer 5'-TCCTCGAGTCAGTC CACGCTCCAGTCGACGG-3' was complementary to the final 23 nucleotides of *UL2* and was introduced with *Xho*I (underlined).

*UL2* was amplified via PCR by KOD-Plus-Neo (TOYOBO), from the genomic DNA of the PRV Becker strain using a genome previously purified from vBecker2-infected PK-15 cells (Smith and Enquist, 2000; Li et al., 2011a,b) as the template. The purified PCR product was digested with *Eco*RI and *Xho*I and ligated into the correspondingly digested prokaryotic expression vector pET28a(+) (Novagen, Madison, WI, USA) to generate pET28a(+)-UL2. The presence of the appropriate insert in the obtained plasmid was then verified using PCR, restriction analysis, and sequencing.

### Bioinformatic analysis of the nucleotide sequence of PRV *UL2*

To determine the nucleotide sequence similarity and identify the ORF, we applied the National Center for Biotechnology Information (NCBI) nucleotide Basic Local Alignment Search Tool [BLAST; http://www.ncbi.nlm.nih.gov/BLAST/ (accessed November 23, 2012)] and ORF Finder [http://www.ncbi.nlm.nih.gov/gorf/gorf.html (accessed November 23, 2012)], respectively. Subsequently, Clustal V in the MegAlign program of DNAStar (version 7.0, DNAStar, Inc.) was used to conduct the nucleotide sequence homology analysis of 61 UL2-like proteins (Table 1).

### Bioinformatic analysis of the deduced amino acid (aa) sequence of PRV *UL2*

For aa sequence comparison, homology search, and conserved domain analysis, the aa sequence of UL2 was analyzed using protein BLAST and the Conserved Domains search tool [http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi (accessed November 23, 2012)], respectively. To compare UL2 with UL2-like proteins of other species (see Table 1), we analyzed the aa sequence homology and phylogenetic relationships using DNAstar 7.0. To predict the signal peptide sequence, transmembrane domain, glycosylation site, phosphorylation site, hydrophobic and hydrophilic regions, B-cell epitope, secondary structure, and 3-dimensional (3-D) structure of UL2 we used SignalP-4.0 Server [http://www.cbs.dtu.dk/services/SignalP/ (accessed November 23, 2012)], TMHMM [http://www.cbs.dtu.dk/services/TMHMM/ (accessed November 23, 2012)], NetNGlyc 1.0 [http://www.cbs.dtu.dk/services/NetNGlyc/ (accessed November 23, 2012)], NetPhos 2.0 [http://www.cbs.dtu.dk/services/NetPhos/ (accessed November 23, 2012)], Bioedit7.0 software, DNAstar7.0 software, PSIpred [http://bioinf.cs.ucl.ac.uk/psipred/ (accessed November 23, 2012)], and CPHmodels 3.2 [http://www.cbs.dtu.dk/services/CPHmodels/ (accessed November 23, 2012)], respectively.

## RESULTS

### PCR amplification and cloning of PRV *UL2*

To obtain the *UL2* gene, we performed PCR based on the DNA template from the purified genome of the PRV Becker strain. As shown in Figure 1, a target fragment of 948 bp, which is consistent with the expected size, was amplified from DNA purified from PRV-

**Table 1.** Abbreviations and accession Nos. of 61 UL2-like proteins from different species.

| Family/subfamily | Genus | Virus/bacterium name (Abbreviation) | Strain | Natural host | GenBank accession No. |
|---|---|---|---|---|---|
| Alphaherpesvirinae | *Varicellovirus* | Suid herpesvirus 1 (SuHV-1) | Becker | *Sus scrofa* (Pig) | AAA18856 |
| | | Pseudorabies virus (PRV) | Bartha | | AEM63999 |
| | | | Kaplan | | AFI70835 |
| | | | Indiana-Funkhauser | | AAA16422 |
| | | Equid herpesvirus 1 (EHV-1) | Ab4 | *Equus caballus* (Horse) | YP_053105 |
| | | Equine abortion virus (EAV) | | | |
| | | Equine herpesvirus 4 (EHV-4) | NS80567 | | NP_045278 |
| | | Equine rhinopneumonitis virus (ERV) | | | |
| | | Equid herpesvirus 8 (EHV-8) | wh | | YP_006273041 |
| | | Equid herpesvirus 9 (EHV-9) | P19 | | YP_002333542 |
| | | Gazelle herpesvirus 1 (GHV-1) | | | |
| | | Cercopithecine herpesvirus 9 (CeHV-9) | Delta | *Erythrocebus patas* (Monkey) | NP_077473 |
| | | Simian varicella virus (SVV) | | | |
| | | Felid herpesvirus 1 (FeHV-1) | C-27 | Felidae (Cat) | YP_003331580 |
| | | Bovine herpesvirus 1 (BoHV-1) | Cooper | *Bos taurus* (Cattle) | AAC54556 |
| | | Infectious bovine rhinotracheitis virus (IBRV) | | | |
| | | Bovine herpesvirus 5 (BoHV-5) | SV507/99 | | YP_003662524 |
| | | Bovine encephalitis herpesvirus (BEHV) | | | |
| | | Human herpesvirus 3 (HHV-3) | Dumas | *Homo sapiens* (Human) | NP_040181 |
| | | Varicella-zoster virus (VZV) | | | |
| | *Simplexvirus* | Human herpesvirus 1 (HHV-1) | E03 | | ADM23318 |
| | | Herpes simplex virus 1 (HSV-1) | | | |
| | | Human herpesvirus 2 (HHV-2) | HG52 | | NP_044471 |
| | | Herpes simplex virus 2 (HSV-2) | | | |
| | | Cercopithecine herpesvirus 1 (CeHV-1) | E2490 | *Macaca mulatta* (Monkey) | AAP41420 |
| | | Macacine herpesvirus 1 (McHV-1) | | | |
| | | Monkey B virus | | | |
| | | Cercopithecine herpesvirus 2 (CeHV-2) | B264 | *Cercopithecus aethiops* (Monkey) | YP_164444 |
| | | Simian agent 8 (SA8) | | | |
| | | Cercopithecine herpesvirus 16 (CeHV-16) | X313 | *Papio cynocephalus* (Baboons) | ABA29255 |
| | | Papiine herpesvirus 2 (PaHV-1) | | | |
| | | Herpesvirus papio 2 (HVP-2) | | | |
| | | Saimiriine herpesvirus 1 (SaHV-1) | MV 5-4 | Saimiri (Squirrel monkeys) | YP_003933838 |
| | | Marmoset herpesvirus (MarHV) | | | |
| | *Iltovirus* | Psittacid herpesvirus 1 (PsHV-1) | 97-0001 | *Amazona oratrix* (Parrot) | NP_944438 |
| | | Pacheco's disease virus (PDV) | | | |
| | | Gallid herpesvirus 1 (GaHV-1) | SA-2 | White Leghorn (Chicken) | AAA64610 |
| | | Infectious laryngotracheitis virus (ILTV) | | | |
| | *Mardivirus* | Gallid herpesvirus 2 (GaHV-2) | Md5 | *Gallus domesticus* (Chicken) | YP_001033930 |
| | | Marek's disease virus type 1 (MDV-1) | | | |
| | | Gallid herpesvirus 3 (GaHV-3) | HPRS24 | *Gallus gallus* (Chicken) | NP_066831 |
| | | Marek's disease virus type 2 (MDV-2) | | | |

Continued on next page

**Table 1.** Continued.

| Family/subfamily | Genus | Virus/bacterium name (Abbreviation) | Strain | Natural host | GenBank accession No. |
|---|---|---|---|---|---|
| | | Meleagrid herpesvirus 1 (MeHV-1)<br>Turkey herpesvirus (HVT) | FC126 | *Meleagris gallopavo* (turkey) | NP_073295 |
| | | Anatid herpesvirus 1 (AnHV-1)<br>Duck enteritis virus (DEV) | 2085 | *Anatid* species (Duck) | AEN80123 |
| Betaherpesvirinae | *Cytomegalovirus* | Human herpesvirus 5 (HHV-5)<br>Human cytomegalovirus (HCMV) | Towne | *Homo sapiens* (Human) | ACS32409 |
| | | Saimiriine herpesvirus 3 (SaHV-3)<br>Squirrel monkey cytomegalovirus (SMCMV) | SqSHV | *Saimiri sciureus* (Squirrel monkey) | AEV80957 |
| | | Cercopithecine herpesvirus 5 (CeHV-5)<br>Geen monkey cytomegalovirus (GMCMV)<br>Simian cytomegalovirus (SCMV) | 2715 | *Cercopithecus aethiops* (Vervet monkey) | YP_004936073 |
| | *Muromegalovirus* | Murid herpesvirus 1 (MuHV-1)<br>Murine cytomegalovirus (MCMV) | Smith | *Mus domesticus* (House mice) | ADD10477 |
| | | Murid herpesvirus 2 (MuHV-2)<br>Rat cytomegalovirus (RCMV) | Maastricht | *Rattus norvegicus* (Norway rat) | NP_064216 |
| | *Roseolovirus* | Human herpesvirus 7 (HHV-7) | RK | *Homo sapiens* (Human) | YP_073819 |
| | | Human herpesvirus 6 (HHV-6) | HST | | BAA04967 |
| | | Caviid herpesvirus 2 (CavHV-2)<br>Guinea pig cytomegalovirus (GPCMV) | 21222 | *Cavia porcellus* (Guinea pig) | YP_002321292 |
| Gammaherpesvirinae | *Rhadinovirus* | Murid herpesvirus 4 (MuHV-4)<br>Murine gammaherpesvirus 68 (MuHV-68) | g2.4 (WUMS) | *Apodemus sylvaticus* (Wood mice) | NP_044883 |
| | | Ateline herpesvirus 3 (AtHV-3)<br>Herpesvirus ateles (HVA) | 73 | *Ateles* (Spider monkey) | AAC95571 |
| | | Saimiriine herpesvirus 2 (SaHV-2)<br>Herpesvirus saimiri (HVS) | 11 | *Saimiri sciureus* (Squirrel monkey) | NP_040248 |
| | | Bovine herpesvirus 4 (BoHV-4)<br>Human herpesvirus 8 (HHV-8) | | *Bos taurus* (Cattle) | NP_076538 |
| | | Kaposi's sarcoma-associated herpesvirus (KSHV) | GK18 | *Homo sapiens* (Human) | ABD28896 |
| | | Cercopithecine herpesvirus 17 (CeHV-17)<br>Macacine herpesvirus 5 (McHV-5) | 17577 | *Macaca mulatta* (Rhesus monkey) | NP_570787 |
| | | Macaca fuscata rhadinovirus (MFRV) | | *Macaca fuscata* (Japanese monkey) | AEW87772 |
| | *Macavirus* | Ovine herpesvirus 2 (OvHV-2) | BJ1035 | *Ovis aries* (Sheep) | YP_438168 |
| | | Alcelaphine herpesvirus 1 (AlHV-1)<br>Malignant catarrhal fever virus (MCFV) | C500 | *Connochaetes taurinus* (Wildebeest) | NP_065544 |
| | *Lymphocryptovirus* | Callitrichine herpesvirus 3 (CalHV-3)<br>Marmoset lymphocryptovirus (MLCV) | CJ0149 | *Callithrix jacchus* (Common marmoset) | NP_733890 |
| | | Human herpesvirus 4 (HHV-4)<br>Epstein-Barr virus (EBV) | B95-8 | *Homo sapiens* (Human) | CAD53429 |
| | | Cercopithecine herpesvirus 15 (CeHV-15)<br>Macacine herpesvirus 4 (McHV-4)<br>Rhesus lymphocryptovirus (LCV) | LCL8664 | *Macaca mulatta* (Rhesus monkey) | YP_067975 |
| | *Percavirus* | Equid herpesvirus 2 (EHV-2) | 86/67 | *Equus caballus* (Horse) | NP_042643 |
| | | Rodent herpesvirus Peru (RHVP) | | *Oligoryzomys microtis* (Pygmy rice rat) | YP_004207882 |

Continued on next page

**Table 1.** Continued.

| Family/subfamily | Genus | Virus/bacterium name (Abbreviation) | Strain | Natural host | GenBank accession No. |
|---|---|---|---|---|---|
| Alloherpesviridae | *Cyprinivirus* | Cyprinid herpesvirus 3 (CyHV-3) Koi herpesvirus (KHV) | TUMST1 | *Cyprinus carpio* (Common carp) | BAF48912 |
| Neisseriaceae | *Neisseria* | *Neisseria lactamica* | 020-06 | *Homo sapiens* (Human) | YP_004048620 |
| | | *Neisseria cinerea* | ATCC 14685 | | ZP_05983780 |
| | | *Neisseria meningitidis* | alpha710 | | YP_005882767 |
| | | *Neisseria gonorrhoeae* | FA 1090 | | YP_207912 |
| | | *Neisseria polysaccharea* | ATCC 43768 | | ZP_06864356 |
| | | *Neisseria mucosa* | ATCC 25996 | | ZP_05977860 |
| | *Kingella* | *Kingella kingae* | ATCC 23330 | | ZP_08466865 |
| Enterobacteriaceae | *Shigella* | *Shigella flexneri* | K-315 | | EIQ19264 |
| | *Sodalis* | *Sodalis glossinidius* | morsitans | *Glossina morsitans morsitans* (Tsetse fly) | YP_455477 |
| | *Escherichia* | *Escherichia fergusonii* | ATCC 35469 | *Gallus gallus domesticus* (Chicken) | YP_002381685 |
| | *Candidatus Moranella* | *Candidatus Moranella endobia* | PCIT | *Planococcus citri* (Mealybug) | YP_004706663 |
| | *Proteus* | *Proteus mirabilis* | H14320 | *Homo sapiens* (Human) | YP_002151626 |
| Pasteurellaceae | *Haemophilus* | *Haemophilus influenzae* | 10810 | | YP_005178054 |

infected PK-15 cells (see Figure 1, lane 1), whereas no specific band was amplified from the mock-infected control (see Figure 1, lane 2). The DNA fragment of *UL2* was then cloned into the prokaryotic expression vector pET28a(+) to yield pET28a(+)-UL2 (see Figure 1, lane 3), which was confirmed via restriction digestion analysis (see Figure 1, lanes 4 and 5), PCR amplification (see Figure 1, lane 6), and DNA sequencing. The sequencing result demonstrated that no aa mutation was present in the clone compared to the PRV Becker strain (accession No. U02512).

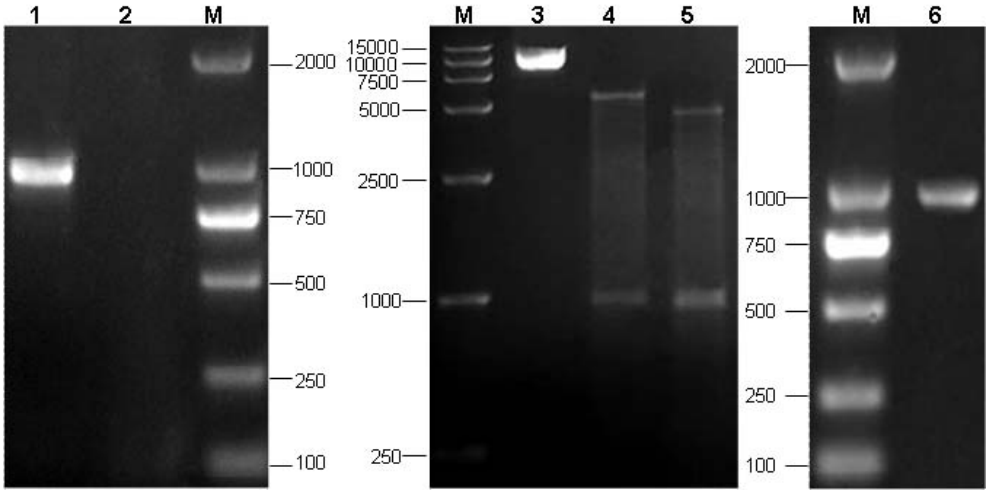


**Figure 1.** Polymerase chain reaction (PCR) amplification and restriction analysis of the recombinant plasmid pET28a(+)-UL2. *Lanes 1* and *2* = PCR amplification product of *UL2* using DNA purified from pseudorabies virus (PRV)- and mock-infected PK-15 cells as the template, respectively; *lane 3* = recombinant plasmid pET28a(+)-UL2; *lane 4* = restriction digestion product (approximately 5335 and 954 bp) of pET28a(+)-UL2 with *Eco*RI and *Xho*I; *lane 5* = restriction digestion product (approximately 4457, 925, and 907 bp) of pET28a(+)-UL2 with *Bam*HI and *Mlu*I; *lane 6* = PCR amplification product of *UL2* from pET28a(+)-UL2. Samples were electrophoresed using a 1% agarose gel and stained with ethidium bromide. The electrophoresis migration of molecular mass marker (M, TaKaRa) is also shown.

## Bioinformatic analysis of the PRV *UL2* nucleotide sequence

ORF Finder analysis revealed an integrated PRV *UL2* ORF consisting of 948 bp. In addition, nucleotide sequence similarity search using nucleotide BLAST (see Table 1) yielded 3 nucleotide sequences (accession Nos. JF797217, JQ809328, and L13855) with strong similarity to the PRV Becker strain *UL2* (up to 98.9, 99.1, and 99.4%, respectively; Table 2), and these accessions corresponded to *UL2* of the PRV Bartha, Kaplan, and Indiana-Funkhauser strains, respectively. Multiple alignment of PRV *UL2* with 58 homologous reference species demonstrated a remarkably high homology of 50.4 to 65.5% with members of the subfamily Alphaherpesvirinae, that is, HSV-1, HSV-2, cercopithecine herpesvirus 1 (CeHV-1), CeHV-2, bovine herpesvirus 1 (BoHV-1), and BoHV-5. However, low homology (less than 30%) was detected between PRV and other members of subfamilies Betaherpesvirinae, Gammaherpesvirinae, and Alloherpesviridae, that is, human herpesvirus 6, human herpesvirus 7, caviid herpesvirus 2, saimiriine herpesvirus 2, rodent herpesvirus Peru, ateline herpesvirus 3, murid herpesvirus 4, and cyprinid herpesvirus 3 (see Table 2).

**Table 2.** Multiple nucleic acid sequence and amino acid sequence alignments of the PRV *UL2* gene with its reference species.

| Virus name | AlHV-1 | AnHV-1 | AtHV-3 | BoHV-1 | BoHV-4 | BoHV-5 | CalHV-3 | *Candidatus Moranella endobia* | CavHV-2 | CeHV-1 |
|---|---|---|---|---|---|---|---|---|---|---|
| PNH[a] (%) | 33.7 | 41.3 | 26.6 | 65.5 | 31.4 | 61.6 | 35.7 | 37.2 | 25.0 | 60.5 |
| PAH[b] (%) | 36.9 | 49.8 | 33.3 | 65.2 | 34.4 | 57.4 | 35.5 | 43.2 | 18.7 | 59.3 |
| Virus name | CeHV-2 | CeHV-5 | CeHV-9 | CeHV-15 | CeHV-16 | CeHV-17 | CyHV-3 | EHV-1 | EHV-2 | EHV-4 |
| PNH[a] (%) | 52.7 | 39.2 | 36.7 | 37.5 | 58.8 | 36.8 | 24.5 | 48.9 | 36.1 | 43.2 |
| PAH[b] (%) | 50.8 | 35.1 | 45.3 | 36.1 | 60.1 | 36.9 | 15.2 | 51.3 | 36.5 | 48.1 |
| Virus name | EHV-8 | EHV-9 | *Escherichia fergusonii* | FeHV-1 | GaHV-1 | GaHV-2 | GaHV-3 | *Haemophilus influenzae* | HHV-1 | HHV-2 |
| PNH[a] (%) | 46.4 | 47.2 | 40.9 | 39.6 | 32.9 | 38.1 | 40.6 | 33.8 | 50.4 | 52.3 |
| PAH[b] (%) | 50.2 | 50.2 | 43.7 | 46.8 | 36.7 | 42.5 | 42.9 | 45.2 | 49.5 | 51.7 |
| Virus name | HHV-3 | HHV-4 | HHV-5 | HHV-6 | HHV-7 | HHV-8 | *Kingella kingae* | MeHV-1 | MFRV | MuHV-1 |
| PNH[a] (%) | 41.2 | 38.3 | 43.3 | 29.7 | 24.2 | 36.5 | 38.0 | 36.2 | 35.5 | 39.4 |
| PAH[b] (%) | 44.9 | 36.5 | 39.6 | 31.4 | 31.9 | 33.7 | 45.7 | 39.4 | 36.1 | 35.5 |
| Virus name | MuHV-2 | MuHV-4 | *Neisseria cinerea* | *Neisseria gonorrhoeae* | *Neisseria lactamica* | *Neisseria meningitidis* | *Neisseria mucosa* | *Neisseria polysaccharea* | OvHV-2 | *Proteus mirabilis* |
| PNH[a] (%) | 35.9 | 29.7 | 41.5 | 40.9 | 41.7 | 42.0 | 40.7 | 41.8 | 36.8 | 30.2 |
| PAH[b] (%) | 32.8 | 34.5 | 47.0 | 46.1 | 47.0 | 46.6 | 43.0 | 46.6 | 37.8 | 41.2 |
| Virus name | PsHV-1 | RHVP | SaHV-1 | SaHV-2 | SaHV-3 | *Shigella flexneri* | *Sodalis glossinidius* | SuHV-1 Bartha | SuHV-1 Kaplan | SuHV-1 Indiana-Funkhauser |
| PNH[a] (%) | 39.1 | 27.7 | 49.5 | 24.5 | 34.3 | 40.7 | 41.4 | 98.9 | 99.1 | 99.4 |
| PAH[b] (%) | 34.6 | 34.1 | 49.2 | 29.8 | 34.3 | 43.2 | 43.6 | 98.4 | 98.4 | 99.0 |

[a]PNH = multiple nucleic acid sequence alignment of *UL2* gene of PRV Becker strain with its homologous genes of 60 selected species (Table 1) by using the MEGALIGN program in LASERGENE (DNAStar 7.0) with Clustal V Method, and sequence distance was calculated using weight matrix Identity. Gaps had been introduced by the alignment program to maximize the homology. [b]PAH = multiple aa sequence alignment of UL2 of PRV Becker strain with its homologous proteins of 60 selected species (Table 1) by using the MEGALIGN program in LASERGENE (DNAStar 7.0) with Clustal V Method, and sequence distance was calculated using weight matrix PAM250. Gaps had been introduced by the alignment program to maximize the homology.

## Bioinformatic analysis of PRV UL2 polypeptide sequence

An aa sequence similarity search using protein BLAST (see Table 1) yielded 3 aa sequences (accession Nos. AEM63999, AFI70835, and AAA16422) that strongly matched the target sequence of PRV Becker UL2 (accession No. AAA18856; up to 98.4, 98.4, and 99.0%, respectively), and these corresponded to UL2 of the Bartha, Kaplan, and Indiana-Funkhauser strains, respectively (see Table 2). In addition, multiple alignments of UL2 with homologs in 57 other reference species showed relatively high homology of 50.2 to 65.2% between UL2 and its equine herpesvirus 1 (EHV-1), EHV-8, EHV-9, CeHV-2, HSV-2, BoHV-1, BoHV-5, CeHV-1, and CeHV-16 counterparts. However, UL2 shared no substantial homology with UL2-like proteins from cyprinid herpesvirus 3 or caviid herpesvirus 2, with values of only 15.2 and 18.7%, respectively (see Table 2).

Phylogenetic analyses of PRV and other species were performed based on the aa sequences of UL2 and the UL2-like proteins of 60 reference species. The proteins were preliminarily separated into families or subfamilies (Figure 2), that is, the families Neisseriaceae, Enterobacteriaceae, Pasteurellaceae, and the subfamilies Alphaherpesvirinae, Betaherpesvirinae, Gammaherpesvirinae, and Alloherpesviridae, which was consistent with the existing classification within the assigned Herpesviridae family. Furthermore, the PRV Becker, Bartha, Kaplan, and Indiana-Funkhauser strains were different from other species (see Figure 2). They clustered together and formed a separate branch and then clustered with members of subfamily Alphaherpesvirinae such as BoHV-1 and BoHV-5 of *Varicellovirus*; CeHV-1, CeHV-2, CeHV-16, HSV-1, HSV-2, and saimiriine herpesvirus of *Simplexvirus*; and other members of *Mardivirus*. Subsequently they clustered with other members of families Neisseriaceae, Enterobacteriaceae, Pasteurellaceae, and members of subfamilies of Betaherpesvirinae, Gammaherpesvirinae, and Alloherpesviridae. Therefore, PRV might have a closer evolutionary relationship to members of *Varicellovirus* of subfamily Alphaherpesvirinae than to members of other herpesvirus subfamilies or the Neisseriaceae, Enterobacteriaceae, and Pasteurellaceae families.

Signal polypeptide prediction indicated no signal polypeptide cleavage site (Figure 3A) and no potential transmembrane domain in UL2 (Figure 3B). However, the N-linked glycosylation site (Asn-X-Ser/Thr) prediction demonstrated a potential N-glycosylation site in UL2 (Figure 3C). Interestingly, 12 potential phosphorylation sites were found in UL2 (Figure 3D), including 3 serine, 8 threonine, and 1 tyrosine residues. In addition, hydrophobicity analysis revealed 8 hydrophobic regions located at aa positions 13-45, 65-76, 91-105, 113-123, 148-159, 171-204, 208-221, and 232-262 (Figure 4A). Compared with the hydrophobic region, the hydrophilic region was slightly smaller (Figure 4B). Analysis of a potential B-cell epitope determinant demonstrated several potential B-cell epitopes in UL2 situated in or adjacent to aa positions 1-18, 45-66, 72-89, 96-102, 106-115, 120-139, 144-152, 156-166, 172-183, 208-213, 221-232, 242-252, 262-275, 280-297, and 304-315 (Figure 4C).

Conserved domain analysis indicated that UL2 contained an apparent conserved domain of UDG_F1 (Figure 5A), which is a UDG-like superfamily. Secondary structure analysis (Figure 5B) suggested that UL2 consisted primarily of random coil (51.11%) and α-helix (40.64%), whereas a β-strand accounted for the least prevalent component (8.25%). 3-D structure prediction for UL2 revealed a known 3-D structure model with a relatively high homology with UDG, which was predominantly composed of 15 α-helices and 9 β-strands (Figure 5C).

 

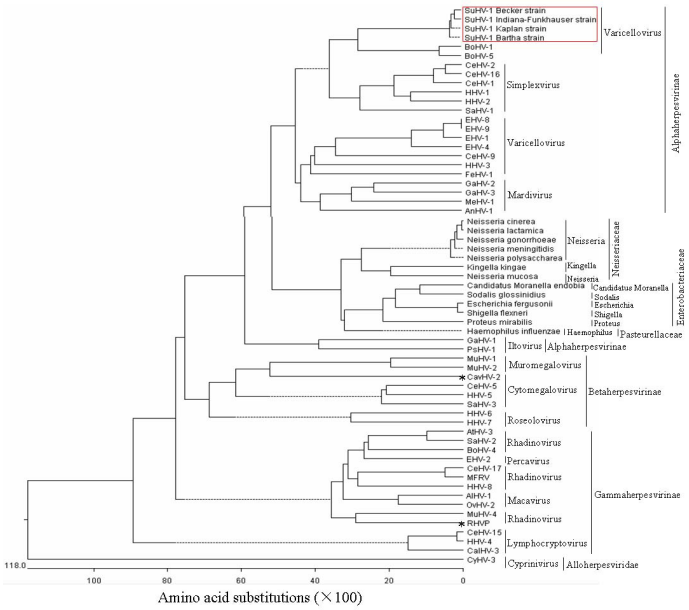


**Figure 2.** Evolutionary relationships of the putative PRV UL2 protein with 60 reference species (see Table 1). A phylogenetic tree of these proteins was generated using the MegAlign program in LASERGENE (DNAStar 7.0) with the Clustal V method, and sequence distance was calculated using weight matrix PAM250. Gaps were introduced by the alignment program to maximize the homology.

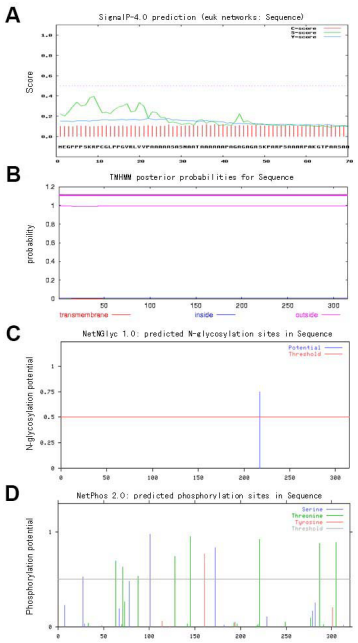


**Figure 3.** Prediction of signal peptide sequence, transmembrane domain, glycosylation site, and phosphorylation site of PRV UL2 analyzed using: **A.** SignalP-4.0 Server, **B.** TMHMM, **C.** NetNGlyc 1.0, **D.** NetPhos 2.0, respectively.

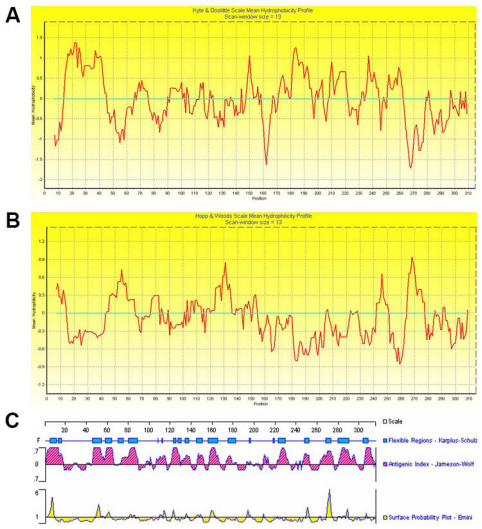


**Figure 4.** Hydrophobicity, hydrophilicity, and antigenic analyses of PRV UL2. The hydrophobicity (**A**) or hydrophilicity (**B**) profile was determined using values of Kyte and Doolittle (1982) or Hopp and Woods (1981), respectively, with a 13-amino acid window. Peaks pointing up represent the most hydrophobic (**A**) and hydrophilic (**B**) regions, respectively. **C.** Antigenic analysis of PRV UL2 was carried out through determination of its primary structure using the PROTEAN software of DNAStar based on flexibility, surface probability, and antigenic index.

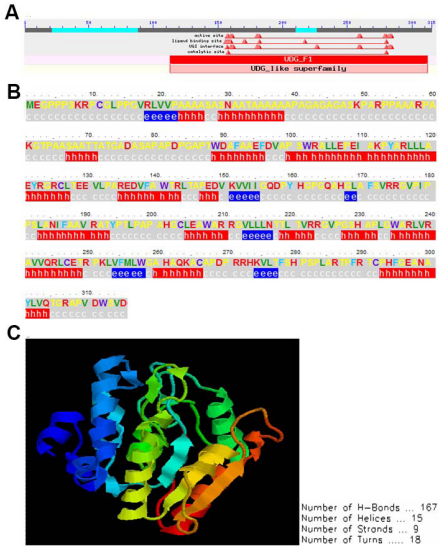


**Figure 5.** Conserved domain analysis, secondary structure, and three-dimensional (3-D) structure prediction of PRV UL2. **A.** Conserved domain analysis of PRV UL2 using the National Center for Biotechnology Information Conserved Domains search tool. The conserved active site, ligand binding site, uracil-DNA glycosylase inhibitor interface, and catalytic site are also shown. **B.** Secondary structure of PRV UL2 predicted using the PSIpred program. The letters h, e, and c indicate alpha helix, extended (beta strand), and coil, respectively. **C.** 3-D structure of PRV UL2 predicted using the protein modeling server database CPHmodels 3.2. The number of H-bond, helices, strands, and turns included in this model were 167, 15, 9, and 18, respectively.

## DISCUSSION

PCR amplification, cloning, and sequencing confirmed evidence of *UL2* in the PRV Becker strain, and the molecular properties of UL2 were analyzed using several bioinformatic tools. During evolution, viruses are generally conserved, and only a few genes undergo mutation (Antunes et al., 2010). Thus, viral evolution can be evaluated at the molecular level. Our analysis (see Table 2) revealed that the nucleotide sequence similarity of *UL2* of PRV Becker to that of the Bartha, Kaplan, and Indiana-Funkhauser strains was up to 98.9, 99.1, and 99.4%, respectively, and the aa sequence similarity was up to 98.4, 98.4, and 99.0%, respectively. These results revealed a close relationship among the studied PRV strains.

Mutation in viruses is well known to occur in response to environmental stress. Therefore, the small mutation of a different strain may have endowed the variance of PRV virulence. Multiple nucleotide sequences and aa alignments in PRV UL2 and the UL2-like proteins (see Table 2) showed that PRV UL2 has greater homology with the members of subfamily Alphaherpesvirinae, especially BoHV-1 (65.5 and 65.2%, respectively). Accordingly, PRV UL2 has relatively high homology with Alphaherpesvirinae but not with Betaherpesvirinae, Gammaherpesvirinae, Alloherpesviridae, Neisseriaceae, Enterobacteriaceae, or Pasteurellaceae. Furthermore, phylogenetic analysis (see Figure 2) unequivocally demonstrated that PRV belongs to the subfamily Alphaherpesvirinae, consistent with the conclusions of a previous report (McGeoch et al., 2000). Moreover, the results showed that UL2 is conserved among families of Herpesviridae, Neisseriaceae, Enterobacteriaceae, and Pasteurellaceae, possibly because it is a UDG protein.

Conserved domain analysis (see Figure 5) also indicated that UL2 obviously contains the conserved active site, ligand-binding site, UDG inhibitor interface, and catalytic site of the UDG enzyme within the UDG_F1 domain (Dean and Cheung, 1993; Pearl, 2000; Geoui et al., 2007). Consequently, UL2 might have a close relationship with the UDG family and high similarity with its counterparts encoded by UDG_F1 genes. Thus, UL2 may belong to the UDG_F1 family, consistent with homologs that have been shown to display UDG activity (Mullaney et al., 1989; Winters and Williams, 1993; Reddy et al., 1998; Lu et al., 2007). However, the PRV UL2 homolog of HCMV UL114 has been shown to play an important role in the replication rather than the repair of the viral genome (Ranneberg-Nilsen et al., 2008). Therefore, determining whether PRV UL2 can remove uracil from both U→G mispairs and U→A base pairs in double-stranded DNA or uracil in single-stranded DNA requires further study.

Protein phosphorylation is among the most normal and essential protein modifications, and certain aspects of cell process modulation are regulated through this mechanism. Signal transduction, proliferation, differentiation, and metabolism are controlled by a balance of the activities of protein kinases and protein phosphatases on pivotal target proteins. Phosphorylation site prediction (see Figure 3) revealed 12 potential phosphorylation sites in PRV UL2, including 3 serine, 8 threonine, and 1 tyrosine residues. Tyrosine phosphorylation is well known to be involved in the modification of protein translocation from the cytoplasm to the nucleus during productive viral infection (Pomeranz and Blaho, 1999) and the replication of several herpesviruses (Geiss et al., 2001; Ren et al., 2001). Phosphorylation of UDG at threonine is reported to be important for base excision repair (Lu et al., 2004), and

various phosphoforms of threonine and serine sites of the non-catalytic domain confer distinct functional properties to UDG, such as protein turnover, different activities, association with replication protein A, and nuclear and mitochondrial genomic integrity (Caradonna and Muller-Weeks, 2001; Hagen et al., 2008). Therefore, the phosphorylation of UL2 may also play an important role during PRV infection, perhaps modulating its subcellular localization or carrying out other uncharacterized functions such as improving base excision repair.

Sequence analysis revealed no potential transmembrane domain or signal peptide in UL2, indicating that it might be neither a transmembrane protein nor a secretory protein (Davis et al., 2006). However, a potential N-linked glycosylation site (Asn-Thr-Thr-Leu) located at aa 218 was present in UL2 (see Figure 3). N-linked glycosylation is an important post-translation modification that profoundly affects protein folding, oligomerization, and stability (Mitra et al., 2006). In addition, it is involved in trafficking modifications, ligand interactions (Meng et al., 2008), and enzyme activity (Hausmann et al., 1997; Sugahara et al., 2001). Because the potential N-linked glycosylation of UL2 occurs near the predicted beta-strand and within a potential strongly hydrophobic region, it might be associated with UDG activity. However, the possibility that this site was buried in the protein to avoid glycosylation cannot be excluded because a few UDG proteins were glycosylated.

Hydrophobicity analysis indicated that the hydrophobic regions were slightly larger than the hydrophilic regions, which might represent the internal and surface regions of the protein, respectively (see Figure 4). Furthermore, some predicted strong hydrophobic regions were located within the UDG_F1 domain, indicating that they may contribute to the α-helix or hydrophobic side chain that enables enzyme activity (Slupphaug et al., 1996; Xiao et al., 1999; Sartori et al., 2002). Secondary structure and 3-D structure predictions (see Figure 5) revealed that the overall structure of UL2 is almost the same as that of UDG; however, differences between UL2 and UDG have also been found, which may be attributed to different viral gene sequences or the diversity of uracil recognition mechanisms used by viral and bacterial enzymes (Savva et al., 1995; Slupphaug et al., 1996; Parikh et al., 1998; Xiao et al., 1999; Kaushal et al., 2010).

As more information becomes available on protein antigens, predictions of the locations of antigenic determinants may be possible before any immunological testing has been performed (Hopp and Woods, 1981). However, the elucidation of protein antigenic structures is presently a difficult, uncertain, and time-consuming task. Earlier methods relied on the assumption that the antigenic region is primarily the hydrophilic region at the surface of the protein molecule (Hopp and Woods, 1981; Welling et al., 1985). However, these methods have inaccuracies and limitations. To improve accuracy, the B-cell epitopes of UL2 were predicted through determinations of their primary structure using DNAStar PROTEAN programs based on flexibility, antigenic index, and surface probability (see Figure 4D). The results suggested that the enhanced knowledge of the antigenic and structural properties of UL2 resulting from this study might yield methods for developing new antibodies and immunoassays for application in the clinical diagnosis of PRV.

In conclusion, in this study we carried out the cloning and molecular characterization of PRV *UL2*. Elucidating the relationship between the molecular characterization and genetic evolution of PRV *UL2* will contribute to an understanding of this virus at the molecular level and enrich the herpesvirus database. The results herein will also provide insights for further research on the function and mechanism of *UL2* during PRV infection.

## ACKNOWLEDGMENTS

Research supported by grants from the Science and Technology New Star in Zhu Jiang, Guangzhou City (#2013J2200018); the Natural Science Foundation of Guangdong Province (#S2013040016596); the National Natural Science Foundation of China (#31200120); the Medical Scientific Research Foundation of Guangdong Province, China (#B2012165); the Scientific Research Foundation for the Ph.D., Guangzhou Medical University (#2011C20); the Foundation for Distinguished Young Talents in Higher Education of Guangdong, China; the First Batch of Youth Learning Backbone Teacher in Guangzhou Medical University; and the Students' extracurricular scientific and technological activities in Guangzhou Medical University (#2012A039 and #2012C007). We thank Dr. Lynn W. Enquist for the generous gift of pBecker2.

## REFERENCES

Antunes RS, Gomes VN, Prioli SM, Prioli RA, et al. (2010). Molecular characterization and phylogenetic relationships among species of the genus *Brycon* (Characiformes: Characidae) from four hydrographic basins in Brazil. *Genet. Mol. Res.* 9: 674-684.

Bogani F, Chua CN and Boehmer PE (2009). Reconstitution of uracil DNA glycosylase-initiated base excision repair in herpes simplex virus-1. *J. Biol. Chem.* 284: 16784-16790.

Caradonna S and Muller-Weeks S (2001). The nature of enzymes involved in uracil-DNA repair: isoform characteristics of proteins responsible for nuclear and mitochondrial genomic integrity. *Curr. Protein Pept. Sci.* 2: 335-347.

Davis MJ, Hanson KA, Clark F, Fink JL, et al. (2006). Differential use of signal peptides and membrane domains is a common occurrence in the protein output of transcriptional units. *PLoS Genet.* 2: e46.

Dean HJ and Cheung AK (1993). A 3' coterminal gene cluster in pseudorabies virus contains herpes simplex virus UL1, UL2, and UL3 gene homologs and a unique UL3.5 open reading frame. *J. Virol.* 67: 5955-5961.

Geiss BJ, Tavis JE, Metzger LM, Leib DA, et al. (2001). Temporal regulation of herpes simplex virus type 2 VP22 expression and phosphorylation. *J. Virol.* 75: 10721-10729.

Geoui T, Buisson M, Tarbouriech N and Burmeister WP (2007). New insights on the role of the gamma-herpesvirus uracil-DNA glycosylase leucine loop revealed by the structure of the Epstein-Barr virus enzyme in complex with an inhibitor protein. *J. Mol. Biol.* 366: 117-131.

Hagen L, Kavli B, Sousa MM, Torseth K, et al. (2008). Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. *EMBO J.* 27: 51-61.

Hausmann J, Kretzschmar E, Garten W and Klenk HD (1997). Biosynthesis, intracellular transport and enzymatic activity of an avian influenza A virus neuraminidase: role of unpaired cysteines and individual oligosaccharides. *J. Gen. Virol.* 78: 3233-3245.

Hopp TP and Woods KR (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U. S. A.* 78: 3824-3828.

Kaushal PS, Talawar RK, Varshney U and Vijayan M (2010). Structure of uracil-DNA glycosylase from *Mycobacterium tuberculosis*: insights into interactions with ligands. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 66: 887-892.

Kramer T, Greco TM, Enquist LW and Cristea IM (2011). Proteomic characterization of pseudorabies virus extracellular virions. *J. Virol.* 85: 6427-6441.

Kyte J and Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol*. 157: 105-132.

Li M, Wang S, Cai M, Guo H, et al. (2011a). Characterization of molecular determinants for nucleocytoplasmic shuttling of PRV UL54. *Virology* 417: 385-393.

Li M, Wang S, Cai M and Zheng C (2011b). Identification of nuclear and nucleolar localization signals of pseudorabies virus (PRV) early protein UL54 reveals that its nuclear targeting is required for efficient production of PRV. *J. Virol.* 85: 10239-10251.

Lu CC, Huang HT, Wang JT, Slupphaug G, et al. (2007). Characterization of the uracil-DNA glycosylase activity of Epstein-Barr virus BKRF3 and its role in lytic viral DNA replication. *J. Virol.* 81: 1195-1208.

Lu X, Bocangel D, Nannenga B, Yamaguchi H, et al. (2004). The p53-induced oncogenic phosphatase PPM1D interacts with uracil DNA glycosylase and suppresses base excision repair. *Mol. Cell* 15: 621-634.

McGeoch DJ, Dolan A and Ralph AC (2000). Toward a comprehensive phylogeny for mammalian and avian herpesviruses. *J. Virol.* 74: 10401-10406.

Meng J, Parroche P, Golenbock DT and McKnight CJ (2008). The differential impact of disulfide bonds and N-linked glycosylation on the stability and function of CD14. *J. Biol. Chem.* 283: 3376-3384.

Mitra N, Sinha S, Ramya TN and Surolia A (2006). N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* 31: 156-163.

Mullaney J, Moss HW and McGeoch DJ (1989). Gene UL2 of herpes simplex virus type 1 encodes a uracil-DNA glycosylase. *J. Gen. Virol.* 70: 449-454.

Parikh SS, Mol CD, Slupphaug G, Bharati S, et al. (1998). Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.* 17: 5214-5226.

Pearl LH (2000). Structure and function in the uracil-DNA glycosylase superfamily. *Mutat. Res.* 460: 165-181.

Pomeranz LE and Blaho JA (1999). Modified VP22 localizes to the cell nucleus during synchronized herpes simplex virus type 1 infection. *J. Virol.* 73: 6769-6781.

Prichard MN, Lawlor H, Duke GM, Mo C, et al. (2005). Human cytomegalovirus uracil DNA glycosylase associates with ppUL44 and accelerates the accumulation of viral DNA. *Virol. J.* 2: 55.

Pyles RB and Thompson RL (1994). Evidence that the herpes simplex virus type 1 uracil DNA glycosylase is required for efficient viral replication and latency in the murine nervous system. *J. Virol.* 68: 4963-4972.

Ranneberg-Nilsen T, Dale HA, Luna L, Slettebakk R, et al. (2008). Characterization of human cytomegalovirus uracil DNA glycosylase (UL114) and its interaction with polymerase processivity factor (UL44). *J. Mol. Biol.* 381: 276-288.

Reddy SM, Williams M and Cohen JI (1998). Expression of a uracil DNA glycosylase (UNG) inhibitor in mammalian cells: varicella-zoster virus can replicate *in vitro* in the absence of detectable UNG activity. *Virology* 251: 393-401.

Ren X, Harms JS and Splitter GA (2001). Tyrosine phosphorylation of bovine herpesvirus 1 tegument protein VP22 correlates with the incorporation of VP22 into virions. *J. Virol.* 75: 9010-9017.

Sartori AA, Fitz-Gibbon S, Yang H, Miller JH, et al. (2002). A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. *EMBO J.* 21: 3182-3191.

Savva R, McAuley-Hecht K, Brown T and Pearl L (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature* 373: 487-493.

Slupphaug G, Mol CD, Kavli B, Arvai AS, et al. (1996). A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* 384: 87-92.

Smith GA and Enquist LW (2000). A self-recombining bacterial artificial chromosome and its application for analysis of herpesvirus pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 97: 4873-4878.

Strang BL and Coen DM (2010). Interaction of the human cytomegalovirus uracil DNA glycosylase UL114 with the viral DNA polymerase catalytic subunit UL54. *J. Gen. Virol.* 91: 2029-2033.

Sugahara K, Hongo S, Sugawara K, Li ZN, et al. (2001). Role of individual oligosaccharide chains in antigenic properties, intracellular transport, and biological activities of influenza C virus hemagglutinin-esterase protein. *Virology* 285: 153-164.

Ward TM, Williams MV, Traina-Dorge V and Gray WL (2009). The simian varicella virus uracil DNA glycosylase and dUTPase genes are expressed *in vivo*, but are non-essential for replication in cell culture. *Virus Res.* 142: 78-84.

Welling GW, Weijer WJ, van der Zee R and Welling-Wester S (1985). Prediction of sequential antigenic regions in proteins. *FEBS Lett.* 188: 215-218.

Winters TA and Williams MV (1993). Purification and characterization of the herpes simplex virus type 2-encoded uracil-DNA glycosylase. *Virology* 195: 315-326.

Xiao G, Tordova M, Jagadeesh J, Drohat AC, et al. (1999). Crystal structure of *Escherichia coli* uracil DNA glycosylase and its complexes with uracil and glycerol: structure and glycosylase mechanism revisited. *Proteins* 35: 13-24.