Stepwise Evolution of the Herpes Simplex Virus Origin Binding Protein and Origin of Replication.

Monica Olsson¹, Ka-Wei Tang¹, Cecilia Persson², L. Marcus Wilhelmsson³, Martin Billeter⁴ and Per Elias¹

1. Institute of Biomedicine, Department of Medical Biochemistry and Cell Biology, Sahlgrenska Academy, University of Gothenburg, Box 440, S-405 30 Gothenburg, Sweden
2. The Swedish NMR Centre at University of Gothenburg, Box 465, S-405 30 Gothenburg, Sweden
3. Chalmers University of Technology, Department of Chemical and Biological Engineering/Physical Chemistry, Kemivägen 10, S-412 96
4. Biophysics Group, Department of Chemistry, University of Gothenburg, Box 462, S-405 30 Gothenburg, Sweden

Running title: Evolutionarily conserved properties of HSV1 replicon

Address correspondence to: Per Elias, Institute of Biomedicine, Department of Medical Biochemistry and Cell Biology, Sahlgrenska Academy, University of Gothenburg, Box 440, S-405 30 Gothenburg, Sweden
Tel 46-31-7863486; Fax 46-31-416108; E-mail: per.elias@medkem.gu.se

The Herpes simplex virus replicon consists of cis-acting sequences, oriS and oriL, and the origin binding protein, OBP, encoded by the UL9 gene. Here, we identify essential structural features in the initiator protein OBP and the replicator sequence oriS, and relate the appearance of these motifs to the evolutionary history of the alphaherpesvirus replicon. Our results reveal two conserved sequence elements in HSV-1 OBP: the RVKNL motif, common to and specific for all alphaherpesviruses, is required for DNA binding, and the WPxxxGAXxFxxL motif, found in a subset of alphaherpesviruses, is required for specific binding to the single-strand DNA binding protein ICP8.

A 121 amino acid minimal DNA binding domain containing conserved residues is not soluble and does not bind DNA. Additional sequences present 220 amino acids upstream the RVKNL motif are needed for solubility and function.

We also examine the binding sites for OBP in origins of DNA replication and how they are arranged. NMR and DNA melting experiments demonstrate that origin sequences derived from many, but not all, alphaherpesviruses can adopt stable boxI/boxIII hairpin conformations. Our results reveal a stepwise evolutionary history of the Herpes simplex virus replicon and suggest that replicon divergence contributed to formation of major branches of the herpesvirus family.

Herpesviruses have been found in animal species ranging from molluscs to man. The order of Herpesvirales consists, according to the International Committee on Taxonomy of Viruses, of three families: Alloherpesviridae, Herpesviridae and Malacoherpesviridae (1). The subfamilies Alphaherpesvirinae, Betaherpesvirinae and Gammaherpesvirinae are found within the family of Herpesviridae. The events leading to establishment of a new virus...
species are poorly understood, but in the case of herpesviruses it is commonly assumed that viruses co-evolve with their hosts (2). Herpesviruses have thus become well adapted to their hosts and may reside in a latent state in the host for a lifetime with little or no overt signs of infection. Upon reactivation infectious virus will be released. The viruses remain faithful to their hosts and infections across species borders are rare but may under specific circumstances give rise to fatal disease.

Replication of Herpes Simplex virus type 1, HSV1, requires a cis-acting DNA sequence, the replicator, termed oriS or orIL, an initiator protein, OBP or UL9 protein, and a replisome composed of DNA polymerase, helicase-primase and a single-strand DNA-binding protein referred to as ICP8 or UL29 protein (3). OBP assisted by ICP8 can in an ATP-dependent reaction unwind double-stranded oriS (4,5). The resulting single-stranded DNA adopts a hairpin conformation, which is stably bound by OBP (6,7). The herpesvirus replisome once assembled on DNA is capable of synthesising leading and lagging strands processively in a coordinated fashion (8). DNA replication is likely to start on circular molecules produced by the action of DNA ligase IV/XRCC4 and proceed in a theta-type manner (9,10). Later, a rolling-circle mode of replication dominates giving rise to characteristic head-to-tail concatemers. The initiator protein OBP appears to be strictly required only during the first few hours of the infectious cycle (11,12).

The HSV1 origin binding protein was first isolated using an assay monitoring specific binding to HSV1 oriS (13). A minimal DNA binding site was identified using footprinting techniques as well as binding studies with double-stranded oligonucleotides (14). For alphaherpesviruses the high affinity binding site is always TTCGCAC with a minor exception for CHV1 also referred to as monkey B virus (Table 1). The C-terminal 317 amino acids of alpha herpesvirus OBP can be isolated as a soluble protein, which remains capable of high affinity binding to the sequence TTCGCAC (15). The C-terminal domain of HSV1 OBP, here referred to as ΔOBP binds as a monomer to the major groove in DNA and makes contacts with base-pairs as well as the deoxyribose-phosphate backbone (16, 17). ΔOBP binds DNA specifically with an estimated $K_d$ of 0.3 nM, a value which is highly influenced by the composition of the assay buffer (16). At high protein concentrations ΔOBP form aggregates, which, still in a sequence specific manner, binds DNA (16). A number of studies have attempted to define amino acids involved in DNA binding (18, 19, 20, 21). In addition, sequence comparisons between alphaherpesviruses and roseoloviruses have helped to identify amino acids in OBP potentially involved in DNA binding as well as corresponding recognition sequences in origins of DNA replication (22, 23, 24, 25). However, a comprehensive and quantitative study of evolutionarily conserved amino acids required for DNA binding is still lacking. The single-strand DNA binding protein, ICP8 encoded by the UL29 gene, is involved in initiation of DNA replication, and it also participates in the elongation phase at the replication fork (4, 5, 26).

ICP8 forms a specific complex with OBP (26, 27). Studies with deletion mutants have demonstrated that important sequences are found close to the C-terminus of OBP, but amino acids directly participating in the high affinity interaction have not been identified. The interaction is biologically significant, since deletion of the extreme C-terminus enhances the helicase activity of OBP but reduces origin dependent DNA synthesis (26).

The HSV1 oriS contains three copies of the binding site for OBP; two binding sites, box I and box II are high affinity sites, whereas the third binding site, box III, has a very low affinity for OBP (14)(Figure 1). All sites are required for efficient replication, and in a competitive situation there is a strong selection for the most efficient replicator sequence (28). Box III and box I are arranged in a palindrome, which becomes rearranged upon activation to form an alternative conformation, most likely a hairpin (4, 6,
7). Point mutations that prevent formation of a hairpin reduce replication, and compensatory mutations restore complementary base pairing as well as the ability to replicate (7).

To learn more about the mechanism of virus evolution we have then examined the evolutionary history of some functionally significant features of the HSV1 replicon and related them to a sequence based evolutionary tree. The results indicate that replicon divergence, characterized by the stepwise appearance of the DNA binding RVKNL motif, the WPxxxGxxFxL ICP8 binding motif and the box III- box I palindrome, may have played important roles in establishing major branches of the alphaherpesvirus tree.

**Experimental procedures**

**Oligonucleotides and plasmids**

Oligonucleotides were purchased from Eurofins MWG Operon (Germany). Single-stranded oligonucleotides were annealed to make substrates for DNA-binding assays. They were radiolabelled at the 3’ends with α-32P dATP 3000 mCi/mmol as previously described (16).

The oligonucleotides

5’GATCTGCGAAGCGTTCGCACTTCG TCCCAATG 3’

and

5’GATCCATTGGGACGAAGTGCGACGCTTCGCA 3’

were annealed to make a box I high affinity binding site for HSV1 OBP (Fig.1).

The oligonucleotides

5’GATCTGCGAAGCGTCCGCACTTCG TCCCAATG 3’

and

5’GATCCATTGGGACGAAGTGCGACGCTTCGCA 3’

were annealed to make box I high affinity binding sites for CHV1 OBP (monkey B virus).

The plasmid pORI(boxImut) contains a mutated version of oriS with only one high affinity binding site for HSV1 OBP (Fig.1). A restriction fragment containing oriS was isolated and radiolabelled as previously described (29).

PCR was used to generate DNA fragments containing a tentative minimal DNA binding domain consisting of amino acids 717-814 of HSV1 OBP/UL9. The primers were:

F1 5’CATATGAGCGGCTGTCCCCCGCGCGGGGTCC 3’

and R1 5’GGATCCTTAGCGAACCACAGCGGCCCACGCGGGG 3’

The resulting DNA fragment was cloned in a pTOPO vector (Invitrogen) and verified by DNA sequencing. It was then transferred as Ndel/BamHI fragments to the pET14B (Novagen). The resulting protein product had the sequence

MGSSHHHHHHHSGLVPRGSHMPPRG
SVESETDHAVALKWGFELFVQMTK
STQTQPAGRVRKNTKQTIVGGLLDAH
HIDHSACRTHRQYLLMAHKREFGARFKLRVPWAIRCRCLRT.

Using the same strategy PCR fragments containing amino acids 536-813, 553-813, 556-813, 558-813, 567-813 of HSV1 OBP/UL9 were cloned and expressed in E.coli. A plasmid encoding CHV1 UL9 protein was provided by Dr. Richard Eberle (USA). A PCR fragment encoding the C-terminal DNA binding domain was used to replace HSV1ΔOBP and make pET3aHisΔCHV1UL9. The forward primer pair was

5’GGATCCGACACCAACGTGTCGCTGCAGGC 3’

and

5’CCTAAGGTGCGGTGCCGCCGGGCACCTAT 3’

The resulting DNA fragment was cloned in a pTOPO vector and verified by DNA sequencing. It was then transferred as BamHI/Bsu36I fragment to the pET14B.

To investigate the interaction between OBP and ICP8 a cassette mutagenesis strategy was devised. pET3aHisΔOBP was modified to contain a SnaBI site and duplex oligonucleotides were inserted into SnaBI and SacII cleaved vector to create HisΔOBP with altered C-termini. The resulting constructs are shown in table 2.

HisΔOBP with mutations in the DNA binding domain designated F751/A, P752G753/AA, R756/A, V757/A, K758N759/AA, N759/A, N759L760/AA and L768L768/AA were all made using Quickchange II Site directed mutagenesis kit from Stratagene and verified by DNA sequencing.

**Proteins**
The His-tagged C-terminal domain of HSV1 OBP, here named HisΔOBP, was produced as previously described (16). The DNA binding domain of CHV1 OBP was also isolated as a His-
tagged fusion protein following the same protocol. His\(\Delta\)OBP with altered C-termini, the various truncated versions of His\(\Delta\)OBP and the alanine substitution mutants were also isolated by the same procedure. Mass spectrometry was used for protein identification. ICP8, UL29 protein, was purified using recombinant baculovirus as previously described (30).

**Gel electrophoresis of protein-DNA complexes**—To evaluate sequence specific DNA binding by HSV1 His\(\Delta\)OBP and CHV1 His\(\Delta\)OBP a polyacrylamide gel mobility shift assay in was used (16). Samples, 20 \(\mu\)l, containing 1 nM radiolabelled duplex oligonucleotides and indicated amounts of protein were incubated on ice before gel electrophoresis. The dried gels were analysed by a phosphorimager.

To evaluate the specificity of the OBP-ICP8 interaction we employed a previously described gel electrophoresis assay in which a ternary complex between double-stranded oriS, His\(\Delta\)OBP and ICP8 could be detected (29). Sonicated calf thymus DNA, 1 \(\mu\)g/ml, was included to suppress interactions between ICP8 and DNA. In control experiments we made sure that the ternary complex was disrupted following addition of a 65-mer of oligodeoxyxymidylate T65 to leave the binary complex oriS- His\(\Delta\)OBP and ICP8 as demonstrated in ref 29 (results not shown).

**NMR**—Single-stranded oligonucleotides containing sequences from oriS supporting formation of the putative box III-box I hairpin were analyzed by NMR. The palindrome oligonucleotides 5´GAAGTGAGAACGCGTTCGCACTTC 3´, including a central loop of three nucleotides and a mismatch at position 7, and 5´AAGTGCAACATGTAGTGTTCGCACTTC 3´, with a five nucleotide loop and no mismatch, represent HSV1 and EHV4 respectively.

For all spectra, 0.5-1 mM samples of oligonucleotide in 50 mM potassium phosphate buffer, pH 6 were used. Before recording the sample was heated to a temperature where the imino signals were not visible and then cooled down to the temperature used, i.e. 15°C or 25°C. All NMR spectra were recorded on Varian Inova 500 or 600 MHz spectrometers equipped with triple resonance probes. A set of DQF-COSY TOCSY and NOESY in \(D_2O\) was recorded for assignment purposes (31). Typically, 64 transients with each 1024 complex points in t2 and 256 increments in t1 were acquired. Two different mixing times were used in TOCSY experiments (50 and 100 ms) as well as in NOESY experiments (100 and 250 ms). The sweep width was set to 5000 Hz and 6000 Hz for spectra recorded at 500 MHz and 600 MHz, respectively. A weak presaturation pulse was used for water suppression. NOESY spectra were also recorded in a mixture of 90%/\(H_2O:10\%\ D_2O\) using the watergate sequence for water suppression (32) but otherwise the same parameters as above. A \(^{1}H^{13}P\) spectrum (144 transients; 256 complex points in t1; 1024 complex points in t2; 1000 Hz sweep width in phosphor dimension; 5000 Hz sweep width for protons) was recorded at 500 MHz using a triple resonance probe.

All spectra were processed with nmrPipe (33). The time-domain data sets were typically zero-filled once in each dimension and Fourier transformed. A polynomial baseline correction was used. For the spectral analysis, resonance assignments, peak integration and measurement of J couplings the program XEASY (34) was used. Structure calculations were performed with the program DYANA following standard procedures (35).

**DNA melting experiments**—Oligonucleotides representing putative Box III-Box I hairpins from a selected set of alphaherpesviruses were as follows:

- **BHV1**
  5´GTGAGAAGTCGCGAATGCAGCGTTCGCAC 3´ (14 basepairs/1 mismatch/3 nucleotides in the loop)

- **HSV1**
  5´AAGTGCAACATGTAGTGTTCGCAC 3´ (14 basepairs/1 mismatch/3 nucleotides in the loop)

- **EHV1**
  5´AAGTGCCGAACATGTAGTGTTCGCAC 3´ (14 basepairs/1 mismatch/3 nucleotides in the loop)

- **EHV4**
  5´AAGTGCAACATGTAGTGTTCGCAC 3´ (14 basepairs/1 mismatch/3 nucleotides in the loop)

For all spectra, 0.5-1 mM samples of oligonucleotide in 50 mM potassium phosphate buffer, pH 6 were used. Before recording the sample was heated to a temperature where the imino signals were not visible and then cooled down to the temperature used, i.e. 15°C or 25°C. All
A Cary 4B spectrophotometer equipped with a programmable multicell temperature block was used for the Tm measurements. Changes in the absorbance were measured at the maximum of the DNA absorption band (260 nm). The concentration of the six different hairpin forming DNA strands was chosen so that the absorbance at 15°C was between 0.69-0.72. The buffer used was 20 mM HEPES, 10 mM Na+, pH 7.5. The temperature of the samples was continuously changed at a rate of 0.5°C per minute from 15°C to 92°C and back to 15°C. To make it easier to compare the melting temperatures of the hairpins, the minute differences in absorption at 15°C were ignored (i.e. all start absorptions are set equal). Tm values were determined from the temperature at peak height maximum of the derivative of the melting curves.

The varying slopes of the melting curves were reflected in the thermodynamic parameters ΔH° and ΔS° in Table 1 estimated according to the van’t Hoff equation for a two-state melting transition (36): 

$$\Delta H^o = 4RT_m^2 \frac{d\alpha}{dT} \bigg|_{max}$$

$$\Delta S^o = \Delta H^o / T_m$$

where α is the fraction of DNA molecules present as hairpins.

Results.

The DNA binding domain of the alphaherpesvirus origin binding protein—

Recently published genome sequences from alphaherpesviruses such as the infectious laryngotracheitis virus or GHV1, the psittacid herpesvirus, PsHV1, and the fibropapillomatosis virus from sea turtles, CPFHV, have shown that GHV1 and PsHV1 retain a putative replication origin containing the sequence TTCGCAC (Table 1) (25,37). The origins of replication in CPFHV have not been sequenced. Sequence comparisons including the evolutionary distant alphaherpesviruses reveal a highly conserved region in the C-terminus of OBP between amino acids L731 and L812 in HSV1 OBP (Supplement figure 1). Acting on the assumption that conserved amino acids would be responsible for binding to a conserved DNA sequence we examined closer a RVKNL motif found in OBPs from all alphaherpesviruses (Table 1 and Supplement figure 1) as well as conserved amino acids surrounding this motif. First, R756/A and K758N759/AA mutants were made and their DNA binding properties examined under optimized binding conditions. The R/A mutant had a greatly reduced affinity for the TTCGCAC sequence (Figure 2). At the highest concentration the complex of lower mobility indicated binding of multiple copies of HisΔOBP (16). The K758N759 mutant did not bind DNA at all. A previous study of a K758/A mutation arrived at a similar conclusion (18). The results indicate that the evolutionarily RVKNL motif was intimately involved in sequence specific binding of OBP to DNA. We next examined an additional set of mutant proteins (Figure 3). The evolutionarily conserved amino acids F751, P752G753, V757, N759, N759L760 and L768L769 were changed to alanines. Binding experiments show that neither the NL/AA mutant nor the F/A mutant bound DNA. The LL/AA mutant had a severely reduced affinity for DNA, but the PG/AA, V/A and N/A mutants revealed a more modest reduction of DNA binding (Figure 3A). A more detailed study was then performed with the latter mutant proteins using the wild type for comparison Figure 3B, C and D). We note that the N/A mutant is very similar to the wild type, and it also supports formation of a higher order complex. The V/A mutant has reduced affinity for DNA and does not support formation of a higher order complex. The PG/AA mutant has intermediate affinity and does not readily form higher order complexes.
The conserved RVKNL motif is unique to alphaherpesviruses using TTCGCAC as recognition sequence and it is therefore most likely involved in sequence specific recognition of DNA. The arginine and lysine residues are probably interacting with the phosphate residues previously identified by ethylation interference studies (16). Other conserved residues common to alphaherpesviruses and roseoloviruses may be needed for proper folding of the DNA binding domain. Conserved residues specific for alphaherpesviruses are candidates for specific base recognition. Further insight into these matters may be gained once structural information can be combined with a mutagenesis study.

We have also attempted to express a minimal version of the DNA binding domain, which would still retain the ability to bind DNA. The borders of the domain were chosen taking previously published suggestions, the results referred to above as well as recently published OBP sequences into account (18)(Supplement figure 1). The 121 amino acid large domain, amino acids 717-814, was expressed as a His-tagged fusion protein and purified. Only a small amount of soluble protein was obtained. Moreover, we could not detect any sign of sequence specific DNA binding (results not shown). This observation indicates that additional sequences are required for DNA binding or proper folding. We therefore made a series of recombinant proteins containing amino acids 536-813, 553-813, 556-813, 558-813, 567-813 of HSV1 OBP. Strikingly, only the protein containing amino acids 536-813 was soluble and bound DNA. All other constructs were insoluble when expressed under standard conditions in E. coli (Supplement figure 2). It therefore appears that a region of HSV1 OBP containing a short hydrophobic stretch of amino acids together with the immediate downstream F553xKYL motif is essential for proper folding of the DNA binding domain (Figure 1).

Since CHV1 or monkey B virus oriS contains a variant, TCCGCAC, of the canonical recognition sequence for OBP we examined the DNA binding properties of corresponding protein (38,39,40). We found that the DNA binding properties of HSV1 OBP and CHV1 OBP are identical with respect to the HSV-1 sequence, TTCGCAC, and the CHV1 sequence TCCGCAC (Figure 4). It is also worth noting that the DNA-binding motif, RVKNL, remains unchanged (Table 1). These observations indicate that the sequence difference contributes little, if any, to species specific replication by closely related alphaherpesviruses. It is unfortunate that CHV1 oriS cannot be stably propagated in common cloning vectors and therefore we do not yet know whether other subtle properties allow discrimination between HSV1 and CHV1 origins of replication.

The ICP8 binding motif- Here we searched for the ICP8 binding motif found in the C-terminus of OBP using a cassette mutagenesis technique. The formation of complexes between HisOBP and ICP8 is measured using a gel mobility shift assay (29). In this instance a supershift of the HisOBP–oriS complex can be seen upon addition of ICP8. A restriction fragment containing a mutated version of oriS containing only one high affinity site for OBP was used to avoid melting of DNA by ICP8. The specificity of complex formation was verified by addition of a 65-mer of oligodeoxythymidylate, which has been shown to disrupt HisOBP/ICP8 complexes (results not shown). We first found that deletion of the last 13 amino acids, WPMMQGAVNFSTL, in HisOBP completely prevented the formation of an HisOBP/ICP8/oriS complex (Figure 5). We also compared the binding of ICP8 to an altered version of HisOBP now containing the last 13 amino acids from VZV OBP. We noted that a HisOBP/ICP8/oriS could still be formed although the mobility of the complex was slightly altered possibly reflecting a reduced affinity (Figure 6). It is therefore likely that primarily the evolutionarily conserved amino acids contribute to the OBP/ICP8 interaction. These amino acids were then systematically replaced by alanines (Table 2 and Figure 7). The mutant proteins were then analyzed using...
the mobility shift assay. We found that only the hydrophobic amino acids, W839 and F848, were essential for complex formation. Replacing P840 and G844 with A did not affect complex formation. Interestingly, we found, after performing a complete survey of OBP sequences hitherto made available in GenBank, that the sequence motif WPxxGxxFxxL/I is absent in OBP from infectious lymphotracheomaotisis virus, GHV1, the fibropapillomatosis virus from sea turtles, CFPHV, the psittacid herpesvirus, PsHV1, and the roseoloviruses, HHV6a, HHV6b and HHV7 (Table 1). We therefore suspect that these proteins do not interact with ICP8 in the same way as other alphaherpesviruses.

**The box III-box I hairpin in oriS** Here we investigate by NMR the ability to form a hairpin of selected single-stranded oligonucleotides containing the box III-box I region of oriS from HSV1, with one mismatch, and from fully palindromic EHV4. Following nearly complete resonance assignments, including all imino protons except for terminal nucleotides, the formation of base pairs is indicated. This is confirmed by the observation of sequential connectivities for all base pairs except for the mismatch in the HSV1 hairpin. Furthermore, measurements of scalar couplings between $^{31}$P and $^1$H were consistent with B-DNA conformation. Finally, NOE distance restraints were obtained that cover the entire molecule. Although their number was not sufficient to warrant an independent structure calculation, they are in agreement with B-DNA except for loop regions and the mismatch as illustrated for HSV1 in figure 8A.

We have also performed DNA melting experiments with single-stranded oligonucleotides representing a larger number of alphaherpesviruses (Figure 8b). In all instances stable hairpins are formed with $T_m$ values ranging from 66 °C for EHV1 to 84 °C for GHV3. The melting temperatures of hairpins depend, among other things, upon the number of base pairs in the stem, the number of mismatches, the GC-content and the number of bases included in the loop region. It can be seen from the $T_m$ values and from the parameters in Table 3 that $T_m$ is influenced positively by the number of base pairs in the hairpin stem and the GC-content and negatively by the number of mismatches. The varying slopes of the melting curves are reflected in the thermodynamic parameters $\Delta H^0$ and $\Delta S^0$ in Table 1 estimated according to the van’t Hoff equation for a two-state melting transition (see experimental procedures). Although the thermodynamic parameters should be interpreted with some caution, it is interesting to note the low $\Delta H^0$ values for EHV1, HSV1/2 and BHV1, which have the smallest hairpin loops (3 bases). A plausible explanation is that a tight loop favors formation of partially melted intermediates, which may broaden the melting intervals and consequently lead to a lowering of the calculated $\Delta H^0$. We have also looked at melting and annealing of the EHV4 hairpin to investigate differences in these processes. Importantly, there was no hysteresis, i.e. the rate of temperature change (0.5°C per minute) was slow enough for the system to reach equilibrium. Not even an increase in the temperature change by a factor of five gave any signs of hysteresis confirming that the DNA strands are forming hairpins rather than intermolecular duplexes (results not shown) (41, 42).

Finally, we examined origins of DNA replication in genomes of completely sequenced herpesviruses and tabulated the arrangement of OBP binding sites (Table 1). We found that three different designs of the origins of replication. First, in at least 12 completely sequenced alphaherpesvirus origins of replication allow formation of a box III-box I hairpin. Second, VZV and the related CHV9 monkey virus have only a single copy of the OBP binding site. Finally, the evolutionarily most distant viruses, infectious laryngotraceheitis virus, GHV1, the fibropapillomatosis virus from sea turtles, CFPHV, the psittacid herpesvirus, PsHV1, and the roseoloviruses, HHV6a, HHV6b and HHV7 have two OBP binding sites. Since it has been demonstrated that the arrangement of OBP binding sites strongly affects the efficiency
of replication it is likely that rearrangement of DNA binding sites might contribute to specificity of DNA replication (43, 44, 45, 46, 47, 48).

Discussion.
We have examined functionally significant and evolutionarily conserved motifs in the initiator protein OBP from alphaherpesviruses as well as in the corresponding replicator sequences recognized by OBP. Our observations highlight some general mechanisms of origin activation and suggest a series of distinct steps, which have contributed to the evolution of the alphaherpesvirus replicon.

Mechanism of origin activation—Activation of an origin of DNA replication ends by recruitment of the complete replisome. The steps preceding the final events include site selection and separation of the two DNA strands. For HSV1 the origin binding protein, OBP or UL9 protein, performs these two reactions (4, 5). Experiments in vitro and in vivo also argue in favour of an essential role of the single-strand DNA binding protein ICP8 (4, 5, 26). The latter seems to perform its function through a specific interaction with OBP (26, 27). We have here defined a sequence motif, WPxxxGAxxFxxL/I present in the extreme C-terminus of OBP, required for binding to ICP8. Once ICP8 binds single-stranded DNA the interaction with OBP is broken, and it is assumed that the remaining components of the replisome associate with the activated origin (29). We have previously demonstrated that an OBP dimer may simultaneously bind two double-stranded oligonucleotides containing the recognition sequence TTCGCAC but only one copy of a hairpin with a single-stranded tail resembling activated oriS (49). Since helicase activity of OBP is enhanced after deletion of the extreme C-terminus it is possible that ICP8 may facilitate a conformational change in OBP, from an origin binding conformation to a helicase conformation, through the specific interaction with the WPxxxGAxxFxxL/I motif (26). Together these observations point towards an active role for the C-terminal DNA binding domain of OBP, here referred to as ΔOBP, during activation of oriS.

We have now delineated the DNA binding domain and identified an evolutionarily conserved amino acid motif, RVKNL, which is unique to alphaherpesviruses and essential for binding to DNA. R756 and K758 are probably interacting with the phosphates previously identified by ethylation interference analysis (16). A NL/AA mutant cannot bind DNA, and a V/A mutant has a moderately reduced affinity for DNA. These amino acids may contact base-pairs, but further studies are needed to clarify their contribution, if any, to the sequence specificity of OBP. In addition we investigate that an additional set of evolutionarily conserved amino acids and demonstrate that some of these, F751 and L768L769, are essential for DNA binding. These amino acids, however, also appear in roseoloviruses and may therefore have a more general role in establishing a proper structure of the DNA binding domain.

To put these studies to test we expressed the minimal domain, amino acids 717-814, containing the evolutionarily conserved amino acids discussed above (Supplement figure 1. Only a small amount of soluble protein was obtained, and it did not bind double-stranded DNA (Results not shown). A considerably larger version, amino acids 536-813, was needed to obtain soluble protein. Interestingly, a slightly shorter version, 553-813 was completely insoluble and further truncation did not alleviate the situation (Supplement figure 2). We therefore argue that a conserved F553xxKYL motif preceded by a few additional amino acids assists in the proper folding of the protein and possibly also in the formation of the DNA binding site (Fig. 1). We speculate that this motif might also be part of a conformational switch allowing the OBP dimer to control the number of bound DNA ligands. This model has some support also from the observation that a K-A mutation in the FxxKYL motif appears to have a temperature-sensitive phenotype for sequence specific DNA-binding (18).
A second aspect of origin activation concerns the DNA sequence. The number of binding sites for OBP, the orientation and affinity of the OBP binding sites, the distance between the binding sites and the composition of the spacer sequences are all features that govern efficiency and specificity of initiation of DNA replication. We have demonstrated that efficient origins are rapidly selected from a library with partially random sequences arguing that the initiator, OBP, and the replicator, oriS, have co-evolved to facilitate efficient initiation of DNA synthesis (28). We have argued that many of these observations can be explained if activated oriS forms a DNA hairpin captured by OBP (4,6,7,49). Indeed, mutations that disrupt hairpin formation reduce replication efficiency and compensatory mutations restore the ability to initiate DNA replication (7). Here we demonstrate by NMR and DNA melting experiments that a large number of alphaherpesviruses have the capacity to form hairpins in oriS. What is less clear is to what extent the different alphaherpesviruses are able to discriminate between cognate and non-cognate origins. It is, in fact, unclear if discrimination between origins would be meaningful for viruses that are unlikely to exist in the same replication compartment. It would, nevertheless, be interesting to see if minor differences in nucleotide and thermodynamic stability of putative hairpin structures contribute to specific recognition and activation of origins of DNA replication from alphaherpesviruses, and to see how such properties would be reflected in the structure of the initiator protein OBP.

Evolutionary considerations—There are currently 17 completely sequenced genomes from different alphaherpesviruses available. In addition, three complete genome sequences for roseoloviruses, betaherpesviruses with a replicon similar to alphaherpesviruses, have been published. The relative abundance of sequence information now makes room for a discussion about the evolutionary roots of the alphaherpesvirus replicon and the significance of the structural differences observed in replicator sequences and initiator proteins. The prototype replicon makes use of replicator sequences, origins of DNA replication, akin to oriS from Herpes simplex virus type 1 and initiator proteins, which are orthologues to the HSV1 origin binding protein, OBP, encoded by the UL9 gene. A conserved replisome composed by orthologues to HSV1 DNA polymerase, UL30/UL42, helicase-primase, UL5/UL8/UL52, and single strand DNA-binding protein ICP8, UL29, subsequently performs coordinated synthesis of leading and lagging strands (8). Here, we have tried to explore the evolution of the herpes simplex virus replicon by identifying functionally significant features of the replicon and examine when they emerged (Fig. 9). These considerations suggest a stepwise evolution of the Herpes simplex virus replicon starting from an assumed common ancestor to, on one hand, CFPHV, GHV1 and PsHV1, and, on the other hand, the roseoloviruses of the betaherpesviruses family HHV6a, HHV6b and HHV7. The branching resulted in different nucleotide sequence in the origins of DNA replication. This also led to variations in the putative DNA binding motif; RVKNL for most alpha herpesviruses and AYRNL or PFRSL for roseoloviruses. Further branching of the alphaherpesvirus tree depended on the acquisition of the ICP8 binding motif WPxxxGxxxFxxI/L. Finally, we observe variations in the arrangement of DNA binding sites in oriS. A large group is characterized by the ability to form a hairpin between the box III and box I sequences in oriS. In contrast, VZV has only a single essential binding site in minimal oriS (45). These changes are likely to be biologically significant since they occurred early in evolution, possibly even before herpesviruses became stringently coupled to a specific host. To gain more insight into these early events complete sequencing of more distant relatives to HSV1 ought to be encouraged (50).
References.

1. http://www.ictvonline.org/virusTaxonomy.asp
2. Fields Virology, 5th Edition. (2007) Eds.: Knipe, D.M., and Howley, P.M. Lippincot Williams & Wilkins, Philadelphia, USA
3. Lehman, I.R., and Boehmer, P.E (1999) J. Biol. Chem. 274, 28059-62
4. Aslani, A., Olsson, M., and Elias, P. (2002) J. Biol. Chem. 277, 41204-12
5. Makhov, A.M., Lee, S.S., Lehman, I.R., and Griffith, J.R. (2003) Proc. Natl. Acad. Sci. USA 100, 898-903.
6. Aslani, A., Simonsson, S., Elias, P. (2000) J. Biol. Chem. 275, 5880-5887
7. Aslani, A., Macao, B., Simonsson, S., and Elias, P. (2001) Proc. Natl. Acad. Sci. USA 98, 7194-7199.
8. Falkenberg, M., Lehman, I.R., and Elias, P. (2000) Proc. Natl. Acad. Sci. USA 97, 3896-3900.
9. Mulyaert, I., and Elias, P. (2007) J. Biol. Chem. 282, 10865-10872.
10. Strang, B.L., and Stow, N.D. (2005) J. Virol. 79, 53,587-595.
11. Schildgen, O., Gräper, S., Blumel, J., Matz, B. (2005) J. Virol. 79,7273-7278.
12. Blumel, J., and Matz, B. (1995) J. Gen. Virol. 76, 3119-3124.
13. Elias,P., O’Donnell, M.E., Mocarski, E.S., and Lehman, I.R. (1986)
14. Elias, P., Gustafsson, C.M., and Hammarsten, O. (1990) J. Biol. Chem. 265,17167-17173.
15. Weir, H.M., Calder, J.M., and Stow, N.D. (1989) Nucleic Acids Res. 17,1409-1425.
16. Simonsson, S., Samuelsson, T., and Elias, P. (1998) J. Biol. Chem. 273,2463-24639.
17. Hazuda, D.J., Perry, H.C., Naylor, A.M., and McClements, W.L. (1991) J. Biol. Chem. 266, 24621-24626.
18. Stow, N.D., Brown, G., Cross, A.M., and Abbotts, A.P. (1998) Virology 240,183-192.
19. Deb, S., and Deb, S.P. (1991) J. Virol. 65, 2829-2838.
20. Arbuckle, M.I., and Stow, N.D. (1993) J. Gen. Virol. 74, 1349-1355.
21. Martin, D.W., Munoz, R.M., Oliver, D., Subler, M.A., and Deb, S. (1994) Virology, 65, 4539-4369.
22. Krug, L.T., Inoue, N., and Pellett, P.E. (2001) J. Virol. 75, 3925-3936.
23. Dykes, C., Chan, H., Krenitsky, D.M., and Dewhurst, S. (1997) J. Gen. Virol. 78, 1125-1129.
24. Krug, L.T., Inoue, N., and Pellett, P.E. (2001) Virology 288, 145-153
25. Thureen, D.R., and Keefer, C.L. (2006) J. Virol. 80, 7863-7872.
26. Boehmer, P.E., Craigie, M.C., Stow, N.D., and Lehman, I.R. (1994) J. Biol. Chem. 269, 29329-29334.
27. Boehmer, P.E., and Lehman, I.R. (1993) Proc. Natl. Acad. Sci. USA 90,8444-8448
28. Hammarsten, O., and Elias, P. (1997) Nucleic Acids Res. 25, 1753-1760.
29. Gustafsson, C.M., Falkenberg, M., Simonsson, S., Valadi, H. and Elias, P. (1995) J. Biol. Chem. 270, 19028-19034.
30. Boehmer, P.E., and Lehman, I.R. (1993) Proc. Natl. Acad. Sci. USA 90, 8444-8448.
31. Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., Rance, M., and Skelton, N.J. Protein NMR spectroscopy, Elsevier, 2007.
32. Piotto M., Saudek V., and Sklenar V. (1992) J Biomol NMR, 132, 125-126.
33. Delaglio, F., Grzesiek S., Vuister GW., Zhu G, Pfeifer J., Bax A. (1995) J Biomol NMR 6, 277-293
34. Bartels, C., Xia, T.H., Billetter, M., Guntert, P., and Wuthrich, K. (1995) J. Biomol. NMR 6, 1-10.
35. Guntert, P. (2004) Meth. Mol. Biol. 278, 353-378.
36. Marky, L. A., and Breslauer, K. J. Biopolymers. 1987, 26, 1601-1620.
37. Herbst, L., Ene, A., Su, M., Desalle, R., and Lenz, J. (2004) Curr Biol. 14, R697-R699.
38. Perelygina, P., Zhu, L., Zurkuhlen, H., Mills, R., Borodovsky, M., and Hilliard, J.K. (2003) J. Virol. 77, 6167-6177.
39. Ohsawa, K., Black, D.H., Sato, H., Rogers, K., and Eberle, R. (2003) Arch. Virol 148, 989-997.
40. Ohsawa, K., Black, D.H., Sato, H., and Eberle, R. (2002) J. Virol. 76, 1516-1520.
41. Xodo, L. E.; Manzini, G.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. Biochemistry. 1988, 27, 6321-6326.
42. Scheffler, I. E.; Elson, E. L.; Baldwin, R. L. J. Mol. Biol. 1968, 36, 291-304.
43. Hernandez, T.R., Dutch, R.E., Lehman, I.R., Gustafsson, C.M., and Elias, P. (1991) J. Virol. 65, 1649-1652.
44. Martin, D.W., Deb, S.P., Klauer, J.S., and Deb, S. (1991) 65, 4359-4369.
45. Stow, N.D., Weir, H.M., and Stow, E.C. (1990) 177, 570-577.
46. Weir, H.M., and Stow, N.D. (1990) J. Gen. Virol. 71, 1379-1385.
47. Chen, D., Stabell, E.C., and Olivo, P.D. 69, 4515-4518.
48. Stow, N.D., and Davison, A.J. (1986) J. Gen. Virol. 67, 1613-1623.
49. Macao, B., Olsson, M., and Elias, P (2004) J. Biol. Chem. 279, 29211-29217.
50. Stacy, B.A., Wellehan, J.F.X., Foley, A.M., Coberley, S.S., Herbst, L.H., Manire, C.A., Garner, M.M., Brookings, M.G., Childress, A.L., and Jacobson, E.R. (2008) Vet. Microbiol. 126, 63-73.
Abbreviations.
The following nomenclature, Refsequences and/or accession number have been used for sequence alignments and tabulation of origins of replication:
BHV1, Bovine herpesvirus 1, NC_001847; BHV5, Bovine herpesvirus 5, NC_005261; CHV1, Cercopithecine herpesvirus 1, NC_004812; CHV2, Cercopithecine herpesvirus 2, NC_006560; CHV9, Cercopithecine herpesvirus 9, NC_002686; CHV16, Cercopithecine herpesvirus 16, NC_007653; EHV1, Equid herpesvirus 1, NC_001491; EHV4, Equid herpesvirus 4, NC_001844; GHV1, Gallid herpesvirus 1, NC_006623; GHV2, Gallid herpesvirus 2, NC_002229; GHV3, Gallid herpesvirus 3, NC_002577; HHV6A, Human herpesvirus 6A, NC_001664; HHV6B, Human herpesvirus 6B, NC_000898; HHV7, Human herpesvirus 7, NC_001716; HSV1, Human herpesvirus 1, Herpes simplex virus type 1, NC_001806; HSV2, Human herpesvirus 2, Herpes simplex virus type 2, NC_001798; VZV, Human herpesvirus 3, Varizella zoster virus, NC_001348; MeHV1, Meleagrid herpesvirus 1, NC_002641; PsHV1, Psittacid herpesvirus 1, NC_005264; SHV1, Suid herpesvirus 1, Pseudorabies virus, NC_006151; CFPHV, Chelonid fibropapilloma associated herpesvirus, AY644454.

Footnotes
This work was supported by grants from the Swedish Cancer Foundation, The Swedish Research Council and the Sahlgrenska University Hospital Łakarutbildningsavtal. We are grateful to the Swedish NMR facility for providing excellent support.
### Table 1

| Virus       | Recognition sequence in oriS | Arrangement of OBP binding sites in oriS | C-terminus of OBP (ICP8-binding motif) | DNA-binding motif |
|-------------|------------------------------|------------------------------------------|-----------------------------------------|------------------|
| HHV6a       | CGTTCTCCC                   | >AT<                                     | SVSEVGFQGHASPGVCVSS                      | AYRNL            |
|             | CGTCCACCT                   |                                          |                                         |                  |
| HHV6b       | CGTTCTCCC                   | >AT<                                     | SVSEVGFQGHASPGVCVSS                      | AYRNL            |
|             | CGTCCACCT                   |                                          |                                         |                  |
| HHV7        | CGTCCACCT                   | >AT<                                     | CLFQHADAGMCLSSGCLLRS                      | PFRSL            |
|             | GATCCCTCCT                  |                                          |                                         |                  |
| CFHV        | nd                           | nd                                       | SKTRLSPGILGHATVGVRCP                      | RMKNL            |
| PsHV1       | A   T    TCGCAC             | >AT<                                     | RTPDWMRQLGSNRPQAEQAP                      | RVKNL            |
| GHV1        | A   T    TCGCAC             | >AT<                                     | LKGRCLEDFHLMTPISSHHV                      | RVKNL            |
| CHV9        | A   T    TCGCAC             | >AT<                                     | AELAPSSWPKVMGDFFNL                        | RVKNL            |
| VZV         | G   T    TCGCAC             | >AT<                                     | SELSPGWPRINGAVNFESL                       | RVKNL            |
| GHV2        | G   T    TCGCAC             | <>AT<                                    | AELPSALWPRVEAGDFHRL                       | KYRNK            |
| GHV3        | G   T    TCGCAC             | <>AT<                                    | AEVPPACWPRVEAGDFHSL                       | KYRNK            |
| MeHV1       | G   T    TCGCAC             | <>AT<                                    | AEVPASYWPAIDGADFHR                       | KYRNK            |
| HSV1        | G   T    TCGCAC             | <>AT< (> )                               | SLPTEAWPMMQGAVNFSTL                       | RVKNL            |
| HSV2        | G   T    TCGCAC             | <>AT< (> )                               | SLPTEAWPMMQGAVNFSTL                       | RVKNL            |
| CHV1        | G   T    CCGCAC             | <>AT<> (< ) AT<>                          | SELPAEAWPMTQGALDFTTI                      | RVKNL            |
| CHV2        | G   T    CCGCAC             | <>AT<> (< ) AT<>                          | SELPAEAWPMTQGALDFTTI                      | RVKNL            |
| CHV16       | g/a T    CCGCAC             | <>AT<> (< ) AT<>                          | SELPAEAWPMTQGALDFTTI                      | RVKNL            |
| EHV1        | G   T    TCGCAC             | <>AT< (< ) AT<                           | ADLPASAWPQVYAVDFDAL                      | RVKNL            |
| EHV4        | G   T    TCGCAC             | <>AT< (< ) AT<                           | ADLPASAWPQVYAVDFDAL                      | RVKNL            |
| SHV1        | G   T    TCGCAC             | <>AT< (< ) AT<                           | SLVPAEAWPTEAGDFAL                        | RVKNL            |
| BHV1        | g/a t/c   TCGCAC             | <>AT<> (< ) AT<                          | ARVPALHWPACAGAVDFCAL                      | RVKNL            |
| BHV5        | g/a T    TCGCAC             | <>AT<> (< ) AT<                          | ARVPALHWPACAGAVDFCAL                      | RVKNL            |

Sources and nomenclature for DNA sequences have been presented under Abbreviations. Origins of DNA replication have been identified from sequence analysis. For roseoloviruses the sequences for two slightly different binding sites for OBP have been listed (31). The number and orientation of OBP binding sites in relation to an AT rich spacer sequence is schematically presented by symbols > and <. Note that since a virus often has more than one origin of replication they may exist as variants. This is indicated by symbols within parenthesis. The conserved amino acids within the ICP8 binding motif is shown in bold print.
Table 2

| HisΔOBP variant | C-terminal sequence       | ICP8 binding |
|-----------------|--------------------------|--------------|
| wt (HSV1)       | EAWPMMQGAVNFSTL          | +            |
| A (VZV)         | GSWPRINGAVNFESL          | +            |
| B               | EAAAPMMQGAVNFSTL         | -            |
| C               | EAAAMMQGAVNFSTL          | +            |
| D               | EAWPMMQAAVNFSTL          | +            |
| E               | EAWPMMQGAVNASTL          | -            |
| F               | EA                       | -            |

C-terminal sequences in HSV1 HisΔOBP created by cassette mutagenesis. Conserved aminoacids highlighted in bold print have been replaced alanines in the mutant proteins. In A the HSV1 sequence was replaced by the corresponding sequence from VZV. Binding of ICP8 to OBP was compiled from figures 5, 6 and 7.

Table 3.

| DNA hairpin | Stem base pairs / mis-matches / bases in loop | Total GC content (%) | Tm (°C) | ΔH° (kJ/mol) | ΔS° (J/(mol·K)) |
|-------------|-----------------------------------------------|----------------------|---------|--------------|-----------------|
| EHV1        | 11 / 1 / 3                                    | 45                   | 66      | 169          | 498             |
| EHV4        | 11 / 0 / 5                                    | 45                   | 69      | 335          | 979             |
| HSV1        | 12 / 1 / 3                                    | 58                   | 71      | 220          | 639             |
| BHV1        | 14 / 1 / 3                                    | 64                   | 76      | 277          | 793             |
| GHV2        | 13 / 0 / 4                                    | 69                   | 83      | 354          | 994             |
| GHV3        | 12 / 1 / 4                                    | 83                   | 84      | 398          | 1110            |

Oligonucleotides used for DNA melting experiments. Thermodynamic parameters derived from the melting experiments are tabulated.
Figure legends

Figure 1. Schematic presentation of the Herpes simplex virus replicon. Upper part. HSV1 oriS. The linear genome contains three homologous replication origins, two copies of oriS and one copy oriL, and encodes seven replication proteins. Middle part. HSV1 OBP. OBP or UL9 protein is a superfamily II DNA helicase as well as a sequence specific DNA binding protein. Here the helicase domain is represented by two connected elipsoids and the C-terminal DNA binding is drawn as a circle. The OBP binding sites in oriS are shown. The OBP dimer binds two double-stranded DNA box I oligonucleotides but only one hairpin with a single-stranded tail (48). The figure is intended to demonstrate conformational changes affecting the DNA binding domain, referred to as ΔOBP, during activation of oriS. Lower part. The DNA binding domain ΔOBP. A schematic presentation of three motifs discussed in this publication: the F553xxKYL motif required for proper folding of the DNA binding domain, the R756VKNL motif necessary for DNA binding and the W839PxxxGAXxFxxL motif involved in binding to ICP8.

Figure 2. The RVKNL motif is involved in binding to DNA. Autoradiographs of mobility shift experiments demonstrating binding of HisΔOBP from HSV1 to a duplex oligonucleotide containing the box I sequence. Also shown is a quantification of the same experiment obtained by phosphoimager analysis. The concentration of duplex oligonucleotide was 1 nM. The concentrations of proteins were: (i) RVKNL(wt) 0.027, 0.27, 2.7 and 27 nM; (ii) RVaaL 0.039, 0.39, 3.9, 39 and 3900 nM; (iii) aVKNL 0.026, 0.26, 2.6, 26 and 2600 nM.

Figure 3. Mutations of conserved amino acids in the DNA binding domain affect DNA binding. Autoradiographs of mobility shift experiments demonstrating binding of mutant forms of HisΔOBP from HSV1 to a duplex oligonucleotide containing the box I sequence. The concentration of duplex oligonucleotide was 1nM. A. The lanes represent (from left to right): no addition of protein, the mutant proteins NL/AA, F/A, PG/AA, LL/AA, V/A, N/A and wt HisΔOBP. The protein concentration was 570 nM. B. The lanes represent (from left to right): no addition of protein, PG/AA (0.057, 5.7, 57 and 570 nM), LL/AA (0.057, 5.7, 57 and 570 nM) and wt HisΔOBP (0.57 nM). C. The lanes represent (from left to right): no addition of protein, V/A (0.057, 5.7, 57 and 570 nM), N/A (0.057, 5.7, 57 and 570 nM) and wt HisΔOBP (0.57 nM). D. Quantification of results presented in B and C above. Note that the amount of complex formed by LL/AA is the sum of both complexes seen in the 9th lane in C.

Figure 4. Binding of HisΔOBP from HSV1 and Monkey B virus, CHV1, to cognate and non-cognate sites. Autoradiographs of mobility shift experiments demonstrating binding of HisΔOBP from HSV1 to a duplex oligonucleotide containing the box I sequence. Also shown are quantifications of the same experiments obtained by phosphoimager analysis. A. Binding of CHV1, (Monkey B Virus) HisΔOBP (left part of the autoradiograph) and HSV1 HisΔOBP (right part of the autoradiograph) to box I from CHV1 oriS. B.
Binding of CHV1 HisΔOBP (left part of the autoradiograph) and HSV1 HisΔOBP (right part of the autoradiograph) to box I from HSV1. The concentration of duplex oligonucleotide was 1 nM. The concentrations of CHV1 HisΔOBP were 0.041, 0.41, 4.1, 41 and 4100 nM and HSV1 HisΔOBP were 0.027, 0.27, 2.7 and 27 nM.

Figure 5. The C-terminal 13 amino acids in OBP are required for specific interaction with ICP8.
Autoradiograph of a mobility shift experiment demonstrating binding of ICP8 to a complex between HisΔOBP and a duplex oligonucleotide containing the box I sequence. The concentration of the Box I duplex oligonucleotide was 1 nM. The concentration of HSV1 HisΔOBP (wt) was 0.027 nM and of HisΔOBP (F), a mutant lacking the C-terminal 13 amino acid, was 0.027 nM. Increasing amounts of ICP8, 1.6, 16, 160, 1600 nM, were added to reactions with HisΔOBP (wt) and 1.6, 16, 160 nM to HisΔOBP (F). The C-termini for HisΔOBP (wt) and HisΔOBP (F) are shown in Table 2.

Figure 6. The C-terminal 13 amino acids from VZV OBP support a heterologous interaction with HSV-1 ICP8.
Autoradiograph of a mobility shift experiment demonstrating binding of ICP8 to a complex between HisΔOBP and a duplex oligonucleotide containing the box I sequence. The concentration of HSV-1 HisΔOBP (wt) was 0.027 nM I (second and third lanes) and 0.27 nM (fourth and fifth lanes). The concentration of HisΔOBP (A), with a C-terminus corresponding to OBP from VZV, was 0.054 nM (sixth and seventh lanes) and 0.54 nM (eight and ninth lanes). ICP8, 1500 nM, was added to reactions as indicated. The C-termini for HisΔOBP (wt) and HisΔOBP (A) are shown in Table 2.

Figure 7. Evolutionarily conserved tryptophan and phenylalanine residues are required for stable interaction between ICP8 and OBP.
Autoradiograph of a mobility shift experiment demonstrating binding of ICP8 to a complex between HisΔOBP and a duplex oligonucleotide containing the box I sequence. The concentration of HSV-1 HisΔOBP (wt) was 0.027 nM. The concentration of HisΔOBP(B) was 0.043 nM, 0.081 nM for HisΔOBP (C), 0.052 nM for HisΔOBP(D) and 0.032 nM for HisΔOBP(E). ICP8, 530 nM, was added to reactions as indicated. The C-termini for HisΔOBP (wt) and the mutant derivatives B (W839/A), C (P840/A), D (G844/A) and E (F848/A) are shown in Table 2.

Figure 8. Hairpin structures formed by box III and box I sequences in oriS.
A. NMR structure of an oligonucleotide corresponding to HSV1 box III and Box I sequences. The 29-mer oligonucleotide, 5’ GAAGTGAGAAGCGAAGCCGTTTCGCACTTC3’, representing HSV1 is shown in a B-DNA conformation (dark blue) except for the mismatch and loop nucleotides 7,14-16 and 23 (light blue). Distance restraints that are consistent with B-DNA are shown in yellow, while restraints violated by more than 1.0 Å are coloured as orange. B. DNA melting experiments performed using oligonucleotides derived from EHV1,
EHV4, HSV1, BHV1, GHV1 and GHV3 each represented by a graph in order left to right (see Experimental procedures for details). Absorption was monitored at 260 nm. Melting temperatures and thermodynamic parameters are given in table 3.

**Figure 9. Evolutionary tree of the alphaherpes virus replicon.**
Alignments of currently available OBP sequences were made using Clustal-W. The result is displayed using Treeview. The branches are marked to illustrate the stepwise acquisition of the alphaherpesvirus DNA recognition motif, the ICP8 binding motif and the BoxIII-Box I palindrome. Thin lines denote the DNA binding domain from roseoloviruses. Thick lines denotes a DNA binding domain from alpha herpesviruses with the RVKNL motif indicative of binding to TTCGCAC. Black lines show the presence of the ICP8 binding motif. Bold letters shows the presence of the Box III-box I palindrome in oriS. Italics denote the presence of only one binding site for OBP. A phylogenetic tree of all herpesviruses as predicted by partial DNA polymerase sequences has recently been published and could be consulted for a broader perspective (50).
Figure 1

BoxIII    BoxI    AT    BoxII

GTGAGAA  TTCGCAC  GTGCGAG

FxxKYL

WPxxxGAxxFxxL

RVKNL
Figure 2

Protein concentration (pM) vs. Complex formation (%)

- RVKNL (wt)
- RVaL
- aVKNL

Complex formation (%)

Protein concentration (pM)
Figure 3

A

| Complex | Free DNA |
|---------|----------|
| HisΔOBP | - | NL/AA | PG/AA | LL/AA | V/A | N/A | wt |

B

| Complex | Free DNA |
|---------|----------|
| HisΔOBP | - | PG/AA | LL/AA | wt |

C

| Complex | Free DNA |
|---------|----------|
| HisΔOBP | - | V/A | N/A | wt |

D

![Graph showing complex formation vs. protein concentration](image-url)
Figure 4

A

HisΔOBP

CHVΔOBP HSV1ΔOBP

Complex

Free DNA

10 100 1000 10000

50

GTCCGCAC

Complex formation (%)

Protein concentration (pM)

B

HisΔOBP

CHVΔOBP HSV1ΔOBP

Complex

Free DNA

10 100 1000 10000

50

GTTCGCAC

Complex formation (%)

Protein concentration (pM)
Figure 5

ICP8
ΔOBP
DNA/ΔOBP/ICP8 complex
DNA/ΔOBP complex
Free DNA
Figure 6

ICP8  
\(\Delta\)OBP  
DNA/\(\Delta\)OBP/ICP8 complex  
DNA/\(\Delta\)OBP complex  
Free DNA
Figure 7

ICP8
ΔOBP
DNA/ΔOBP/ICP8 complex
DNA/ΔOBP complex
Free DNA

- + - + - + - + - +
wt B C D E
Figure 8b
Stepwise evolution of the herpes simplex virus origin binding protein and origin of replication
Monica Olsson, Ka-Wei Tang, Cecilia Persson, L. Marcus Wilhelmsson, Martin Billeter and Per Elias

J. Biol. Chem. published online April 7, 2009

Access the most updated version of this article at doi: 10.1074/jbc.M807551200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2009/04/09/M807551200.DC1