© 1997 Oxford University Press Nucleic Acids Research, 1997, Vol. 25, No. 9 **1753–1760**

Herpes simplex virus: selection of origins of DNA replication

Ola Hammarsten and Per Elias*

Department of Medical Biochemistry, University of Göteborg, Medicinaregatan 9A, S-413 90 Göteborg, Sweden

Received January 17, 1997; Revised and Accepted March 18, 1997

ABSTRACT

A selection procedure was devised to study the role of *cis*-acting sequences at origins of DNA replication. Two regions in Herpes simplex virus oriS were examined: an AT-rich spacer sequence and a putative binding site, box III, for the origin binding protein. Plasmid libraries were generated using oligonucleotides with locally random sequences. The library, amplified in *Escherichia coli*, was used to transfect BHK cells followed by superinfection with HSV-1. Replicated plasmids resistant to *Dpn*I cleavage were amplified in *E.coli*. The selection scheme was repeated. Plasmids were isolated at different stages of the procedure and their replication efficiency was determined. Efficiently replicating plasmids had a high AT content in the spacer sequence as well as a low helical stability of this region. In contrast, this was not seen using the box III library. We also noted that the wild type sequence invariably dominated the library after five rounds of selection. These plasmids arose from recombination between plasmids and viral DNA. Our results imply that a large group of sequences can mechanistically serve as origins of DNA replication. In a viral system, however, where the initiation process might be rate-limiting, this potentially large group of sequences would always converge towards the most efficient replicator.

INTRODUCTION

DNA synthesis is initiated on cellular and viral chromosomes at specialized regions termed origins of DNA replication. Several *cis*-acting sequences with discrete functions are generally required in order to support efficient DNA synthesis from a unique origin. Studies have so far indicated that the essential sequences serve to bind initiator proteins and to facilitate local unwinding of the DNA helix (1). Additional functions for *cis*-acting sequences might be to regulate the formation of initiation complexes or serve as start sites for primer synthesis.

We have focused our interest on the replication of Herpes simplex virus type I. Two highly homologous viral origins of replication, oriS and oriL, are found in the HSV-1 genome. The minimal HSV-1 oriS consists of ~90 bp. It has, however, been noted that sequence elements outside the minimal origin of DNA replication may enhance the replication efficiency ~100-fold (2).

OriS contains two strong binding sites, boxes I and II, for the viral origin binding protein, OBP (3–5), a DNA helicase encoded by the UL9 gene (6–8). A third closely homologous site, box III, is also found in oriS and oriL. It might also be able to bind OBP (5). Boxes I and II are separated by a spacer sequence containing a stretch of 18 alternating AT nucleotides. Mutations in boxes I, II and III lower the replication efficiency of oriS (9). It has also been demonstrated that alterations of the length and sequence of the spacer dramatically influence the replication efficiency (10,11).

The activation of oriS is probably initiated by the cooperative binding of two OBP dimers (4,5,10). This interaction is strongly dependent on the distance between the sites and seems to cause a bending and distortion of the spacer sequence (10,12). A specific interaction between OBP and the single-strand DNA binding protein, ICP8, may serve to position ICP8 at the spacer sequence (13). Observations using the electron microscope have shown that active complexes between OBP and oriS extrude the spacer sequence in a profoundly altered conformation (14). This spacer sequence might therefore resemble the DNA unwinding elements, DUEs, described in other replication origins (15–21).

We are currently interested in determining how the specific protein–DNA interactions and cooperative protein–protein interactions are used to initiate a local melting of the DNA helix in HSV-1 oriS. We therefore wanted to identify the group of sequences that could succesfully support origin dependent DNA synthesis. If a sufficiently large group of sequences could be retrieved, one could examine in a systematic way how the nucleotide sequence and physical properties of these sequences would affect initiation of DNA replication. In order to collect a group of viral replicators we have devised a method for selecting replicating oriS-containing plasmids. Briefly, a plasmid library of high complexity was generated using oligonucleotides with locally random sequences. The library was transfected into BHK cells and superinfected with HSV-1. Replicated plasmid DNA was retrieved following *Dpn*I digestion of isolated DNA. The replicated plasmids were amplified in *Escherichia coli* and the selection scheme was repeated.

Individual plasmids were then isolated at different stages during the procedure. They were sequenced and their replication efficiencies were determined. The results obtained from these studies have been incorporated in a model for the activation of HSV-1 oriS. They might also be used to evaluate attempts to predict the location of origins of DNA replication on chromosomes.

*To whom correspondence should be addressed. Tel: +46 31 773 3486; Fax: +46 31 416 108; Email: per.elias@medkem.gu.se

MATERIALS AND METHODS

Cells and viruses

BHK cells were propagated as described (22). Virus from the strain Glasgow strain 17 syn+ was prepared as before (22). Virus stocks containing 10^8 plaque forming units (p.f.u.)/ml were stored in aliquots at –70°C until use. Transfection of cells was carried out using liposomes (22).

Plasmids

All constructions were made using pTZ19 (Pharmacia) as the parental vector. The following plasmids were made. p35/36 contains the complementary oligonucleotides OH35 and OH36 inserted at the *Xba*I and *Hin*dIII sites in pTZ19. The plasmid was used for construction of the libraries as described below. pWT was isolated during the selection experiments. It is derived from the library oligonucleotide PE161 and has acquired the wild type sequence probably by recombination and gene conversion since two independent isolates of this plasmids were obtained during the course of these investigations. pControl was also isolated during the selection experiments. It contains HSV-1 oriS and neighbouring sequences from the HSV-1 genome (nucleotides 131740–138785 in the X14112 entry in the EMBL database). The plasmid was used as an internal control in plasmid replication assays.

pOS822 was a generous gift from P.Schaffer (Harvard University) and contains a 822 bp *Bam*HI fragment containing HSV-1 oriS (2). pCG5 contains a minimal version of HSV-1 oriS (9). p160, p161 and p49 designate plasmid libraries derived from the oligonucleotides PE160, PE161 and OH40.

Unless otherwise stated, mini preparations of plasmids were performed using Genomed Jetprep following the protocol provided by the manufacturer. Plasmids were sequenced using the thermo sequenase fluorescent labelled primer cycle sequencing kit (Amersham RPN 2436) supplemented with 1.6 µM of the appropriate dideoxynucleotide, 15 000 c.p.m. (3000 Ci/mmol) ^{32}P-labelled PE68 sequencing primer and 0.025 µg plasmid DNA per reaction. The samples were incubated in a thermocycler. After the sequencing reaction they were heat denatured in formaldehyde as suggested by the suppliers and subsequently run on an 8% denaturing sequencing gel.

Oligonucleotides

The sequences of the oligonucleotides used were as follows:

OH35 5′-AGCTAGATCTGAATTCCGGGCACGGCGCCAGT-3′
OH36 5′-CTAGACTGGCGCCGTGCCCGGAATTCGAATCT-3′
OH37 5′-GGGATCCTCTAGACTGGCGC-3′
PE68 5′-CCAGGGTTTTCCCAGTCACGAC-3′
PE160 5′-CCCCCGGGA TCCGCCCGGG TAAAAGAAGT GAGAACGCGA AGCGTTCGCA CTTCGTCCCN NNNNNNNNNN NNNNNNNGGG CGAAG-TGCGA GCACTGGCGC CGTGCCCGGA ATTCAGATCT-3′
PE161 5′-GGATCCAAGC TTGCCCGGGT AAAAGAAGTG AGAACGCGAC TTCGTCCCNN NNNNNNNNNN NNNNNNGGGC GAAGTGCGAG CACTG-GCGCC GTGCCCGGAA TTCAGATCT-3′
OH40 5′-GGGATCCCTC TAGAGCCCGG GTANNNNNNN NNNNNNNNNN AAGCGTTCGC ACTTCGTCCC AATATATATA TATTATTAGG GCGAAGTG-CG AGCACTGGCG CCGTGCCC GGAATTCAGA TCT-3′

PE160 and PE161 were from Oligonucleotides etc; OH35, OH36 and OH40 were purchased from Scandinavian Gene Systems. PE160 contains the minimal OriS including a random sequence of 18 nt with an average of 50% A/T corresponding to the spacer sequence (nt 45–62 in Fig. 1). In PE161 the random sequence has an average of 80% A/T in the spacer sequence. OH40 contains a random sequence of 17 nt with an average of 50% A/T at box III in OriS (nt 7–24 in Fig. 1)

Oligonucleotide OH37 was used to create a double-stranded *Xba*I site in the single-stranded version of pOH35/36 (see below).

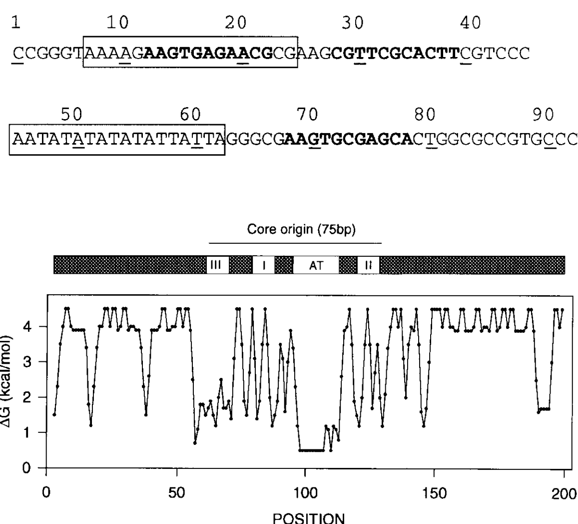

Figure 1. HSV-1 oriS. The sequence of a minimal version of HSV-1 oriS and a schematic presentation of the *cis*-acting elements addressed in this study. The binding sites for the HSV-1 origin binding protein, boxes I, II and III, are shown in bold print. They are defined as in refs 4 and 5. The spacer sequence is defined as in ref. 10. The underlined nucleotides are numbered. The p40 library contained a random sequence at the positions 7–24 corresponding to box III. The p160 and p161 libraries contained random sequences at positions 45–62 corresponding to the putative DNA unwinding element. Below is a calculation of the helix stability of oriS performed as described in ref. 25. The parameters used were 10 mM monovalent ions and 37°C. The window was 2 bp.

Plasmid libraries

The method is presented schematically in Figure 2a. Single-stranded DNA from pOH35/36 was prepared using M13K07 superinfection of TG1 cells following the protocol supplied by the manufacturer (Promega). ssDNA was purified using standard procedures (23). OH37 (0.4 nmol) was annealed to 50 µg of pOH35/36 ssDNA in 100 µl of restriction enzyme buffer A (Boehringer). The partially double stranded DNA was cleaved with 50 U *Xba*I, purified by phenol extraction and precipitated with ethanol in the presence of 40 µg glycogen. Phosphorylated PE160, PE161 or OH40 (13 pmol), all containing random sequences at defined positions in oriS, was annealed to 5 µg of linearized single-stranded p35/36 in 50 µl water and heated to 95°C for 3 min. The reaction was allowed to cool to 65°C and incubated another 30 min. An aliquot of 50 µl of a prewarmed buffer containing Vent buffer (Promega), 4 U Vent DNA polymerase and 0.4 mM dNTPs was added and the temperature was increased to 70°C . The DNA synthesis reaction was allowed to proceed for 50 min. Five separate reactions were pooled and DNA was purified by phenol extraction and precipated with ethanol. The DNA pellet was dissolved in 100 µl of water and applied to a 1% agarose gel. DNA corresponding in size to linear plasmid molecules was cut out from the gel. DNA was recovered from the gel slice by electroelution, extracted by phenol and precipitated with ethanol. The double-stranded linear DNA, ~25 µg,

was circularized using 90 Weiss units of T4 DNA ligase in 50 µl of T4 DNA ligase buffer (Promega) at room temperature overnight. The reaction mixture was extracted by phenol/chloroform and precipitated with ethanol. DNA was dissolved in 30 µl water. An aliquot of 3 µl circularized plasmid DNA was mixed with 30 µl competent TOP10F′ *E.coli* cells (24). The mixture was subjected to electroporation. The bacteria from 10 transformations were pooled and allowed to recover for 1 h at 37°C in SOC medium (24). They were subsequently spread out on 10 LB-plates supplemented with 125 µg/ml ampicillin. The plates was incubated at 37°C overnight. The number of bacterial colonies formed was determined separately from serial dilutions. Bacterial colonies were recovered from the plates and plasmid DNA was prepared using Wizard midiprep as described by the manufacturer (Promega). The complexity of the library was determined by multiplying the fraction of plasmids containing full length inserts with the total number of bacterial colonies obtained. The frequency of plasmids with full length inserts, which varied between 30 and 50%, was determined by cleavage with *Eco*RI of plasmid DNA prepared from 20 randomly picked bacterial colonies. In experiments starting with a reaction mixture of 10 µg DNA we obtained 2×10^7 full length plasmids derived from p160, 1×10^6 from p161 and 1×10^7 from p40.

Plasmid replication assays

The replication efficiency was determined by monitoring transient replication of a plasmid containing HSV-1 oriS in BHK cells superinfected with HSV-1. The assay was performed essentially as described (22). In each experiment, 0.5 µg of plasmid DNA from a minipreparation was mixed with 0.03 µg pControl. Superinfection was with 10 p.f.u./cell. The transformed and superinfected cells were incubated 18 h at 37°C. Total cellular DNA was isolated and subjected to Southern blotting of agarose gels as described (11). Replicated plasmid DNA was measured by phosphorimager analysis. In each experiment duplicate samples of pWT were included as positive controls. The replication efficiency was determined by comparing the amount of replicated library plasmid with the amount of replicated pWT. Moderate variations in transfection efficiencies were corrected by comparison with the amount of replicated pControl in individual samples.

Selection procedures

An aliquot of 0.5 µg of the amplified plasmid library was transfected into 20 multi well monolayers containing 4.5×10^5 BHK cells . The cells were infected with 10 p.f.u. HSV-1 per cell. Total cellular DNA was isolated at 18 h post-infection as described (22). DNA (~40 µg) was cleaved overnight with 100 U *Dpn*I and 450 U *Bgl*II in 200 µl cleavage buffer (33 nM Tris-acetate pH 7.9, 10 mM magnesium acetate, 66 mM potassium acetate, 0.5 mM dithiothreitol, 0.05 mg/ml bovine serum albumin). DNA was then run on 1% agarose gels and DNA corresponding in size to linear monomeric plasmid DNA was recovered by electroelution. It was extracted by phenol/chloroform and precipitated with ethanol. The DNA was dissolved in 300 µl ligase buffer (Promega) and mixed with 6 Weiss units of T4 DNA ligase. Incubation was overnight at room temperature. Recircularized plasmid DNA was recovered after phenol/chloroform extraction and ethanol precipitation. The resulting DNA pellet was dissolved in 15 µl water. Plasmid DNA (8 µl; ~10 µg) was mixed

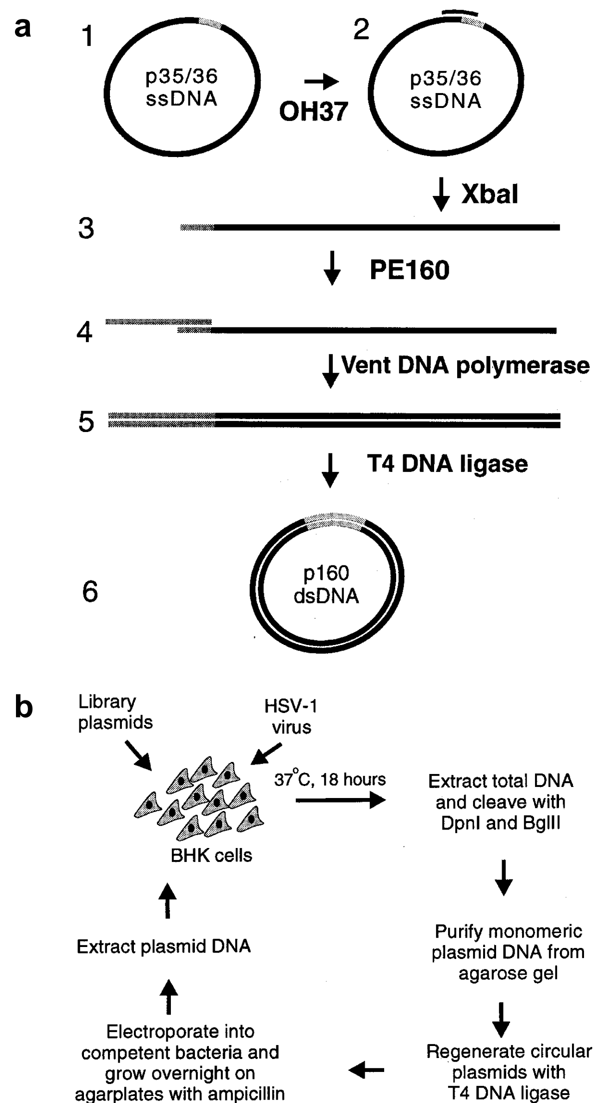

Figure 2. A protocol for selecting replicating plasmids from a library based on HSV-1 oriS. (**a**) Construction of the plasmid library as described in Materials and Methods. The libraries were p160 with 50% AT in the spacer sequence, p161 with 80% AT in the spacer sequence and p40 with 50% AT in box III. (**b**) The selection procedure.

with 60 µl of competent Top10F′ *E.coli*. The cells were electroporated and grown on plates as described above. Plasmid DNA was prepared from bacteria recovered from the plates using the Wizard midi kit (Promega) when the number of bacterial colonies obtained exceeded 5×10^5. If fewer colonies were obtained this step was repeated. The selection protocol never proceeded with $<5 \times 10^5$ colonies. The plasmid preparation was used to transfect a new set of 20 monolayers of BHK cells and the selection procedure was repeated.

RESULTS

Plasmid libraries and selection strategy

Three different plasmid libraries were generated for this study. In two of these the AT-rich spacer sequence of HSV-1 oriS was replaced with a random sequence containing 80 or 50% adenines

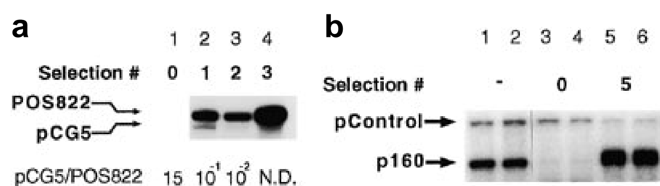

Figure 3. Efficiently replicating plasmids can be isolated by repeated selections. (**a**) Competition between the plasmids pCG5 and pOS822. Before transfection the plasmids were mixed in a molar ratio of 15:1 (selection# 0). The plasmid mixture was transfected into BHK cells followed by superinfection with HSV-1. Replicated DNA was measured on autoradiographs of Southern blots using the Phosphorimager as described in Materials and Methods. (**b**) Selection of replicating plasmids from a library. Before transfection 0.03µg pControl was mixed either with 0.5 µg pWT or 0.5 µg of the plasmid library p160 amplified once in *E.coli*. The results from two transfection experiments carried out in parallell are shown. The first two lanes show replicated DNA from experiments when pControl and pWT were mixed before transfection. The third and fourth lanes show replicated DNA from an experiment in which pControl was mixed with the starting plasmid library p160. The fifth and sixth lanes show replicated DNA in which pControl was mixed with plasmid DNA from the fifth selection cycle.

and thymines (Fig. 1). They were designated p160 and p161 respectively (Fig. 2). The third library, p40, contained a random sequence in box III of oriS (Fig. 1). This region of oriS contains a putative low-affinity binding site for the HSV-1 origin binding protein, OBP (4,5).

A method using annealing of a long oligonucleotide to a linear single-stranded vector was designed in order to obtain a library of high sequence complexity and a high proportion of full-length inserts (Fig. 2).

The selection scheme was designed to avoid disproportionate amplification of plasmids in *E.coli* and to avoid contamination of the libraries with plasmids arising from recombination with cellular or viral DNA. The structure of two plasmids isolated during these investigations, pControl and pWT, illustrate the latter case. pControl contains a 7 kb oriS fragment from the viral genome. pWT is homologous to p161 but has the complete wild type sequence of HSV-1 oriS. The probability that this plasmid would exist in the library used for the selection procedure is estimated to be <0.1%. Since it was isolated in two independent experiments we consider it likely that it occurred by homologous recombination and gene conversion.

Efficient selection of replicating plasmids

In order to verify the efficiency of the selection protocol we performed a control experiment where two different plasmids were mixed at a known ratio before the first transfection took place. The plasmid pOS822, containing oriS plus adjacent regions from the viral chromosome, has been estimated to replicate between 10 and 100 times better than a plasmid, pCG5, containing the minimal oriS (2,9). In this case the plasmids were mixed at a molar ratio of 1:15 before the first transfection. The ratio of the replicated plasmids were measured after each step in a consecutive series of three selection cycles. The results demonstrate the ratio of pCG5 to pOS822 was reduced 1500-fold after two selection cycles. In fact, pCG5 was not detectable after three cycles (Fig. 3a).

We also established that the plasmid libraries behaved in a similar way. In this instance 0.5 µg of the plasmid library p160 was mixed with 0.03 µg of a plasmids containing a large insert with oriS, pControl. An identical experiment was also performed using the plasmid pWT, which was isolated from the p160 library. pWT contains the wild type oriS sequence. Our results demonstrate that we can readily detect replication of pWT under these experimental conditions (Fig. 3b, lanes 1 and 2). In contrast, the library p160 replicated very poorly in the first transfection experiment (Fig. 3b, lanes 3 and 4). Our results further demonstrated that a library obtained after five complete selection cycles replicated very succesfully (Fig. 3b, lanes 5 and 6).

In Table 1 we have summarized the results from the final step in one complete selection experiment. We retrieved plasmid DNA giving rise to 10^6 colonies after five selection cycles. Plasmids from 20 randomly picked colonies were analyzed by DNA sequencing and analyzed in transient replication assays. We found that 25% of these plasmids had a wild type sequence. We also found more than one copy of several replicating non-wild type plasmids. Together these constituted 55% of the plasmids analyzed. Two replicating plasmids, 10%, were found in only one copy. These results indicate that repeated selection cycles dramatically reduced the sequence complexity of the library. We also found that wild type sequences occurred most frequently in the library. Inasmuch as the probability for the wild type sequence to occur in the starting library is <0.1% we suggest that these sequences were introduced by homologous recombination and gene conversion. Finally, in this particular experiment we altogether identified five different replicating non-wild type sequences whose properties were further analysed as described below.

Table 1. Plasmids obtained after five selection cycles

Properties of the selected plasmids	Number of colonies	Frequency (%)	Number of analyzed plasmids
All plasmids	10^6	100	20/20
Wt plasmids[a]	2.5×10^5	25	5/20
Non-wt plasmids[b]	7.5×10^5	75	15/20
Redundant plasmids[c]	5.5×10^5	55	11/20
Unique plasmids[d]	10^5	10	2/20
High replication efficiency[e]	7.5×10^5	75	15/20
Low replication efficiency[f]	2.5×10^5	25	5/20

Twenty bacterial colonies obtained after five selection cycles using the p160 library were picked at random. They were sequenced and tested for replication efficiency as described in Materials and Methods.
[a]Plasmids containing wild type oriS.
[b]Plasmids with a non-wild type spacer sequence.
[c]Plasmids found more than once.
[d]Plasmids found only once.
[e]Replication efficiency >50% of pWT.
[f]Replication efficiency <50% of pWT.

Properties of selected plasmids

We isolated individual plasmids from different stages during the selection scheme (Fig. 4). As mentioned before, the appearance of plasmids containing the wild type oriS sequence and plasmids with alterations in the length of the insert were discarded. Similarily, plasmids that did not have the proper restriction sites were also ignored. The remaining plasmids were sequenced. We examined

120 bp covering the entire oriS from 150 different plasmid isolates. In this way we identified 53 unique plasmids (Tables 2, 3 and 4). The relative replication efficiencies were determined using 0.5 µg of the plasmid mixed with 0.03 µg of pControl in the transient replication assay (Fig. 4). We also included duplicate samples of pWT mixed with pControl (Fig. 4). The replication efficiency was calculated by comparing the amount of replicated DNA for each plasmid to the amount of replicated pWT DNA. The amount of replicated pControl was used to normalize the amount of replicated DNA before the calculations were made.

Table 2. Plasmids with spacer sequences that inefficiently support DNA replicatoin

Sequence of the spacer	Replication efficiency (%)	A+T (%)	Origin	Selection cycle	Helix stability ΔG (kcal/mol)
TGTTCTAAACCTACTAAT	6	72	p160	4	12.6
AATATACCGTCATCTATA	0.4	72	p160	4	12.8
CTCTGCGGTTTGTAGATT	2	56	p160	4	17.8
ACGTTCTTGTATCAGGAA	0.2	61	p160	4	16.1
AGACTACTCTTTGCTGGG	3	50	p160	1	17.4
TTCAAACCTGCGCCGATC	5.3	44	p160	1	23.2
ACATATTTATTTAATTCA	6	89	p160	1	11.8
ACCACGTGCCAACAACCT	6	44	p160	1	21.6
ATTGAAAAACCGTGCTAA	5	67	p160	1	18.1
AATTTATGTTTTAAAATA	6	94	p161	1	11.6
ATTTCATATAACATTTCT	7	83	p161	1	12.4
AATTTTAATACCTTTTTT	4	88	p161	1	13.1
AAAATTATGATCAGAATT	6	83	p161	1	13.6
TATATAAATTTTTATAAC	6	94	p161	1	9.1
AATACCTATTTATTTCAC	4	78	p161	1	12.3
TAGATTACATTTCGTATT	5	78	p161	1	13.0
TAATAATAAATAATGATT	7	94	p161	1	10.3
TTAATAAGGAAAGAAATA	5	83	p161	1	12.8
TTTTTAAAACTAATTTCC	2	83	p161	0	14.2
CTATTAATAATTGCGACA	3	72	p161	0	14.9
TTAGGTTCTGACGATTTT	4.3	67	p161	0	16.4
CGTTCGATCCGCCATTAT	0.3	50	p160	3	21.6
GTCATTTTTATTTGGTAC	2	72	p161	3	14.1
AATATTAAATTGCAAAAT	2	89	p161	3	13.8
TTCTACAATTTCCGTTTA	0.2	72	p161	3	15.8

The plasmids were retrieved from the selection cycle indicated in the table. The replication efficiency was measured as described in Materials and Methods. Helix stability was calculated according to ref. 25 assuming 10 mM monovalent ions and a temperature of 37°C.

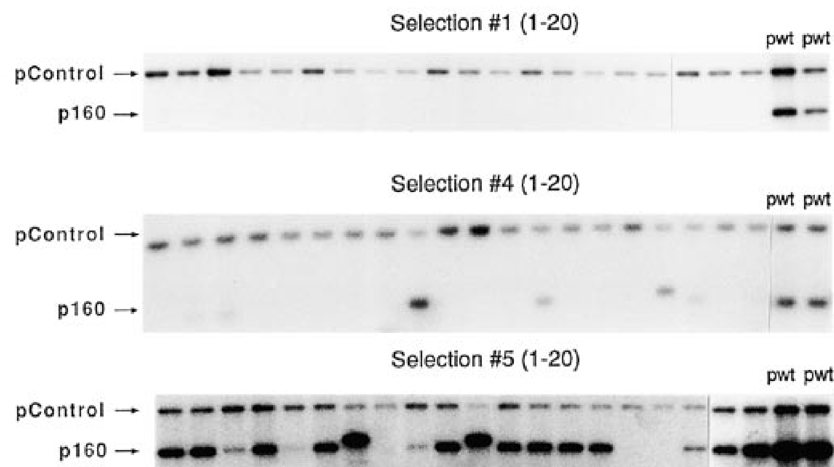

Figure 4. Selection of efficiently replicating plasmids. BHK cells were transfected with 0.5 µg of the plasmid library amplified in *E.coli*. Following each transfection experiment, individual plasmids were retrieved after recircularization and transformation of *E.coli*. The plasmids were sequenced. The replication efficiencies of a random samples of these plasmids were examined in transient replication experiments using pControl is as an internal standard as described in Materials and Methods. In the autoradiograph p160 indicates the position of pWT as well as plasmids derived from the library. In the two rightmost lanes pControl was transfected with pWT in order to normalize the results as described in Materials and Methods.

Table 3. Plasmids with spacer sequences that efficiently support DNA replicatoin

Sequence of the spacer	Replication efficiency (%)	A+T%	Origin	Selection cycle ΔG(kcal/mol)	Helix stability
AATATATATATATTATTA	100	100	Wt	–	6.4
AAAAACATATAAAATAAT	114–127	94	p160	5	11.1
ATAACTAAAAGATTCGGA	115	72	p161	5	14.5
CTATCTATAGAAACAAAT	73–93	78	p161	5	11.2
CTTTCTTTTCCCAATTTT	76–107	72	p160	5	17.1
CTTATTTTAACACTTAAA	78	83	p161	5	11.9
TTTTTCTAAAACTTCTGA	77	77	p160	5	14.3
TTACCATATTTACCCGA	53–71	65	p160	5	15.5
TTTGTAAAATAGTCTTTA	49–59	83	p160	5	12.0
TATAATATACTAATATTA	32	94	p161	5	6.9
GAAACTCTAGCAAAACGT	27	61	p161	5	16.3
TTTCCTACATCCGGACTT	26	56	p160	4	17.7
TAAGTTTTCGAGGATCTA	19	67	p160	4	15.2
AAAACTTTAATAAAATTA	11–28	94	p161	5	11.3
TTTCCATACTTCTCAACG	16	61	p160	5	17.0
ACTATCGCTTCTGTATAT	10–16	67	p160	5	13.7
GTATATATACTTTCTGAA	10	78	p160	5	10.4
TTCCCAAGAAGTGAATT	11	67	p160	5	17.5

Results were obtained as described for Table 2.

Table 4. Plasmids selected from the p40 library

Sequence of the box III region	Replication efficiency (%)	A+T (%)	Origin	Selection cycle	Helix stability ΔG(kcal/mol)
AAAGAAGTGAGAACGCG	100	53	wt	–	18.2
CCGTGTTGGTGGTGGCG	1	30	p40	0	23.9
TCGATGTCTGGTACACC	2	47	p40	0	21.1
GCCCTGGTAATGGGCTT	1.5	41	p40	0	21.1
TTGCGGTAGTTAATAAT	9	70	p40	0	14.6
ACTGACTGTCGTGGGTT	11	47	p40	0	17.9
TCTAGTCTTATACGTCC	20	58	p40	5	13.0
ATCGCGGGGGGCTGGTT	20	29	p40	5	25.9
TACCCGGGGATCCTCTA	25	41	p40	5	20.0
GCAGTCTTCGTAGTGGT	26	47	p40	5	16.8

Results were obtained as described for Table 2.

The results from these experiments are summarized in Tables 2, 3 and 4. In each table the DNA sequence from the library is presented together with the wild type sequence. The replication efficiencies are presented. The percentage of AT base pairs for each sequence is given. We also calculated the thermodynamic properties of the sequence (25).

Plasmids derived from the libraries containing a random sequence replacing the AT-rich spacer, p160 and p161, are presented in Tables 2 and 3. Plasmids that come from the box III library, p40, are shown in Table 4.

For a further analysis we selected plasmids shown in Table 2 and 3 that originated from the p160 library. The starting library had an average content of 50% adenines and thymines in the spacer sequences. The replication efficiencies of these plasmids and of a collection of unselected plasmids that did not replicate were examined. We noted a correlation between an increasing replication efficiency and an increasing AT content of the spacer sequence (Table 5).

Table 5. The replication efficiency and the AT-content of spacer sequences in plasmids derived from the p160 library

Replication efficiency	% AT	Number of sequences
50–100	78.2	5
10–25	65.8	5
2–6	58.8	5
–[a]	56	10

All replicating plasmids from the p160 library were divided into three categories according to their replication efficiencies as shown in Tables 2 and 3.
[a]The average A+T content in the library was determined by sequencing 10 plasmids chosen at random from the library at the outset of the experiment.

A graphical presentation of the properties of the plasmids originating from the p160 library is also shown (Fig. 5). A linear regression analysis indicated that increasing replication efficiency correlated with increasing AT content of the spacer sequence (Fig. 5a). In a similar way, we noted that efficiently replicating

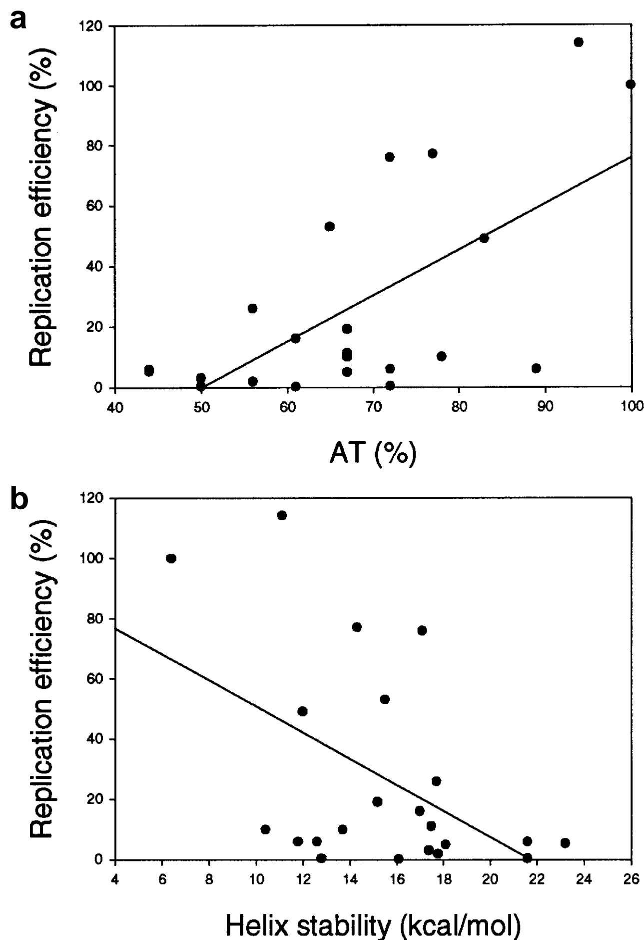

Figure 5. The replication efficiency of selected plasmids correlates with the AT content and helix stability of the spacer sequence. A graphical presentation of the results for plasmids derived only from the p160 library as shown in Table 1. The linear regression analysis was performed using Sigmaplot version 3.0 Jandel Corporation. (**a**) AT content (%) of the spacer sequence versus replication efficiency of the plasmid. The regression line $y = -75.83 + 1.517x$ gives the correlation coefficient $r = 0.66$. (**b**) Calculated helical stability of the spacer sequence (kcal/mol) versus replication efficiency of the plasmid. Helical stability was calculated as for Table 2. The regression line $y = 94.03 - 4.322x$ gives the correlation coefficient $r = 0.47$.

plasmids tended to have a low helix stability (Fig. 5b). It is interesting to note, however, that there are exceptions to these rules. Some plasmids with a high AT content in the spacer replicated poorly. It was also possible to identify plasmids with a predicted low helical stability of the spacer sequence that replicated inefficiently.

The results from selection experiments using the p40 library were different (Table 4). First, the number of plasmids directly derived from the p40 library found after five selection cycles was low and they only had a moderately increased replication efficiency compared with their unselected counterparts. Second, we could not see a correlation between the AT content of the box III sequence and the replication efficiency of the plasmid.

To summarize, our results support the view that the spacer sequence performs the role of a DNA unwinding element whereas box III has a different role. Efficiently replicating plasmids tended to have a spacer sequence with a high AT content and a low predicted helical stability. Several exceptions were, however, found. We believe that the latter observation indicates that initiation of DNA synthesis might require sequence or structure specific interactions between the replication machinery and the DNA unwinding element.

DISCUSSION

We have used a selection method to study the role of two *cis*-acting elements in the HSV-1 origin of DNA replication: an AT-rich spacer sequence and a weak putative binding site for the origin binding protein referred to as box III (4,5,10). We noted that the spacer sequence had properties typical for a DNA unwinding element (17,25). A high replication efficiency correlated well with a high AT content and also with a low predicted helix stability. Exceptions were, however, found. Several plasmids with either a high AT content or low helix stability of the spacer sequence were replicating inefficiently. We believe that this indicates that specific interactions has to occur between components of the replication machinery and the spacer sequence. These interactions might be sequence specific or, perhaps more likely, structure-specific. We have previously noted that the single-stranded DNA binding protein ICP8 could be positioned in the immediate vicinity of the AT-rich spacer sequence by the HSV-1 origin binding protein (13). It is tempting, therefore, to speculate that an essential step during the activation of oriS involves the destabilization of defined base pairs in the spacer sequence followed by the binding of ICP8 precisely at this position.

The major obstacle during these investigations was the continuous contamination of the libraries by plasmids that had recombined with the part of the HSV-1 genome that contained oriS. We tried to eliminate these sequences in two ways. Only plasmids of the right size and with the correct restriction sites were isolated from the cells and further propagated. We believe that these problems also illustrate an important aspect of replication of HSV-1. It seems as if homologous recombination and gene conversion are combined with an efficient selection in order to maintain the integrity of the viral genome.

The role of well-defined *cis*-acting sequences during the initiation of DNA synthesis has been well established for prokaryotic organisms, phages and viruses (1). The same seems to be true also for *Saccharomyces cerevisiae* (1). In the latter case an ARS consensus sequence is required for an interaction with the origin recognition complex, ORC, as well as for supporting the initiation of DNA synthesis (26,27). Studies of higher eukaryotes present a different picture. DNA replication in extracts from *Xenopus* oocytes appears to be sequence-independent but still under strict control (28). Attempts to identify conserved *cis*-acting elements serving as origins of DNA replication on mammalian chromosomes have been unsuccesful (29,30). Does this observation imply that initiation of DNA synthesis is fundamentally different in higher eukaryotes? We believe that the results presented in this communication partially address this question. We have seen that a very large number of sequences can mechanistically serve as origins of DNA replication also in a viral system. Our results indicate that after the completion of five consecutive selection cycles a maximum of ~10^5 different replicating sequences might exist in the population (Table 1). The replication efficiencies of these sequences ranged between 50 and 100% of the wild type oriS sequence (Table 3). At the same time, the wild type oriS and efficiently replicating plasmids occurred much more frequently than other sequences. We suggest that initiation of DNA synthesis

is a rate limiting step during the progation of a virus. This will inevitably result in a strong selection for the most efficient replicator sequence. The situation will be different on mammalian chromosomes. In this instance a coordinated use of multiple start sites for DNA synthesis is of paramount importance. It could perhaps even be argued that overreplication due to extremely efficient origins of DNA replication must be counterselected. The sequence requirement for mammalian origins of DNA replication could then be restricted to play a role only in the binding of the origin recognition complex to DNA (26). The structural determinants for this interaction still remains obscure.

ACKNOWLEDGEMENTS

We thank Anders Odén and Stina Simonsson for valuable comments on the statistical analysis. This work was supported by the Swedish Cancer Society 2552-B95-09XBC, the Åke Wiberg Foundation and the Wilhelm and Martina Lundgren Fund.

REFERENCES

1 Kornberg,A. and Baker,T.A. (1992) *DNA Replication.* Freeman, New York, 2nd Edn.
2 Wong,S.C. and Schaffer,P.W. (1991) *J. Virol.* **65**, 2601–2611.
3 Elias,P., O'Donnell,M.E., Mocarski,E.S. and Lehman,I.R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6322–6326.
4 Elias,P., Gustafsson,C.M. and Hammarsten,O. (1990) *J. Biol. Chem.* **265**, 17167–17173.
5 Elias,P., Gustafsson,C.M., Hammarsten,O. and Stow,N.D. (1992) *J. Biol. Chem.* **267**, 17424–17429.
6 Bruckner,R.C., Crute,J.J., Dodson,M.S. and Lehman,I.R. (1991) *J. Biol. Chem.* **266**, 2669–2674.
7 Fierer,D.S. and Challberg,M.D. (1992) *J. Virol.* **66**, 3986–3995.
8 Olivo,P.D., Nelson,N.J. and Challberg,M.D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5414–5418.
9 Hernandez,T.R., Dutch,R.E., Lehman,I.R., Gustafsson,C. and Elias,P. (1991) *J. Virol.* **65**, 1649–1652.
10 Gustafsson,C.M., Hammarsten,O., Falkenberg,M. and Elias,P. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4629–4633.
11 Lockshon,D. and Galloway,D.A. (1988) *Mol. Cell. Biol.* **8**, 4018–4027.
12 Koff,A., Schwedes,J.F. and Tegtmeyer,P. (1991) *J. Virol.* **65**, 3284–3292.
13 Gustafsson,C.M., Falkenberg,M., Simonsson,S., Valadi,H. and Elias,P. (1995) *J. Biol. Chem.* **270**, 19028–19034.
14 Makhov,A.M., Boehmer,P.E., Lehman,I.R. and Griffith,J.D. (1996) *EMBO J.* **15**, 1742–1750.
15 Kowalski,D., Natale,D.A. and Eddy,M.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9464–9468.
16 Umek,R.M. and Kowalski,D. (1988) *Cell* **52**, 559–567.
17 Kowalski,D. and Eddy,M.J. (1989) *EMBO J.* **8**, 4335–4344.
18 Umek,R.M. and Kowalski,D. (1990) *Nucleic Acids Res.* **18**, 6601–6605.
19 Umek,R.M. and Kowalski,D. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2486–2490.
20 Natale,D.A., Schubert,A.E. and Kowalski,D. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2654–2658.
21 Huang,R.Y. and Kowalski,D. (1993) *EMBO J.* **12**, 4521–4531.
22 Hammarsten,O., Yao,X. and Elias,P. (1996) *J. Virol.* **70**, 4523–4529.
23 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
24 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds) *Current Protocols in Molecular Biology*, Suppl. 8. Wiley, New York.
25 Lin,S. and Kowalski,D. (1994) *J. Mol. Biol.* **235**, 496–507.
26 Bell,S.P. and Stillman,B. (1992) *Nature* **357**, 128–134.
27 Marahrens,Y. and Stillman,B. (1992) *Science* **255**, 817–823.
28 Coverley,D. and Laskey,R.A. (1994) *Annu. Rev. Biochem.* **63**, 745–776.
29 Krysan,P.J., Smoth,J.G. and Calos,M.P. (1993) *Mol. Cell. Biol.* **13**, 2688–26996.
30 Smith,J.G. and Calos,M.P. (1995) *Chromosoma* **103**, 597–605.