MOLECULAR AND CELLULAR BIOLOGY, Oct. 1993, p. 5931–5942
0270-7306/93/105931-12$02.00/0

Vol. 13, No. 10

Analysis of a Replication Initiation Sequence from the Adenosine Deaminase Region of the Mouse Genome

VALERIE J. VIRTA-PEARLMAN,[1] PREETHI H. GUNARATNE,[1] AND A. CRAIG CHINAULT[1,2]*

Institute for Molecular Genetics[1] and Department of Biochemistry,[2] Baylor College of Medicine, Houston, Texas 77030

Received 2 April 1993/Returned for modification 9 June 1993/Accepted 30 June 1993

A 4-kb *Hin*dIII fragment that supported the efficient autonomous replication of plasmid vector pDY⁻, a replication-defective construct based on Epstein-Barr virus sequences, in human K562 cells was rescued from amplified double-minute chromosomes containing the murine adenosine deaminase locus. Polymerase chain reaction assays of size-fractionated nascent strands demonstrated that replication initiation occurred within the same 1- to 2-kb region of this fragment in autonomously replicating plasmids containing the sequence in either orientation, in double-minute chromosomes, and in the single-copy locus at its normal chromosomal location. The complete sequence of this fragment was determined; it contains a 248-bp polypurine tract and consensus binding site sequences for several putative transcription and replication factors.

Our understanding of the molecular mechanisms of DNA replication in mammalian cells is largely based on the study of viruses. Although such studies have been extremely productive, there are certain aspects of chromosomal replication that cannot be addressed with such model systems. For example, although the initiation of viral replication is clearly dependent on specific sequences that are recognized by self-encoded proteins, such as simian virus 40 T antigen, neither the DNA signals nor the proteins involved in the initiation of replication in mammalian chromosomes have been defined. In addition, there are unique temporal and spatial controls that must come into play with chromosomal replication to coordinate the accurate and orderly duplication of each DNA base in the entire genome during the S phase.

It is clear that for the replication of mammalian chromosomes during the S phase at the measured rates of replication fork movement, the initiation of DNA synthesis must occur at many sites. By analogy with prokaryotic systems (40, 44) and simple eukaryotic systems (10, 11, 24, 42, 43, 51, 80), one would expect this initiation to be dependent on specific signals. However, the continuing ambiguity about the identity of such signals must be due to one of two factors: either (i) they do not exist or (ii) the intrinsic complexity and/or redundancy of these signals make them difficult to dissect by standard approaches. The fact that many of the assays available to search for such sequences are both experimentally difficult and dependent on fundamental assumptions concerning the properties of putative replication intermediates further complicates this issue (74). The most promising candidates for mammalian chromosomal replication origins are sites flanking the *DHFR* locus in Chinese hamster cells (4, 12–15, 38, 53, 73, 75), and the c-*myc* locus in the human genome (55, 72), in which sequences have been implicated in origin activity by multiple assays. However, even at the most rigorously studied region 3′ of the *DHFR* gene, the conclusions from different approaches have not been completely consistent; whereas some assays indicate sequence specificity in origin activity (14, 37, 73), others indicate replication initiation over very broad zones (27, 75). Models for the reconciliation of these data have been proposed (26, 52), but experimental resolution is still lacking.

The work described here was designed to take advantage of the availability of murine fibroblast cell lines in which the *ADA* locus has been amplified up to 11,000-fold (65, 82) to search for replication origins near this locus. Because of this extraordinarily high copy number, the amplified sequences in these cells are estimated to represent as much as 20% of the total nuclear DNA (46), a fact that should alleviate sensitivity problems associated with many of the assays used to study replication properties. In addition, the amplified copies are present on autonomously replicating double-minute chromosomes, which must, by definition, contain one or more origins of replication. In this paper, we describe the isolation from the amplified domain of a 4-kb *Hin*dIII fragment that is able to specifically complement a replication-defective derivative of an autonomously replicating, Epstein-Barr virus (EBV)-derived plasmid. A modified protocol for determining the sites of the initiation of replication by polymerase chain reaction (PCR) analysis of nascent DNA strands was then used to demonstrate that replication of the rescued plasmid initiated within the cloned insert regardless of its orientation. This assay was then used to demonstrate that, within experimental error, the same region was used as a replication initiation site in double-minute chromosomes and in its normal single-copy chromosomal context. Comparison of these results with those obtained for other loci may facilitate the identification of common sequence or structural motifs involved in defining replication origins in mammalian cells.

MATERIALS AND METHODS

Cell culture. The original murine *ADA*-amplified cells (82) were modified by the integration of a herpes simplex virus (HSV) thymidine kinase (*TK*) gene to allow the incorporation of thymidine analogs. A transfectant that incorporated thymidine at the wild-type level was used for all subsequent procedures. Cell lines B1/50tk and B1/150tk were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and glutamine (GIBCO Laboratories), 1.1 mM adenosine–50 μM alanosine–1 mM

* Corresponding author. Electronic mail address: chinault@bcm.tmc.edu.

uridine, and 50 to 150 μM deoxycoformycin (also known as pentostatin [Parke-Davis/Warner Lambert Co.]) to maintain the desired amplification level (65). NIH 3T3 cells were grown in DMEM supplemented with 10% fetal bovine serum and glutamine.

Eukaryotic vectors and constructs used in the functional assay. Plasmids p220.2 and pDY$^-$ (28) were generous gifts from M. Calos. p220.2 is an 8.9-kb plasmid that contains the following EBV sequences: *oriP* (EBV coordinates 7333 to 9516) and the *EBNA-1* gene (107567 to 110176). The hygromycin resistance gene is under the control of the HSV *TK* gene promoter and transcription termination sequences. The bacterial origin of replication and the ampicillin resistance gene are from a pBR322 derivative with the poison sequence deleted. pDY$^-$ is an 8.8-kb plasmid derivative of p220.2 in which the dyad symmetry region of *oriP* (*Eco*RV-*Hpa*I fragment) has been deleted (28). *Hin*dIII-digested B1/150tk DNA was cloned into the unique *Hin*dIII site of pDY$^-$ to make a mouse genomic library. This library was used to transform *Escherichia coli* HB101 cells, and a large-scale plasmid preparation was made from a pool of several thousand ampicillin-resistant transformants.

Mammalian cell transfection, DNA isolation, and analysis. Supercoiled plasmid DNA (20 μg) was used to transfect K562 cells (human erythroleukemia cell lineage; American Type Culture Collection) by electroporation with a Gene-Pulser transfection apparatus (Bio-Rad) in accordance with the manufacturer's recommendations and the guidelines of Chu et al. (19). The parameters used were 200 V and 500 μF. Transfection efficiencies for K562 cells were determined by soft-agar cloning as previously described (70). Transfected cells were selected for hygromycin resistance in liquid cultures with 300 μg of hygromycin B (Boehringer or Calbiochem) per ml. Resistant cells were passaged three times by splitting a confluent plate 1:10. Plasmids were rescued from Hirt supernatants (41) by transformation into *E. coli* DH5α cells made competent by the protocol of Hanahan (66). Extrachromosomal DNA for restriction analyses was isolated by an alkaline extraction method as follows. Approximately 10^7 cells were resuspended in 0.6 ml of 50 mM glucose–10 mM EDTA–25 mM Tris-Cl (pH 8.0) and lysed by the addition of 1.2 ml of 0.2 N NaOH–1% sodium dodecyl sulfate (SDS) and incubation on ice for 10 min. Chromosomal DNA was precipitated by the addition of 1.5 volumes of 3 M potassium acetate–0.2 M glacial acetic acid, incubation on ice for 10 to 30 min, and centrifugation with a Sorvall SA-600 rotor at 10,000 rpm for 10 min at 4°C. The supernatant was extracted once with phenol-chloroform (1:1), and the low-molecular-weight DNA was precipitated by the addition of an equal volume of isopropanol and incubated at −20°C for 4 to 24 h. This precipitate was collected by centrifugation as described above and resuspended in TE (10 mM Tris-HCl [pH 8.0], 1 mM EDTA). NaCl was added to a 200 mM final concentration, the solution was extracted once with phenol, and then the DNA remaining in the aqueous phase was precipitated with 2 volumes of ethanol, washed with 70% ethanol, and resuspended in 50 μl of TE. Plasmid replication was monitored by *Dpn*I or *Mbo*I digestion, with subsequent electrophoresis through an agarose gel, Southern transfer onto a GeneScreen Plus membrane (New England Nuclear), and hybridization (66). The probe used for extrachromosomal DNA analysis was plasmid p220.2 minus the *EBNA-1* and *oriP* sequences; autoradiographic exposure times were 20 to 36 h.

DNA sequencing, sequence analysis, and PCR primer production. The 4-kb insert from plasmid pDY/377 (377 fragment) was subcloned into pTZ18R, and exonuclease III deletion constructs were made from both orientations and transformed into *E. coli* XL1-Blue cells for sequencing. The sequence was obtained by use of an automated dideoxynucleotide sequencing system from Applied Biosystems (model 373A) or by manual dideoxynucleotide sequencing. The following parameters were used for consensus binding site sequence analysis: NFI (CTF), minimum 6 of 6 match to GCCAAT (1, 20, 57); NFIII (Oct-1), minimum 9 of 10 match to ATGMWAATGA or ATGATAATGA (57, 62); RIP-60, minimum 9 of 11 match to ATTATTATTAG (23); EBNA-1, minimum 14 of 16 match to GRTAGCNNNNGCTAYC (3); PUR, minimum 15 of 16 match to GGNNGAGGGAGARRRR (9); p53, minimum 10 of 10 match (paired within 0 to 13 nucleotides [nts]) to RRRCWWGYYY (6, 29); and the ARS core consensus sequence, minimum 9 of 11 match to WTTTATRTTTW (45). (International Union of Pure and Applied Chemistry-International Union of Biochemistry ambiguity codes for the consensus sequences are as follows [22]: W = A or T, R = A or G, Y = T or C, M = A or C, and N = A, C, G, or T.) Sequence data bases were searched by use of a previously described program (2), and Monte Carlo simulation was done as described by Barker and Dayhoff (7). GenBank nucleotide computations were performed at the National Center for Biotechnology Information by use of the BLAST network service. Oligonucleotides were made on an Applied Biosystems model 380B synthesizer. Primer sets were designed from published EBV sequence information or from sequences obtained as described above.

Nascent-strand isolation for use in the PCR origin localization assay. The method that was developed by Vassilev and Johnson (71) for mapping replication initiation sites by PCR was used with modifications. Ten 100- to 150-mm plates of B1/50tk cells at 50 to 70% confluency were used for each experiment. These plates were labeled for 15 min at 37°C with 20 μM 5-bromodeoxyuridine (BrdU; Sigma) and 60 pM [^{3}H]thymidine as a tracer. All subsequent procedures were done in low light to limit UV-induced damage. The cells were washed three times with cold phosphate-buffered saline and lysed by the addition of 3 ml of cell lysis buffer (50 mM Tris-HCl, [pH 8.0], 1 M NaCl, 10 mM EDTA, 5% SDS). Lysates were combined and incubated overnight with 100 μg of proteinase K per ml of lysate. The DNA was extracted gently once with phenol-chloroform (1:1), ethanol precipitated, rinsed in 70% ethanol, rinsed in 95% ethanol, and redissolved in 1 ml of TE. NaOH was added to a final concentration of 0.2 N. Alkaline cesium chloride (CsCl) gradients were run as described previously (30) to separate the newly replicated, single-stranded DNA containing BrdU (heavy DNA [H-DNA]) from the unsubstituted DNA (light DNA [L-DNA]). CsCl dissolved in 50 mM NaOH–3 mM EDTA was added to the DNA to a final refractive index (η) of 1.408. This mixture was centrifuged at 85,000 × *g* for 60 to 72 h with a Sorvall T1270 rotor at 25°C. After centrifugation, approximately 48 fractions were recovered from the bottom of each centrifuge tube and placed in a 96-well polyvinyl chloride plate by use of a Hoefer gradient fractionator. The η and radioactive incorporation were assayed with 10- and 50-μl aliquots, respectively, from every other fraction. Gradient fractions containing BrdU-substituted H-DNA (η, 1.411 to 1.413) or parental L-DNA (η, 1.404 to 1.407) were pooled. The pools were diluted by the addition of 2 volumes of water and precipitated with 2.5 volumes of 95% ethanol at 4°C. Precipitates were collected by centrifugation with a Sorvall AH-629 rotor at 10,000 rpm for 30 min at 4°C. The precipitates were washed with 70% ethanol and air dried.

The H-DNA was resuspended in 50 to 100 μl of water (approximately 10 μl/10^8 cells). Twenty five microliters of purified nascent DNA was size fractionated by electrophoresis through a 1% alkaline agarose gel for 16 h at 2 V/cm and 4°C. Four hundred nanograms of λ *Hin*dIII-φX174 *Hae*III (the former from Bio-Rad and the latter from BRL) and a 1-kb ladder (BRL) were used as size markers. The gel was neutralized with several changes of 0.5× TBE for 2 h at room temperature. The standard lanes were cut off and stained with 0.5× TBE containing 0.5 μg of ethidium bromide per ml. Gel slices containing fragments of desired sizes were placed in 1.5-ml Eppendorf tubes, and the single-stranded nascent DNA was isolated by use of a Prep-A-Gene kit and protocol (Bio-Rad). The DNA was collected in a final volume of 20 to 30 μl, and 1 to 4 μl was used for PCR analysis with three to five primer sets. Okazaki fragments should have been present in the smallest fraction isolated (approximately 0.25 to 0.5 kb); however, recovery and PCR detection of these small intermediates were found to be highly variable, as a function of the preparation and primer pair used, and the levels of PCR products produced were frequently inadequate to yield visible bands on positive photographs.

This procedure was also used for the isolation of nascent strands from plasmid-transfected eukaryotic cells. In these instances, after the proteinase K digestion (2 to 16 h), extrachromosomal DNA was isolated by the method of Hirt (41). The DNA was incubated for >8 h at 4°C. Chromosomal DNA and proteins were pelleted at 10,000 × *g*, the resulting supernatant was ethanol precipitated, and the rest of the protocol outlined above was used.

PCR conditions. PCR primers designated Orn, Hyg, and EBNA (see Fig. 2, 3, and 4) amplify segments of 460, 423, and 390 nts, respectively, from the p220.2 and pDY$^-$ vectors. The primer sequences are as follows (the primer pairs are denoted forward [F] and reverse [R]): Orn, F(5′-CAAT GTTGTGTTGCAGTCCACAGACTGC-3′) and R(5′-CTAA ACCTGACTACGGCATCTCTGC-3′); Hyg, F(5′-CGCCAT GTAGTGTATTGACCGATTCC-3′) and R(5′-CGCGACGT CTGTCGAGAAGTTTCTGAT-3′); and EBNA, F(5′-CAGT AGCATCTCTGTCTGGTGACCTT-3′) and R(5′-CTTCTCT CCTAGGCCATTTCCAGGT-3′). The PCR primer sets for segments A, B, C, D, and E (see below) all amplify 230-nt regions of the genomic ADA amplicon, except for the set for B, which yields a product of 370 nts. The sequences are as follows: segment A, F(5′-CTGAGACTATCCTCCAGGTC TTCT-3′) and R(5′-CATGGCTGCCTATGACCAACAGAA-3′); segment B, F(5′-CATTCCTCTATGGCGCTGGCCATG AT-3′) and R(5′-CCAACAACGTCTGTGGTGCAGAGA-3′); segment C, F(5′-CGTTCAATATGCCTCCTTCGTGTG-3′) and R(5′-CAGGGCTGTGAGCAGTAAACCTGTATTG-3′); segment D, F(5′-AACAGCATACTTGGTGGTAGGTG TGGC-3′) and R(5′-TCCGATTTGCAAGACTCCTCTATG GC-3′); and segment E, F(5′-GAACACTCTGAGCCTCAG TTTGTTC-3′) and R(5′-AGTCTCCATCAGTGACACATT GCC-3′).

The PCR was carried out under standard conditions with 1× Cetus buffer (10× Cetus buffer is 500 mM KCl, 100 mM Tris-HCl [pH 8.3], 15 mM $MgCl_2$, and 0.01% gelatin), the four deoxyribonucleotides (200 μM each), and *Taq* polymerase (1 U per reaction mixture) and with a reaction volume of 20 μl. Primers were used at a 0.5 μM final concentration. PCR amplification was completed through 30 cycles (94°C, 30 s; 55 to 57°C, 0.5 to 1 min; 72°C, 1 min) unless otherwise noted. Perkin-Elmer Cetus or MJ Research thermal cyclers were used interchangeably.

Nucleotide sequence accession number. The GenBank accession number of the 377 fragment sequence is L20424.

RESULTS

Functional rescue of a replication initiation sequence. It has been shown elsewhere that plasmids containing a 1.7-kb sequence near the replication origin of EBV (referred to as *oriP*) plus a functional *EBNA-1* (Epstein-Barr nuclear antigen 1) gene can be maintained extrachromosomally at low copy levels in human and rodent cells (49, 77, 81). The *oriP* sequence includes 21 tandem copies of a 30-bp sequence that may be involved in plasmid maintenance, transcriptional enhancement, and replication termination. It also contains a region of dyad symmetry in which four partial copies of this 30-bp sequence form a 65-bp palindrome that appears to be the actual site of replication initiation (64). Krysan et al. (49) demonstrated that by deleting from one such construct (p220.2) a 140-bp sequence at the palindromic region, a replication-defective derivative, pDY$^-$, could be obtained. Interestingly, it was found that pDY$^-$ could still be maintained for prolonged periods in nuclei of growing cells, a property that is referred to as "nuclear maintenance" and that is apparently dependent on an EBNA-1 interaction with the 30-member family of repeats (63, 69).

Restriction fragments from a highly amplified mouse cell line (B1/150tk) were used in a functional complementation assay in an attempt to isolate autonomously replicating derivatives of pDY$^-$. *Hin*dIII fragments of B1/150tk genomic DNA were ligated into the unique *Hin*dIII site of the pDY$^-$ vector, and recombinant plasmids were introduced into *E. coli* by transformation. Ampicillin-resistant cells were pooled, and a mixed-plasmid preparation was made from the pool. Twenty micrograms of purified plasmid DNA was then used for electroporation into human K562 cells. It was assumed that the enormous overrepresentation of plasmids with inserts derived from amplified sequences would result in their uptake into a large number of cells, compared with the uptake of those containing any individual nonamplified sequence, and that any plasmid containing a fragment from the amplified domain that resulted in preferential extrachromosomal replication could be recovered from cells surviving selection. Therefore, after 6 to 9 weeks of hygromycin selection, Hirt supernatants were prepared (41) and plasmids were rescued back into *E. coli*. Two independent electroporations yielded a total of eight *E. coli* transformants. On the basis of mapping and hybridization, five of the eight clones were shown to be identical; the 4-kb insert in this recombinant, referred to as pDY/377, was shown to be from the amplified domain by hybridization (data not shown). The other *E. coli* transformants did not contain sequences from the amplified domain and were not studied further.

When the purified pDY/377 construct was subsequently reintroduced into K562 cells by electroporation and transfection efficiency was measured by soft-agar cloning, it was found in two independent experiments that 0.6 to 0.7% of the cells surviving the electroporation (approximately 62%) were able to form colonies in the presence of hygromycin selection. Similar experiments with p220.2 indicated a transfection frequency of about 1%; pDY$^-$ yielded no surviving colonies. This result indicates that the replication defect in pDY$^-$ is efficiently complemented by the cloned insert.

After transfection of K562 cells with purified plasmids p220.2 and pDY/377, the replication state of the constructs was tested with extrachromosomal DNA preparations. Plas-

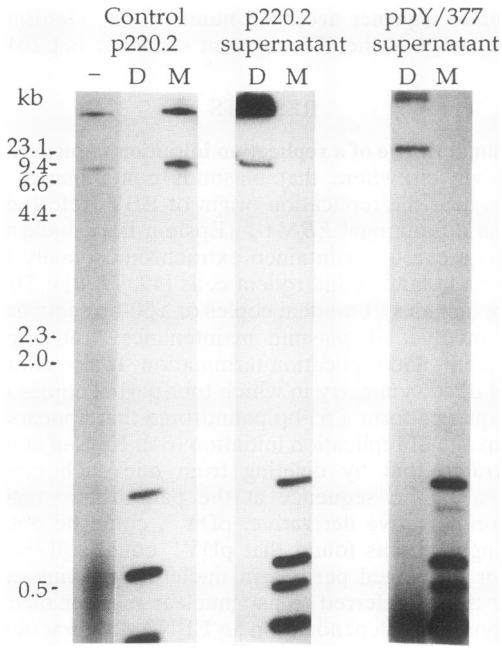

FIG. 1. Determination of replication status by digestion with methylation-sensitive restriction enzymes. Control plasmid p220.2, obtained from bacteria, was analyzed without digestion to establish the migration positions of supercoiled and nicked circular forms and after digestion with *Dpn*I (D) or *Mbo*I (M). Both p220.2 and pDY/377 extrachromosomal supernatants were prepared from 10^6 K562 cells after introduction of the plasmids by electroporation and maintenance for 6 weeks with hygromycin selection. These samples were then digested with either *Dpn*I or *Mbo*I, as indicated, prior to gel electrophoresis, transfer to a nylon filter, and hybridization with vector sequences as outlined in Materials and Methods. There are >20 potential sites of cleavage for these enzymes, and only a few of the small fragments generated are shown; the extra band below 0.5 kb in the supernatant digests is apparently due to a partial digestion product. λ *Hin*dIII DNA size standards are indicated.

mids that undergo replication should lose the bacterial N^6-methyladenine modification and become resistant to digestion by *Dpn*I; conversely, replicated DNA should become sensitive to digestion by *Mbo*I (32, 61). As shown in Fig. 1, control plasmid p220.2 isolated from bacteria was *Dpn*I sensitive and *Mbo*I resistant. However, plasmid p220.2 molecules rescued from transfected K562 cells were *Dpn*I resistant and *Mbo*I sensitive. This same situation was found with rescued plasmid pDY/377 molecules, indicating that these plasmids had also undergone replication.

Nascent-strand PCR assay. A method for defining replication initiation points by analyzing the ability of pulse-labeled DNA to generate amplified products in the PCR was recently published (71). This general approach was adapted with extensive modifications to improve the selectivity and sensitivity of the method. The general outline of our protocol is as follows. Unsynchronized cells are pulse-labeled for 15 min with BrdU and [^{3}H]thymidine. The DNA is extracted, and the newly replicated (BrdU-containing) DNA is separated from unsubstituted DNA by equilibrium centrifugation in an alkaline CsCl gradient. The BrdU-substituted DNA is collected and size fractionated by electrophoresis through an alkaline agarose gel. Nascent DNA fragments of desired sizes are then isolated from appropriate gel slices and subjected to PCR with primer sets of interest. The PCR products obtained from each fraction are then analyzed on agarose gels, with visualization being done directly by ethidium bromide staining or by subsequent Southern hybridization with labeled probes.

EBV-based plasmid p220.2 was used as a positive control to demonstrate that the nascent-strand PCR assay could identify a previously mapped initiation point. Figure 2 shows an ethidium bromide-stained agarose gel in which the PCR products from nascent fragments of different sizes from p220.2-containing human K562 cells were analyzed, as well as a diagram showing the linear map of p220.2 and the results predicted to occur with the primer pairs used. If replication proceeds bidirectionally from the origin, one would expect that the actual distance between the initiation point and the first positive assay point would be half the size of the shortest nascent stand detected. In the EBV plasmids, however, there is a replication terminator that lies upstream from the initiation site (at or near the 21-member family of repeats; 32). Therefore, in this case, replication proceeds bidirectionally until it encounters this terminator, and then the rest of the plasmid is replicated unidirectionally from the other fork. In Fig. 2, the Orn primer set from the p220.2 sequence lies about 0.5 kb from the origin and, as predicted, generates a PCR product in the 0.5- to 1-kb nascent-strand fraction. The Hyg primer set is 2.1 kb from the origin, and an amplification product is first seen in the 3- to 4-kb fraction. Similarly, the EBNA primer set is 5.1 kb from the origin, and an amplification product is first seen in the 6- to 8-kb fraction, a result consistent with the predicted size if the terminator is functional. Thus, the assay as outlined appears to be capable of pinpointing an initiation point to within a reasonable distance.

The assay was then repeated with nascent strands isolated from K562 cells containing pDY/377 plasmid constructs. Figures 3 and 4 show maps and present the results of experiments with the insert in both orientations. Figure 3 shows the results from the A-C orientation. Four primer sets were used; they included the three used previously with the positive control and an additional one (referred to as 377C) developed from the 377 fragment insert sequence. In this case, again taking into account the expected termination of DNA replication in one direction, replication initiates from within the 377 fragment. The Orn primer set from the vector now first amplifies a product in the 5- to 6-kb fraction. The 377C primer set, from within the 377 fragment, first amplifies a product in the 0.25- to 0.5-kb fraction, indicating that it is closest to the region of initiation. The Hyg primer set first amplifies a product in the 4- to 5-kb fraction, and the EBNA primer set first amplifies a product in the 8- to 10-kb range. In a repeat experiment, a primer pair from the other end of the insert (referred to as 377A in Fig. 4) first showed PCR product generation in the 5- to 6-kb fraction, and all other primer sets yielded the same results as those obtained before (data not shown).

Figure 4 shows the analyses of nascent strands from cells containing the reverse C-A orientation construct. In this experiment, the Orn primer set now first amplifies a product in the 2- to 3-kb fraction. The 377C primer set first amplifies a product in the 0.5- to 1-kb fraction, indicating that it is still closest to the region of initiation. The 377A primer set first amplifies a product in the 5- to 6-kb fraction, and the Hyg primer set first amplifies a product in the 6- to 7-kb fraction. In a separate experiment, appropriate PCR products were first detected in the 0.5- to 1-, 2- to 3-, 5- to 6-, and 8- to 10-kb fractions for the 377C, Orn, 377A, and Hyg primer pairs, respectively (data not shown). It should be noted that in Fig.

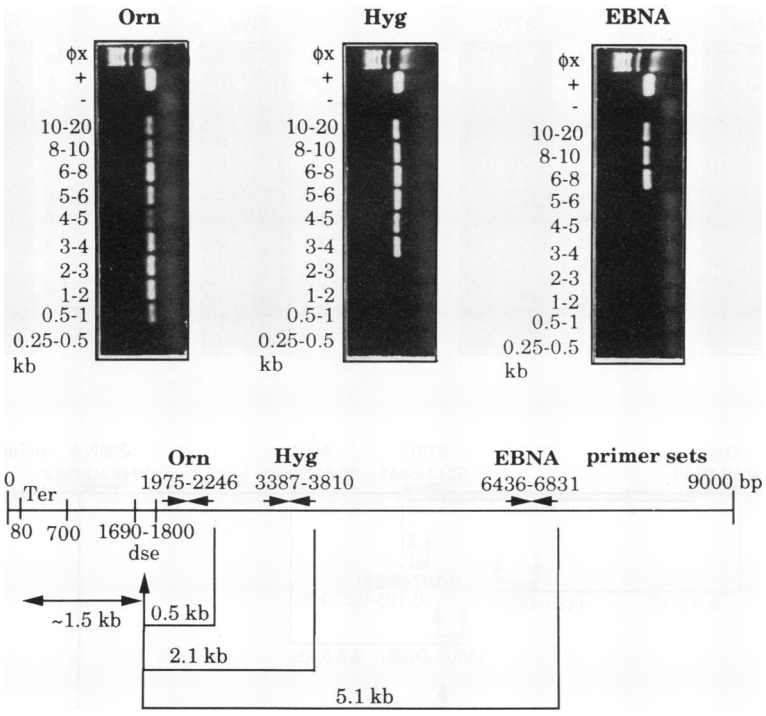

FIG. 2. PCR analysis of nascent strands from cells carrying p220.2. The panels are photographs of ethidium bromide-stained agarose gels showing the PCR amplification products from nascent strands fractionated according to size, as indicated to the left of each panel. A linear map of the plasmid is shown below. The dyad symmetry element at which DNA replication initiates is designated dse. The termination region in the family of repeats is labeled Ter. The PCR primer sets used were Orn, Hyg, and EBNA, the sequences of which are given in Materials and Methods. Lane − shows PCR with no DNA; lane + shows amplification of plasmid DNA (20 ng); lane φx shows φX174 *Hae*III DNA size standards.

4, the predicted positions of initiation are diagrammed on the assumption of equal rates of bidirectional replication. However, depending on the actual initiation point, termination activity at Ter could cause the production of asymmetric nascent strands, so that the actual positions of initiation predicted from the data for the 377A and Hyg primers would be shifted further to the left, making them even more consistent with the region predicted from the results with the other two primer pairs. Therefore, all the data suggest that there is a region within the 377 fragment that is functional for the initiation of DNA replication for plasmid pDY/377 in the K562 human cell line. It is unclear whether the lack of complete coincidence of the predicted initiation positions is due to multiple sites, unequal rates of bidirectional replication, or simply a technical limitation of the assay with respect to distance determinations.

For determination of whether the 377 fragment sequence is used as an initiation site in a chromosomal context, the same approach was used to study nascent DNA isolated from a mouse cell line (B1/50tk) with a ca. 5,000-fold amplification of the *ADA* region on double-minute chromosomes and carrying the HSV *TK* gene. Figure 5 shows the results of the PCR analyses of this DNA after fractionation as described above. The A primer set amplifies a product that first appears in the 4- to 5-kb fraction, indicating that initiation occurs 2 to 2.5 kb downstream, as drawn, assuming that replication is bidirectional. The B primer set first amplifies a product in the 1- to 2-kb fraction. The C primer set first amplifies a product in the 0.5- to 1-kb fraction. The D primer set first amplifies a product in the 4- to 5-kb fraction. The E primer set first amplifies a product in the 7.5- to 10-kb fraction.

Together, these data suggest that initiation occurs nearest primer set C and that primers located distal to this primer, both upstream and downstream, are only able to generate products from larger nascent strands, as diagrammed. Therefore, it appears that initiation occurs on the double-minute chromosomes at the same position as in the autonomous pDY/377 construct, within experimental error, and proceeds bidirectionally.

For assessment of whether the same region functions as a replication initiation site in the normal, nonamplified chromosomal context, the assay was applied to an independent mouse cell line, NIH 3T3. In this case, it was necessary to transfer the PCR products from the agarose gel to a nylon filter and hybridize with radioactive probes to visualize the results because of the lower signal levels. Nevertheless, as shown in Fig. 5 (right-hand panels), the primer sets used yielded essentially the same results as those obtained with the amplified cell line. Taken together, these data map the region of initiation to approximately 1 to 2 kb, between positions 2000 and 3993 of the sequence (see below).

DNA sequencing and sequence analysis. The complete nucleotide sequence of the 4-kb 377 fragment (Fig. 6) was determined by a nested deletion-dideoxynucleotide sequencing approach. Computer analysis revealed several features consistent with the potential role of this fragment as a replication origin. The fragment contains a murine B1 repeat and a very extensive polypurine tract (248 bp interrupted by one thymidine at position 2630) that is located within the

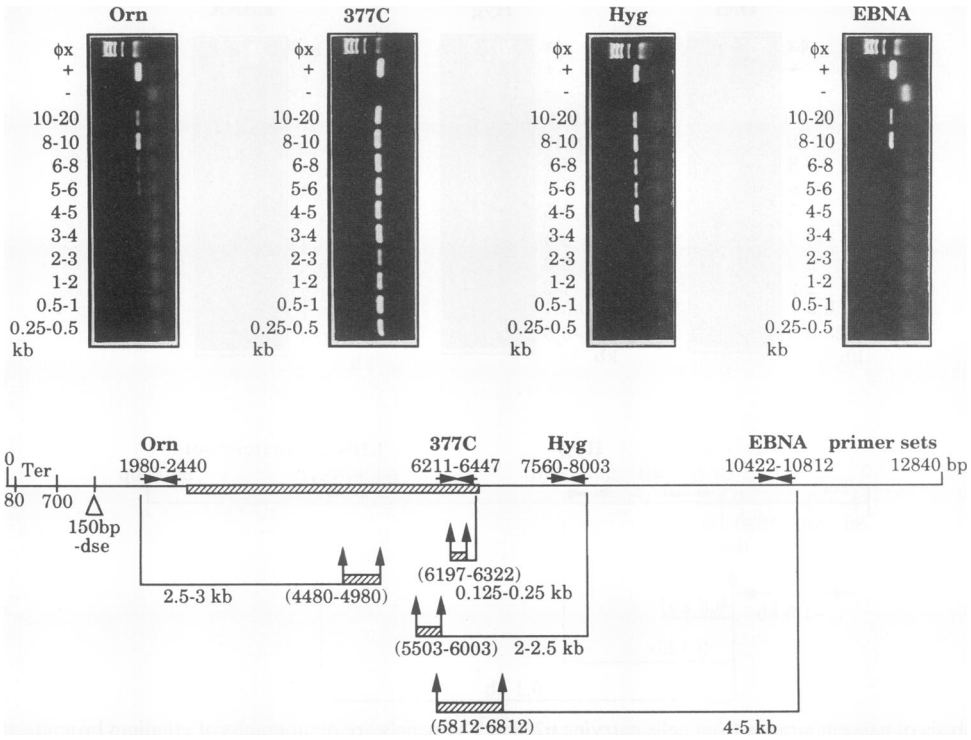

FIG. 3. PCR analysis of nascent strands from cells carrying pDY/377A-C. The panels are photographs of ethidium bromide-stained agarose gels showing the PCR amplification products from nascent strands fractionated according to size. Control and marker lanes are as described in the legend to Fig. 2. Below the results are a linear map of plasmid pDY/377 (in which the hatched box represents the murine genomic insert) and a diagram indicating the positions predicted for replication initiation on the basis of the PCR results obtained with each of the indicated primer pairs.

experimentally mapped initiation region. As indicated in Fig. 6, the sequence also contains three close matches to a purine-rich protein recognition element, PUR (9), and two matches to the consensus sequence for the RIP-60 protein (23), both of which have been noted elsewhere to be within sequences implicated in replication initiation (9, 16). In addition, there are consensus binding sites for a number of transcription factors that also have been suggested to play a role in replication in viral systems (25), including NFI (CTF), NFIII (Oct-1), and p53, as indicated in Fig. 6. For p53, one copy of the recognition motif is insufficient for DNA binding; two copies separated by 0 to 13 bp are required (29). Therefore, although seven perfect matches of the consensus sequence are present in the 4-kb fragment, only the two that satisfy this spacing requirement (positions 2978 to 3005, with an 8-bp separation) are indicated. There are also several occurrences of the AT-rich core sequence found in yeast ARS elements, as has been noted elsewhere in the vicinity of other potential mammalian origins (17, 33). Only two widely spaced sequences resembling the high-affinity EBNA-1 binding site sequence are present.

Interestingly, a search of the GenBank data base for sequences similar to that of the 377 fragment revealed a region, between positions 1483 and 1615, homologous to two previously reported sequences: a human chromosomal DNA sequence, ARS2, which replicates autonomously in *Saccharomyces cerevisiae* (56), and a mouse nuclear matrix DNA sequence, G5 (34). The regions of homology are shown in Fig. 7. The human ARS sequence, HUARS2A, shows two regions of similarity with the 377 fragment sequence. One region is 95% similar over 21 nts, and the other is 60% similar over 67 nts; the Monte Carlo simulation index indicates probable homology, with a score of 9.32 (7). The mouse nuclear matrix sequence, MUSMDG5, is 90% similar over 46 nts, with a Monte Carlo simulation index of 7.46, indicating probable homology. An *Xba*I-*Hin*dIII fragment near the Chinese hamster cell *DHFR* gene (17) also shows two regions of similarity with the fragment 377 sequence, as shown in Fig. 7. Region A is 72% similar over 104 nts, and region B is 68% similar over 98 nts.

DISCUSSION

Attempts to identify replication origins in mammalian chromosomes have been frustrated both by the lack of proven functional assays for the identification of such sequences and by the difficulty in obtaining adequate selectivity and sensitivity to examine the properties of single-copy loci in the context of a complex genome. Although progress has been made on the latter problem as a result of the availability of PCR technology, the development of a functional assay, which is crucial to be able to rapidly isolate and manipulate sequences involved in the initiation of DNA replication, has still not been achieved. Krysan et al. (49) tried to develop such an assay by deleting the replication initiation site of a replication-proficient plasmid containing the minimal elements from EBV needed for autonomous replication and then identifying mammalian DNA fragments able to overcome the resultant replication defect. Because of the preservation of the nuclear retention properties of the deletion plasmid, one may anticipate that complementation would require replacement of the initiation function. How-

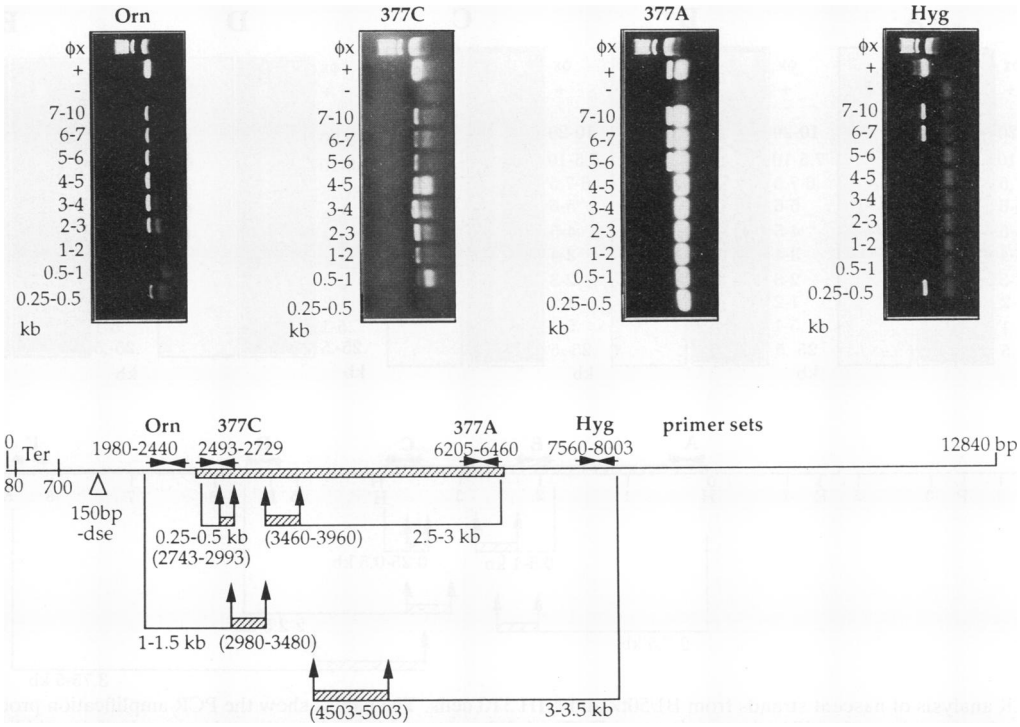

FIG. 4. PCR analysis of nascent strands from cells carrying pDY/377C-A. The panels are photographs of ethidium bromide-stained agarose gels showing the PCR amplification products from nascent strands fractionated according to size. Note that for the analysis described in the text, the signals for the 0.25- to 0.50-kb fractions in this experiment are assumed to be due to amplification from Okazaki fragments (see Materials and Methods). Control and marker lanes are as described in the legend to Fig. 2. Below the results are a linear map of plasmid pDY/377 (in which the hatched box represents the murine genomic insert) and a diagram indicating the positions predicted for replication initiation on the basis of the PCR results obtained with each of the indicated primer pairs.

ever, the validity of this approach was called into question by the finding that complementation showed length dependence, such that many fragments larger than 12 kb preferentially showed activity, and by the failure to demonstrate sequence-dependent initiation (39, 50). Both two-dimensional gel analyses of putative replication intermediates (48, 50) and the inability to obtain functional subclones (39) indicated that the few sequences studied in detail were not acting as simple origins. In contrast, we have been able to rescue a 4-kb fragment that appears to yield very efficient sequence-dependent complementation from the amplified domain that is associated with overexpression of the murine *ADA* gene. It is difficult to reconcile these results without sequence comparisons and further studies to determine the possible effects of differences in cell lines and protocols and to examine further the generality and/or intrinsic limitations of the approach. Nevertheless, it appears that in at least some cases, this functional approach to identifying replication initiation sequences may be applicable.

Interestingly, it has been found with EBV that the EBNA-1 protein mediates the formation of a DNA loop between the dyad symmetry element at which replication initiates and the nearby repetitive enhancer region (31, 68). In the functional pDY/377 constructs studied here, a 140-bp region at the normal initiation site has been replaced by a 4-kb fragment that functions to restore autonomous replication behavior. Although two reasonably close matches (10 of 12) to the preferred EBNA-1 binding site (3) are found within the 4-kb sequence, it seems unlikely that these would fortuitously provide the necessary elements for similar loop formation in the recombinant constructs. The actual role of EBNA-1 or other host proteins in the replication initiation observed within the inserts needs further investigation.

Given the uncertainty of the potential functional approach, it was felt that mapping of replication initiation sites by an independent method was crucial. For this, we adapted the PCR-based analysis of nascent strands developed by Vassilev and Johnson (71). Results from this assay were consistent with the presence of an origin of bidirectional replication within a 1- to 2-kb region of the insert in either orientation in pDY/377 plasmid derivatives. By use of additional PCR primers based on sequences flanking this region in the genome, it was also possible to show bidirectional replication initiation from the same area in the normal chromosomal context, both in amplified double-minute chromosomes and at a single-copy level in an independent cell line. This assay is limited by the fact that most of the nascent strands isolated tend to be smaller than 10 kb. This fact appears to be due to a combination of two factors. First, physically induced or UV-induced damage of the BrdU-substituted DNA may cause random breakage of the DNA. Second, because of the short pulse time, newly initiated DNA is more evenly labeled with BrdU than previously initiated DNA and, because partially substituted strands fractionate between the heavy- and light-density peaks in CsCl gradients, there may be preferential recovery of the more highly substituted smaller strands. Although these factors limit the scope of the analysis and make quantitation difficult, they do not appear to affect significantly the localization of the approximate position of replication initiation

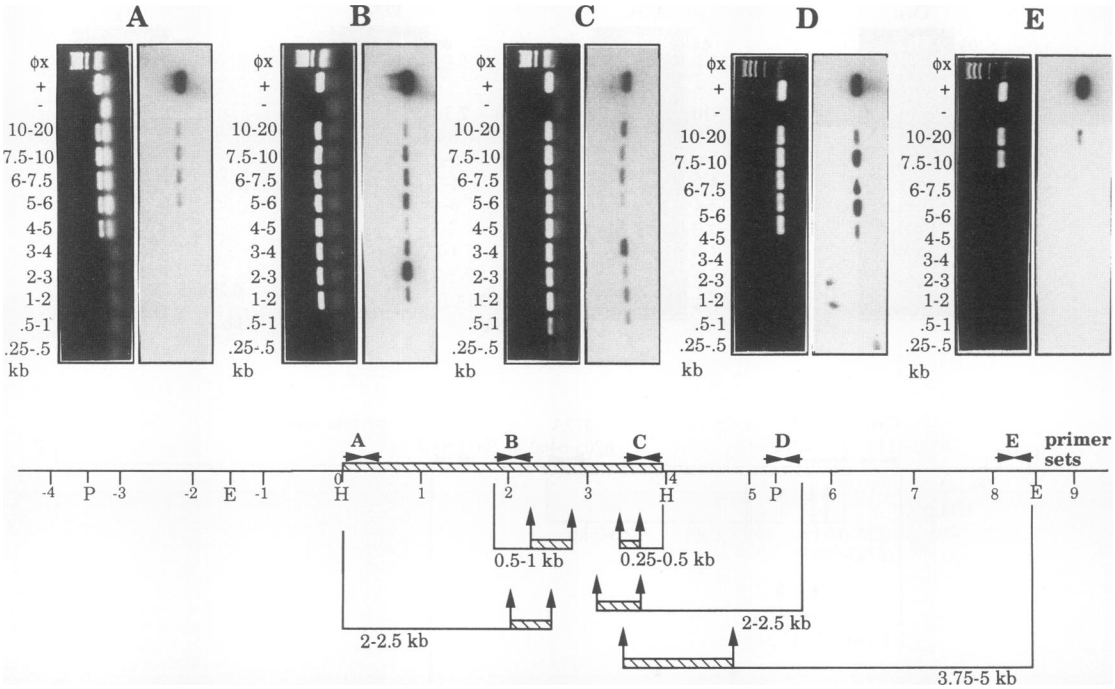

FIG. 5. PCR analysis of nascent strands from B1/50tk and NIH 3T3 cells. The panels show the PCR amplification products obtained from size-fractionated nascent strands with primer pairs A to E. The left-hand panel in each set is a photograph of an ethidium bromide-stained agarose gel showing PCR amplification products obtained from nascent strands from *ADA*-amplified cell line B1/50tk. The right-hand panel in each set is an autoradiograph showing PCR amplification products obtained from nascent strands from an unamplified cell line, NIH 3T3, after Southern transfer and hybridization with the appropriate labeled PCR products. Lane − shows a mock PCR without DNA; lane + shows PCR amplification of 20 ng of B1/150tk genomic DNA; lane φx shows φX174 *Hae*III DNA size standards. The map shows the positions of the primer sets used and the location of replication initiation predicted from PCR analysis with each primer pair. The hatched box on the map is the 4-kb *Hin*dIII fragment in the pDY/377 constructs analyzed in Fig. 3 and 4. P, *Pst*I; E, *Eco*RI; H, *Hin*dIII. Blots B and C were exposed for 6 h. Blot A was exposed for 12 h. Blots D and E were exposed for 30 h. All lanes labeled + and − were from a 6-h exposure.

within a limited (10- to 20-kb) region. It is not clear what the ultimate level of resolution is for this assay, although it appears to localize easily an origin to 1 to 2 kb.

Taken together, the results strongly support the presence of a sequence-specific DNA initiation element within this region of the mouse genome. The 377 fragment sequence appears to map about 150 kb 3′ of the *ADA* gene on the amplified double-minute chromosomes from which it was first isolated (35). It was shown recently that the double-minute chromosomes in a related cell line are primarily 500-kb circles comprising an imperfect inverted duplication of a 250-kb amplicon (59); preliminary characterization of the double-minute chromosomes in the cells used here is consistent with this structure (35). Interestingly, subsequent work to examine the replication timing of the *ADA*-377 fragment region has shown that the 377 fragment sequence replicates later in the S phase than the *ADA* gene itself, suggesting that the double-minute chromosomes must contain more than one initiation sequence (78). In agreement with this prediction, a second potential replication origin was recently independently localized to an 11-kb region lying approximately 30 kb 5′ of the *ADA* gene by use of different approaches designed to detect early firing origins of replication (18). Therefore, it appears that the amplified double-minute chromosomes may contain multiple potential replication origins, and detailed comparisons of their structural and functional properties may be very informative. It will also be important to conduct two-dimensional gel analyses of replication intermediates in the amplified region. Such studies with the *DHFR* locus have indicated nonspecific initiation over a region as large as 55 kb (27), in sharp contrast to the more specific origin localizations predicted from other assays. Similar studies with the system used in this work may help establish whether this apparent paradox is due, for example, to intrinsic qualitative or quantitative differences between the intermediates being studied and whether these results can be reproduced in other genomic regions.

Sequence analyses of putative mammalian chromosomal replication origin sequences have failed to identify simple consensus sequences, although a number of general sequence and structural features frequently have been found; these include consensus sequences for a number of putative transcription and replication factors, polypurine tracts, regions of bent DNA, and potential DNA-unwinding elements (reviewed in reference 8). Many of these features are consistent with findings obtained with simple eukaryotic systems, such as yeasts and viruses. However, without reliable functional analyses to allow genetic dissection of their roles in chromosomal replication, it is difficult to assess their importance in this process.

The most prominent feature of the 377 fragment sequence is the extensive polypurine tract located within the experimentally mapped region of replication initiation. There is evidence that such sequences can form triplex DNA and non-B DNA structures and may cause pausing or arrest of DNA synthesis (5). Also, their distribution is statistically nonrandom and conserved in vertebrate chromosomes (54, 79), also suggesting functional importance. Within this same

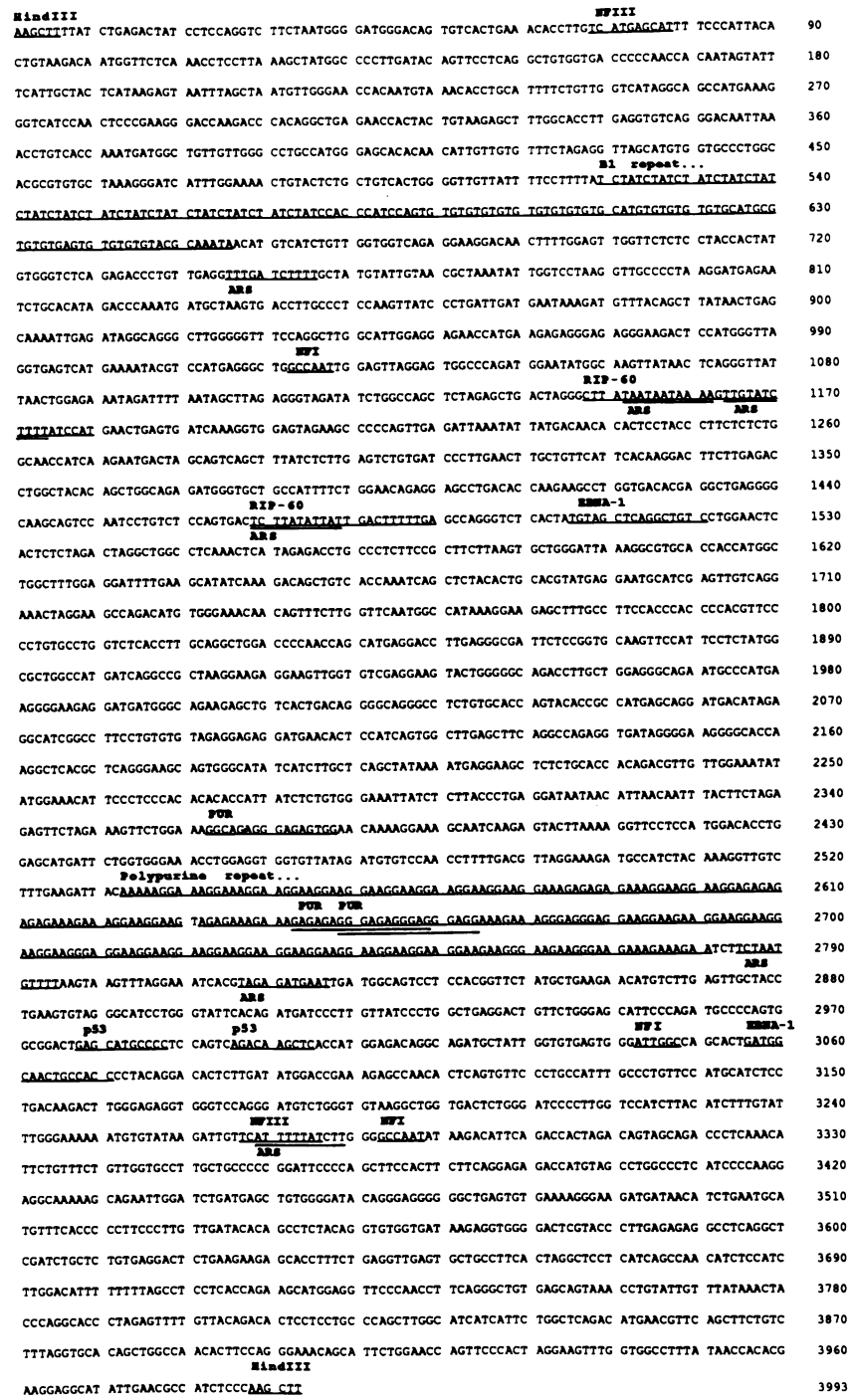

FIG. 6. DNA sequence of the 4-kb *Hin*dIII fragment (377 fragment). Consensus protein recognition sequences (see Materials and Methods for the parameters used) and other sequence features are labeled and underlined.

region are three exact matches to the consensus sequence identified by Bergemann and Johnson (9) as a binding site for a single-stranded DNA binding protein, designated PUR, from HeLa cells, which is hypothesized to act as a helix-destabilizing protein. This PUR element also has been found near other putative mammalian origins, including those near the human c-*myc*, β-globin, and rhodopsin genes and the hamster *DHFR* gene (9, 17, 33). Interestingly, the deletion of a PUR element from a yeast ARS plasmid abolished ARS activity (60), suggesting that this element may play a conserved functional role in replication. One may speculate that if polypurine tracts play an important role in defining preferred initiation sites, the unusual length of the tract in the 377 fragment may contribute to the very high (if not exclusive) level of usage of sequences at or near this region for replication initiation in the assays described in this work.

```
377        1483    TTTTTGA GCCAGGGTCT CACTATGTAG CTCAGGCTGT CCTGGAACTC ACTCTCTAGA
HUARS2A     640    ....... .g..... ....(----------------------46nts----------
MUSMDG5       3                                  ...... ......... ...t.g....
DHFR-A     2333                                       . .....c...t g....gg...
DHFR-B     4271                                    .g .a.....t.. .tcaca...t


377        1540 CTAGGCTGGC CTCAAACTCA TAGAGACCTG CCCTCTTCCG CTTCTTAAGT GCTGGGATTA
HUARS2A     707 --------). ....gg.a.. ggtgat.t.c ..a...cagc ..c.cc.tta .......c..
MUSMDG5      29 .c..... ..g.....
DHFR-A     2354 .c......a. ..g..ct... g........c .tgg.c.tg. .c..ccg... ctg.......
DHFR-B     4293 tc...t.... ..tg..... g......tca .tta.c.tt. .c..cc.c.. .....a....


377        1600 AAGGCGTGCA CCACCATGGC TGGCTTTGGA
HUARS2A     759 c....ac... ...a..                  [95%(-)60% similar]  MC:9.32
MUSMDG5                                            [90% similar]        MC:7.46
DHFR-A     2414 ........t. ......ac.. .c.         [72% similar]        MC:6.23
DHFR-B     4353 .....a..a. ......                  [68% similar]        MC:6.23
```

FIG. 7. 377 fragment chromosomal sequence comparisons and analysis. The numbers flanking the sequence are the 5′ start positions from the GenBank data base. The dots represent identical bases. The percent sequence identity and Monte Carlo (MC) simulation index indicate the level of homology, where >6 indicates probable homology, >3 to <6 indicates possible homology, and <3 indicates no homology.

Another potentially interesting *cis*-acting element for which consensus sequences are present is the recognition site for the protein RIP-60. These sequences were initially found in the hamster *DHFR* origin region (16, 23), in which they were speculated to aid in initiation by enhancing DNA bending and potentially mediating the interaction of dispersed elements. Additionally, RIP-60 copurifies with an ATP-dependent helicase, highly suggestive of a role in an initiation complex. A number of transcription factors also have been implicated as initiation factors for eukaryotic DNA replication in viral systems; these include SP1, AP1, NFI (CTF), NFIII (Oct-1), and p53 (20, 21, 36, 47, 57, 58, 62, 67, 76). Consensus sequences for all these elements can be found in the 377 fragment, but their relatively frequent representation in DNA because of their short recognition site sequences makes it difficult to determine their significance.

In summary, we have identified a short fragment in autonomously replicating double-minute chromosomes that complements the replication defect in an EBV-based plasmid derivative and that has been shown to include an initiation region for DNA replication in several contexts, including the normal single-copy chromosomal location, by a nascent-strand PCR assay. Although resolution with this technique is limited, the results appear to be clearly inconsistent with random initiation. Several potentially important sequence elements are present within the fragment. Further application of the assay developed in this work should help in the analysis of the importance of these features within this sequence and others. Only through such detailed comparisons will the mechanisms for the regulation of DNA initiation in mammalian chromosomes be resolved.

ACKNOWLEDGMENTS

We thank Carol Germain (Parke-Davis/Warner Lambert Co.) for the generous gift of pentostatin. We also thank John Yates (Roswell Park Cancer Institute, Buffalo, N.Y.) for suggesting the use of the alkaline extraction approach for extrachromosomal DNA purification and Rod Kellems (Baylor College of Medicine, Houston, Tex.) for many contributions to this project.

This work was supported by grant CD-478 from the American Cancer Society.

REFERENCES

1. **Adhya, S., P. S. Shneidman, and J. Hurwitz.** 1986. Reconstruction of adenovirus replication origins with a human nuclear factor I binding site. J. Biol. Chem. **261**:3339–3346.
2. **Altschul, S. F., W. Grish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
3. **Ambinder, R. F., W. A. Shah, D. R. Rawlins, G. S. Hayward, and S. D. Hayward.** 1990. Definition of the sequence requirements for binding of the EBNA-1 protein to its palindromic target sites in Epstein-Barr virus DNA. J. Virol. **64**:2369–2379.
4. **Anachkova, B., and J. L. Hamlin.** 1989. Replication in the amplified dihydrofolate reductase domain in CHO cells may initiate at two distinct sites, one of which is a repetitive sequence element. Mol. Cell. Biol. **9**:532–540.
5. **Baran, N., A. Lapidot, and H. Manor.** 1991. Formation of DNA triplexes accounts for arrests of DNA synthesis at $d(TC)_n$ and $d(GA)_n$ tracts. Proc. Natl. Acad. Sci. USA **88**:507–511.
6. **Bargonetti, J., P. N. Friedman, S. E. Kern, B. Vogelstein, and C. Prives.** 1991. Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. Cell **65**:1083–1091.
7. **Barker, W. C., and M. O. Dayhoff.** 1972. Detecting distant relationships: computer methods and results, p. 101–110. *In* M. O. Dayhoff (ed.), Atlas of protein sequences and structures. National Biomedical Research Foundation, Washington, D.C.
8. **Benbow, R. M., J. Zhao, and D. D. Larson.** 1992. On the nature of origins of DNA replication in eukaryotes. Bioessays **14**:661–670.
9. **Bergemann, A. D., and E. M. Johnson.** 1992. The HeLa Pur factor binds single-stranded DNA at a specific element conserved in gene flanking regions and origins of DNA replication. Mol. Cell. Biol. **12**:1257–1265.
10. **Brewer, B. J., and W. L. Fangman.** 1987. The localization of replication origins on ARS plasmids in *S. cerevisiae*. Cell **51**:463–471.
11. **Brewer, B. J., and W. L. Fangman.** 1988. A replication fork barrier at the 3′end of yeast ribosomal RNA genes. Cell **55**:637–643.
12. **Burhans, W. C., J. E. Selegue, and N. H. Heintz.** 1986. Isolation of the origin of replication associated with the amplified Chinese hamster dihydrofolate reductase domain. Proc. Natl. Acad. Sci. USA **83**:7790–7794.
13. **Burhans, W. C., J. E. Selegue, and N. H. Heintz.** 1986. Replication intermediates formed during initiation of DNA synthesis in methotrexate-resistant CHOC 400 cells are enriched for sequences derived from a specific amplified restriction fragment. Biochemistry **25**:441–449.
14. **Burhans, W. C., L. T. Vassilev, M. S. Caddle, N. H. Heintz, and M. L. DePamphilis.** 1990. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell **62**: 955–965.
15. **Burhans, W. C., L. T. Vassilev, J. Wu, J. M. Sogo, F. S. Nallaseth, and M. L. DePamphilis.** 1991. Emetine allows identi-

fication of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. EMBO J. **10**:4351–4360.
16. **Caddle, M. S., L. Dailey, and N. H. Heintz.** 1990. RIP60, a mammalian origin-binding protein, enhances DNA bending near the dihydrofolate reductase origin of replication. Mol. Cell. Biol. **10**:6236–6243.
17. **Caddle, M. S., R. H. Lussier, and N. H. Heintz.** 1990. Intramolecular DNA triplexes, bent DNA and DNA unwinding elements in the initiation region of an amplified dihydrofolate reductase replicon. J. Mol. Biol. **211**:19–33.
18. **Carroll, S. M., M. L. DeRose, J. L. Kolman, G. H. Nonet, R. E. Kelly, and G. M. Wahl.** 1993. Localization of a bidirectional DNA replication origin in the native locus and in episomally amplified murine adenosine deaminase loci. Mol. Cell. Biol. **13**:2971–2981.
19. **Chu, G., H. Hayakawa, and P. Berg.** 1987. Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res. **15**:1311–1326.
20. **Cleat, P. H., and R. T. Hay.** 1989. Co-operative interactions between NFI and the adenovirus DNA binding protein at the adenovirus origin of replication. EMBO J. **8**:1841–1848.
21. **Coenjaerts, F. E. J., E. De Vries, G. J. M. Pruijn, W. van Driel, S. M. Bloemers, M. T. van der Lugt, and P. C. van der Vliet.** 1991. Enhancement of DNA replication by transcription factors NFI and NFIII/Oct-1 depends critically on the positions of their binding sites in the adenovirus origin of replication. Biochim. Biophys. Acta **1090**:61–69.
22. **Cornish-Bowden, A.** 1985. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. Nucleic Acids Res. **13**:3021–3030.
23. **Dailey, L., M. S. Caddle, N. Heintz, and N. H. Heintz.** 1990. Purification of RIP60 and RIP100, mammalian proteins with origin-specific DNA-binding and ATP-dependent DNA helicase activities. Mol. Cell. Biol. **10**:6225–6235.
24. **Danna, K. J., and D. Nathans.** 1972. Bidirectional replication of SV40 DNA. Proc. Natl. Acad. Sci. USA **69**:3097–3100.
25. **DePamphilis, M. L.** 1988. Transcriptional elements as components of eukaryotic origins of DNA replication. Cell **52**:635–638.
26. **DePamphilis, M. L.** 1993. Origins of DNA replication in metazoan chromosomes. J. Biol. Chem. **268**:1–4.
27. **Dijkwel, P. A., and J. L. Hamlin.** 1992. Initiation of DNA replication in the dihydrofolate reductase locus is confined to the early S period in CHO cells synchronized with the plant amino acid mimosine. Mol. Cell. Biol. **12**:3715–3722.
28. **DuBridge, R. B., P. Tand, H. C. Hsia, P. Leong, J. H. Miller, and M. P. Calos.** 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. **7**:379–387.
29. **El-Deiry, W., S. E. Kern, J. A. Pientenpol, K. W. Kinzler, and B. Vogelstein.** 1992. Definition of a consensus binding site for p53. Nature Genet. **1**:45–49.
30. **Epner, E., R. A. Rifkind, and P. A. Marks.** 1981. Replication of α- and β-globin DNA sequences occurs during early S phase in murine erythroleukemia cells. Proc. Natl. Acad. Sci. USA **78**:3058–3062.
31. **Frappier, L., and M. O'Donnell.** 1991. Epstein-Barr nuclear antigen 1 mediates a DNA loop within the latent replication origin of Epstein-Barr virus. Proc. Natl. Acad. Sci. USA **88**:10875–10879.
32. **Gahn, T. A., and C. L. Schildkraut.** 1989. The Epstein-Barr virus origin of plasmid replication, *OriP*, contains both the initiation and termination sites of DNA replication. Cell **58**:527–535.
33. **Gale, J. M., R. A. Tobey, and J. A. D'Anna.** 1992. Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells. J. Mol. Biol. **224**:343–358.
34. **Goldberg, G. I., I. Collier, and A. Cassel.** 1983. Specific DNA sequences associated with the nuclear matrix in synchronized mouse 3T3 cells. Proc. Natl. Acad. Sci. USA **80**:6887–6891.
35. **Gunaratne, P. H., and A. C. Chinault.** Unpublished data.
36. **Guo, Z.-S., and M. L. DePamphilis.** 1992. Specific transcription factors stimulate simian virus 40 and polyomavirus origins of DNA replication. Mol. Cell. Biol. **12**:2514–2524.
37. **Handeli, S., A. Klar, M. Meuth, and H. Cedar.** 1989. Mapping replication units in animal cells. Cell **57**:909–920.
38. **Heintz, N. H., and J. L. Hamlin.** 1982. An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. Proc. Natl. Acad. Sci. USA **79**:4083–4087.
39. **Heinzel, S. S., P. J. Krysan, C. T. Tran, and M. P. Calos.** 1991. Autonomous DNA replication in human cells is affected by the size and the source of the DNA. Mol. Cell. Biol. **11**:2263–2272.
40. **Hirota, Y., M. Yamada, A. Nishimura, A. Oka, K. Sugimoto, K. Asada, and M. Takanami.** 1981. The DNA replication origin (*ori*) of *Escherichia coli*: structure and function of the *ori*-containing DNA fragment. Prog. Nucleic Acid Res. **26**:33–48.
41. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–369.
42. **Huberman, J. A., L. D. Spotila, K. A. Nawotka, S. M. El-Assouli, and L. R. Davis.** 1987. The *in vivo* replication origin of the yeast 2 μm plasmid. Cell **51**:473–481.
43. **Huberman, J. A., J. Zhu, L. R. Davis, and C. S. Newlon.** 1988. Close association of a DNA replication origin and an ARS element on chromosome III of the yeast, *Saccharomyces cerevisiae*. Nucleic Acids Res. **16**:6373–6385.
44. **Kagumi, M., and A. Kornberg.** 1984. Replication initiated at the origin (*oriC*) of the *E. coli* chromosome reconstituted with purified enzymes. Cell **38**:183–190.
45. **Kearsey, S.** 1984. Structural requirements for the function of a yeast chromosomal replicator. Cell **37**:299–307.
46. **Kellems, R. E., R. Lin, and A. C. Chinault.** 1992. Adenosine deaminase: a dominant amplifiable genetic marker, p. 207–221. *In* R. E. Kellems (ed.), Gene amplification in mammalian cells: techniques and applications. Marcel Dekker, Inc., New York.
47. **Kern, S. E., K. W. Kinzler, A. Bruskin, D. Jarosz, P. Friedman, C. Prives, and B. Vogelstein.** 1991. Identification of p53 as a sequence-specific DNA-binding protein. Science **252**:1708–1711.
48. **Krysan, P. J., and M. P. Calos.** 1991. Replication initiates at multiple locations on an autonomously replicating plasmid in human cells. Mol. Cell. Biol. **11**:1464–1472.
49. **Krysan, P. J., S. B. Haase, and M. P. Calos.** 1989. Isolation of human sequences that replicate autonomously in human cells. Mol. Cell. Biol. **9**:1026–1033.
50. **Krysan, P. J., J. G. Smith, and M. P. Calos.** 1993. Autonomous replication in human cells of multimers of specific human and bacterial DNA sequences. Mol. Cell. Biol. **13**:2688–2696.
51. **Linskens, M. H. K., and J. A. Huberman.** 1988. Organization of replication of ribosomal DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8**:4927–4935.
52. **Linskens, M. H. K., and J. A. Huberman.** 1990. The two faces of higher eukaryotic DNA replication origins. Cell **62**:845–847.
53. **Ma, C., T.-H. Leu, and J. L. Hamlin.** 1990. Multiple origins of replication in the dihydrofolate reductase amplicons of a methotrexate-resistant Chinese hamster cell line. Mol. Cell. Biol. **10**:1338–1346.
54. **Manor, H., B. S. Rao, and R. G. Martin.** 1988. Abundance and degree of dispersion of genomic $d(GA)_n \cdot d(TC)_n$ sequences. J. Mol. Evol. **27**:96–101.
55. **McWhinney, C., and M. Leffak.** 1990. Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human c-*myc* gene. Nucleic Acids Res. **18**:1233–1242.
56. **Montiel, J. F., C. J. Norbury, M. F. Tuite, M. J. Dobson, J. S. Mills, A. J. Kingsman, and S. M. Kingsman.** 1984. Characterization of human chromosomal DNA sequences which replicate autonomously in *Saccharomyces cerevisiae*. Nucleic Acids Res. **12**:1049–1068.
57. **Mul, Y. M., C. P. Verrijzer, and P. C. van der Vliet.** 1990. Transcription factors NFI and NFIII/Oct-1 function independently, employing different mechanisms to enhance adenovirus DNA replication. J. Virol. **64**:5510–5518.
58. **Murakami, Y., M. Satake, Y. Yamaguchi-Iwai, M. Sakai, M. Muramatsu, and Y. Ito.** 1991. The nuclear protooncogenes c-*jun* and c-*fos* as regulators of DNA replication. Proc. Natl. Acad.

Sci. USA **88**:3947–3951.
59. **Nonet, G. H., S. M. Carroll, M. L. DeRose, and G. M. Wahl.** 1993. Molecular dissection of an extrachromosomal amplicon reveals a circular structure consisting of an imperfect inverted duplication. Mol. Cell. Biol. **15**:543–558.
60. **Palzkill, T. G., and C. S. Newlon.** 1988. A yeast replication origin consists of multiple copies of a small conserved sequence. Cell **53**:441–450.
61. **Peden, K. W., J. M. Pipas, S. Pearson-White, and D. Nathans.** 1980. Isolation of mutants of an animal virus in bacteria. Science **209**:1392–1396.
62. **Pruijn, G. J. M., W. van Driel, and P. C. van der Vliet.** 1986. Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. Nature (London) **322**:656–659.
63. **Reisman, D., and B. Sugden.** 1986. *trans*-Activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr viral nuclear antigen 1. Mol. Cell. Biol. **6**:3838–3846.
64. **Reisman, D., J. Yates, and B. Sugden.** 1985. A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two *cis*-acting components. Mol. Cell. Biol. **5**:1822–1832.
65. **Rowland, P., III, J. Pfeilsticker, and P. A. Hoffee.** 1985. Adenosine deaminase gene amplification in deoxycoformycin-resistant mammalian cells. Arch. Biochem. Biophys. **239**:396–403.
66. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
67. **Santoro, C., N. Mermod, P. C. Andrews, and R. Tjian.** 1988. A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) **334**:218–224.
68. **Su, W., T. Middleton, B. Sugden, and H. Echols.** 1991. DNA looping between the origin of replication of Epstein-Barr virus and its enhancer site: stabilization of an origin complex with Epstein-Barr nuclear antigen 1. Proc. Natl. Acad. Sci. USA **88**:10870–10874.
69. **Sugden, B., K. Marsh, and J. Yates.** 1985. A vector that replicates as a plasmid and can be efficiently selected in B-lymphoblasts transformed by Epstein-Barr virus. Mol. Cell. Biol. **5**:410–413.
70. **Symington, B. E.** 1990. Fibronectin receptor overexpression and loss of transformed phenotype in a stable variant of the K562 cell line. Cell. Regul. **1**:637–648.
71. **Vassilev, L., and E. M. Johnson.** 1989. Mapping initiation sites of DNA replication in vivo using polymerase chain reaction amplification of nascent strand segments. Nucleic Acids Res. **17**:7693–7705.
72. **Vassilev, L., and E. M. Johnson.** 1990. An initiation zone of chromosomal DNA replication located upstream of the *c-myc* gene in proliferating HeLa cells. Mol. Cell. Biol. **10**:4899–4904.
73. **Vassilev, L. T., W. C. Burhans, and M. L. DePamphilis.** 1990. Mapping an origin of DNA replication at a single-copy locus in exponentially proliferating mammalian cells. Mol. Cell. Biol. **10**:4685–4689.
74. **Vassilev, L. T., and M. L. DePamphilis.** 1992. Guide to identification of origins of DNA replication in eukaryotic cell chromosomes. Crit. Rev. Biochem. Mol. Biol. **27**:445–472.
75. **Vaughn, J. P., P. A. Dijkwel, and J. L. Hamlin.** 1990. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. Cell **61**:1075–1087.
76. **Verrijzer, C. P., A. J. Kal, and P. C. van der Vliet.** 1990. The DNA binding domain (POU domain) of the transcription factor oct-1 suffices for stimulation of DNA replication. EMBO J. **9**:1883–1888.
77. **Vidal, M., C. Wrighton, S. Eccles, J. Burke, and F. Grosveld.** 1990. Differences in human cell lines to support stable replication of Epstein-Barr virus-based vectors. Biochim. Biophys. Acta **1048**:171–177.
78. **Virta-Pearlman, V. J.** 1992. Ph.D. thesis. Baylor College of Medicine, Houston, Tex.
79. **Wong, A. K. C., H. A. Yee, J. H. van de Sande, and J. B. Rattner.** 1990. Distribution of CT-rich tracts is conserved in vertebrate chromosomes. Chromosoma **99**:344–351.
80. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A *cis*-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc. Natl. Acad. Sci. USA **81**:3806–3810.
81. **Yates, J. L., N. Warren, and B. Sugden.** 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature (London) **313**:812–815.
82. **Yeung, C.-Y., D. E. Ingolia, C. Bobonis, B. S. Dunbar, M. E. Riser, M. J. Siciliano, and R. E. Kellems.** 1983. Selective overproduction of ADA in cultured mouse cells. J. Biol. Chem. **258**:8338–8345.