JOURNAL OF VIROLOGY, Feb. 1997, p. 1405–1416 Vol. 71, No. 2
0022-538X/97/$04.00+0

A Novel Cellular Site-Specific DNA-Binding Protein Cooperates with the Viral NS1 Polypeptide To Initiate Parvovirus DNA Replication

JESPER CHRISTENSEN,[1]† SUSAN F. COTMORE,[1] AND PETER TATTERSALL[1,2]*

Departments of Laboratory Medicine[1] and Genetics,[2] Yale University School of Medicine, New Haven, Connecticut 06510

Received 1 August 1996/Accepted 15 October 1996

Replication of linear single-stranded parvovirus DNA proceeds by a rolling-hairpin mechanism which generates long, palindromic, duplex concatamers. Processing to monomer length requires initiation from origins of DNA replication located at the 3′ and 5′ ends of each embedded monomer, reactions which can be recapitulated in vitro for minute virus of mice (MVM). To determine which cellular proteins were essential for replication from these origins, S100 extracts from 293S cells were fractionated on phosphocellulose. When recombined, these fractions were able to support replication in vitro, dependent on the viral initiator protein NS1, using plasmid forms of the 5′ origin or the minimal 3′ origin as templates. Fraction P-cell 1 contains two factors, replication protein A (RPA) and proliferating-cell nuclear antigen (PCNA), known to be essential for simian virus 40 replication in vitro. When P-cell 1 was replaced with purified recombinant RPA and PCNA, NS1-mediated MVM replication initiated from the 5′ origin but not from the 3′ origin. The 3′ origin is a 50-bp sequence containing three distinct recognition elements, an NS1 binding site, a site at which NS1 nicks the DNA to generate the priming 3′ OH, and a region containing a consensus activated transcription factor (ATF) binding site. To identify the missing factor(s) for 3′ origin replication, P-cell 1 was fractionated by further chromatography and active fractions were identified by their ability to complement RPA, PCNA, and P-cell 2 for NS1-mediated, origin-specific replication. Gel shift and UV cross-linking analysis of the replication-competent fractions revealed a novel 110-kDa sequence-specific DNA binding protein which recognized the consensus ATF binding site region of the origin and which we have termed parvovirus initiation factor, or PIF. Binding of PIF appears to activate the endonuclease function of NS1, allowing efficient and specific nicking of the 3′ minimal origin under stringent conditions in vitro.

Parvoviruses have small (5-kb), linear, single-stranded DNA genomes which encode two or three overlapping structural polypeptides and a small number of sequence-related, nonstructural proteins involved in genome replication (4). Of the latter, only one, a large pleiotropic nuclear phosphoprotein variously designated NS1 or Rep 68 (or Rep 78), is absolutely required for productive replication in all cell types (20, 26). Their severely limited coding potential means that parvoviruses must rely on diverting the replication machinery of the host cell for their own preferential replication, a dependence which is the more remarkable because they employ a single-strand-specific rolling-hairpin mechanism unique in mammalian cells (15). This mode of replication closely resembles, and probably evolved from, the rolling-circle (RCR) mechanisms previously characterized in certain prokaryotic replicons (22). As in prokaryotic RCR systems, replicon-encoded initiator proteins, in this case NS1 or Rep 68/78, serve as site-specific DNA binding proteins which recognize sequences within specialized viral origins and initiate replication by introducing a single-stranded nick at a specific sequence located close to the core recognition site. This cleavage leaves the initiator protein covalently attached to the 5′ end at the nick site via a phosphotyrosine bond and generates a base-paired 3′ nucleotide which serves as a primer for DNA synthesis.

The negative-sense genome of minute virus of mice MVM is 5,172 nucleotides long and contains 4.8 kb of single-stranded coding sequence bracketed by short, unique terminal palindromes which fold back on themselves to form imperfect hairpin duplexes (2). These complex palindromes play a central role in the replication process, in part because they contain the sequence information required to generate NS1-dependent origins. The 3′ hairpin of MVM assumes a Y-shaped structure in which two internal palindromes form the ears while a third palindrome creates the stem (Fig. 1A). The latter contains a mismatched bubble sequence, where the triplet GAA on the inboard arm is opposed by the dinucleotide GA on the outboard arm, which is important because it appears to regulate protein interactions on the DNA and, in so doing, precludes the direct use of this hairpinned form of the palindrome as an NS1-cleavage substrate (3, 13, 15a). In the stem, the 3′ nucleotide of incoming virion DNA is paired with an internal base (Fig. 1A), creating a DNA primer which allows a host polymerase to synthesize a complementary copy of the internal sequences of the viral genome. This synthesis, which precedes viral gene expression, creates a monomer-length, duplex molecule in which the two strands are covalently continuous at one end, linked by a single copy of the original 3′ telomere. Following expression of the NS1 gene, this duplex monomer is opened and copied by a unidirectional leading-strand-specific replication fork, generating a dimer duplex intermediate in which two unit-length copies of the genome are joined in a head-to-head orientation by a single duplex copy of the origi-

* Corresponding author. Mailing address: Department of Laboratory Medicine, Yale University School of Medicine, 333 Cedar St., New Haven, CT 06510. Phone: (203) 785-4586. Fax: (203) 785-7340.

† Present address: Department for Virology and Immunology, The Royal Veterinary and Agricultural University of Copenhagen, 1870 Frederiksberg C, Denmark.

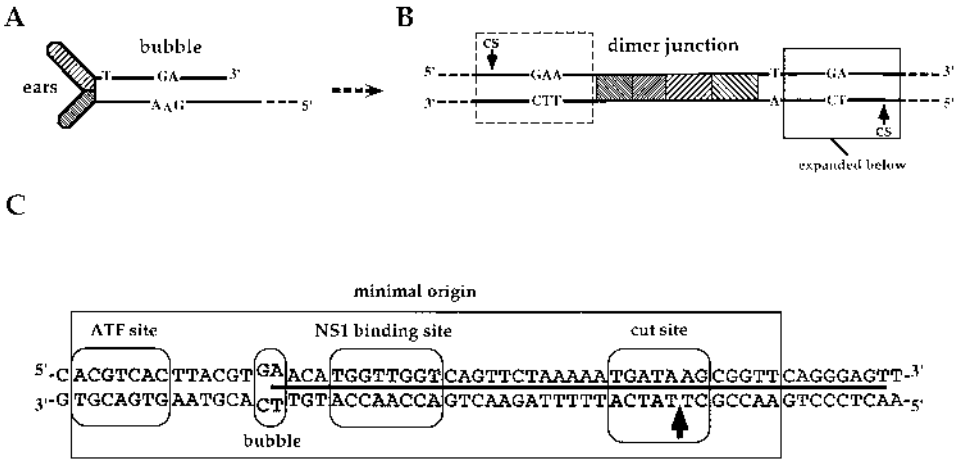

FIG. 1. Formation and organization of the MVM 3′ DNA replication origin. (A) Structure of the left-end hairpin showing the 3′ OH used for priming DNA replication and the mismatched bubble sequence as present in the parental single-stranded viral genome. (B) Organization of the left-end hairpin sequences within the dimer junction generated after replication through the hairpin. Hatched boxes represent the palindromic sequences which were originally folded to give the internal ears in the hairpin form shown in panel A. The boxed sequence (expanded below) represents the minimum active replication origin on the outboard arm (TC sequence) of the dimer junction, while the sequence in the dashed box represents the corresponding origin sequence on the inboard arm (GAA sequence) which is inactive because of the insertion of a single extra nucleotide from the other strand of the bubble sequence. The potential nicking sites on each side of the junction are denoted cs. (C) Sequence of the minimal 3′ replication origin (boxed). The different elements involved in replication are indicated: ATF site, consensus binding site for the family of ATFs; bubble, di- or trinucleotide spacer element described in panel B; NS1 binding site, NS1 core recognition sequence; cut site, the specific sequence necessary for nicking and covalent attachment of NS1. The heavy underline indicates sequences protected by NS1 from DNase I digestion.

nal 3′ palindrome (14). The junction region thus formed contains an active NS1-driven origin (11, 24). Figure 1B illustrates how the various elements of the 3′ hairpin are disposed and copied in the dimer bridge structure to create this site. Genetic mapping experiments showed that the minimal active origin is a 50-bp fragment containing the dinucleotide bubble sequence (generally designated the TC sequence), located on one side of the palindromic symmetry axis (13). An almost identical sequence containing the triplet GAA is located on the other side of the axis but is totally inactive. The active origin (Fig. 1C) contains three distinct recognition elements: the NS1 cut site; the ACCA repeat sequence, which is a high-affinity NS1 binding site (10); and a consensus activated transcription factor (ATF) binding site. The last two elements are separated by the bubble dinucleotide, which is not a specific recognition element, since the actual sequence of the bubble is unimportant, but instead serves as a precise spacer element such that insertion of any third nucleotide here abolishes nicking and hence initiation (13). DNase I protection analysis revealed that NS1 extends over a region of approximately 43 nucleotides, including much of the origin and extending over flanking sequences beyond the nick site, but that this footprint ends abruptly in the middle of the bubble, leaving the region harboring the ATF site unprotected (10). Although NS1 binds readily to both the inactive (GAA) and active (TC) origin sequences in vitro, under physiological assay conditions it could not be shown to nick either sequence (28). Taken together, these observations suggested that to activate the endonuclease function of NS1, an unidentified cellular protein was required which could bind specifically to the consensus ATF site in the origin and interact with NS1 in a precise way across the bubble dinucleotide.

In an attempt to identify this cellular protein, we have fractionated competent replication extracts from 293 cells by procedures originally developed for identifying cellular proteins involved in simian virus 40 (SV40) in vitro replication (35, 42). As in the SV40 system, we have reconstituted NS1-driven replication in vitro by using specific fractions and various purified recombinant proteins. In this report, we describe the isolation of a highly enriched protein fraction which is not required by SV40 but which is absolutely required for NS1-mediated replication initiation at the 3′ origin. This fraction contains a novel human sequence-specific DNA binding protein we have termed PIF (parvovirus initiation factor), which binds to the ATF region of the MVM 3′ minimal origin but is not a known member of the ATF family. In the presence of this semipurified PIF fraction, purified NS1 is now able to introduce a site-specific, single-strand nick into the minimal active MVM 3′ origin under stringent conditions in vitro, leaving NS1 covalently attached to the 5′ end generated at the nick site.

MATERIALS AND METHODS

Chemicals and reagents. Nucleotides and ribonucleotides, chelating Sepharose fast flow and Q-Sepharose were obtained from Pharmacia (Uppsala, Sweden). Poly(dI-dC), initially used as a presumptive noncompetitive inhibitor, was a double-stranded, alternating copolymer poly(dI-dC)-poly(dI-dC) and was obtained from Sigma (St. Louis, Mo.). Phosphocellulose and DE81 paper were obtained from Whatman (Springfield Mill, Maidstone, United Kingdom).

Cells. Human 293S cells, obtained from Nicholas Muzyczka, University of Florida, Gainesville, Fla., were grown at 37°C in spinner bottles containing Joklik-modified minimal essential medium supplemented with 5% newborn calf serum. *Spodoptera frugiperda* (Sf9), "high five" insect cells, and recombinant baculovirus *Autographa californica* nuclear polyhedrosis virus were grown as previously described (9).

Recombinant baculoviruses. cDNAs encoding human ATF 1, ATF 2, and ATF 3 were kindly provided by Tsonwin Hai, Ohio Biotechnology Center, Ohio State University, Columbus, Ohio, and subcloned into the baculovirus transfer vector pBluebacbII (InVitrogen, San Diego, Calif.). Recombinant baculoviruses expressing ATF 1 through 3 were generated after cotransfection with baculovirus DNA and purified essentially as previously described (9). Baculoviruses expressing an amino-terminal histidine-tagged form of the major subunit of replication protein A (His-RPA1) and the two minor subunits (RPA2 and RPA3) were generous gifts from Chris Umbrecht and Tom Kelly, John Hopkins University Medical School, Baltimore, Md. (39, 40). A recombinant baculovirus expressing CREB327 was kindly provided by James Hoeffler, University of Colorado School of Medicine, Denver, Colo. (5), and baculoviruses expressing amino-terminal histidine-tagged Ku p86 and Ku p70 subunits were kindly provided by J. Donald Capra, University of Texas Southwestern Medical School, Dallas, Tex. (30). A recombinant baculovirus expressing MVM NS1 with an amino-terminal histidine tag was the generous gift of David Pintel, University of Missouri Medical School, Columbia, Mo., and Natalie Salome, DKFZ, Heidelberg, Germany. Its use as a source of recombinant NS1 has been described previously (7).

Purification of recombinant polypeptides. Recombinant ATF 1, ATF 2, ATF 3, and CREB327 were expressed by high-multiplicity infection of Sf9 cells or high five cells (for ATF 2) and purified by nuclear extraction essentially as described previously (8). The purity ranged from 10 to 50% as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie brilliant blue staining. All recombinant ATFs were tested for activity by gel mobility shift analysis and were found to bind efficiently and in a site-specific manner to a double-stranded oligonucleotide covering the MVM consensus ATF motif. The three-subunit histidine-tagged RPA complex was expressed by coinfection of Sf9 cells at a multiplicity of infection of 5 for His-RPA1 and RPA2 baculoviruses and a multiplicity of infection of 10 for the RPA3 baculovirus. The infected cells were harvested 44 h postinfection, and the expressed histidine-tagged RPA complex was purified by Ni^{2+}-chelate chromatography, as previously described for the purification of histidine-tagged MVM NS1 (7), and further purified by salt gradient elution from a fast protein liquid chromatography (FPLC) Mono-Q column (HR5/5; Pharmacia, Uppsala, Sweden). The recombinant RPA was approximately 95% pure as judged by SDS-PAGE and Coomassie brilliant blue staining. Human proliferating-cell nuclear antigen (PCNA) was expressed in bacterial strain BL21(DE3)(pT7/PCNA), kindly provided by Bruce Stillman, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., and purified as described by Fien and Stillman (18).

Plasmids. The plasmid DNAs used as templates for the in vitro replication reactions were pL1-2TC and pL1-2GAA, containing the minimal active 3′ MVM origin and the corresponding inactive origin, respectively (13). Plasmid p5′AGA contains a duplex copy of one arm of the 5′ palindrome of MVM and provides an active template for NS1-mediated in vitro replication (13).

Cell extracts and fractionation. S100 extracts from 293S cells were prepared and fractionated essentially as described by Tsurimoto and Stillman (36). To obtain the fractions P-cell 1, P-cell 2 and P-cell 3 described in Fig. 2, S100 extracts were adjusted to 150 mM NaCl and applied to a phosphocellulose column (3 ml of resin per liter of original cell culture) equilibrated in buffer A (25 mM Tris-HCl, 1 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride, 0.01% Nonidet P-40 [NP-40], 1 mM dithiothreitol [DTT], 150 mM NaCl, 10% glycerol). The flowthrough from the column was called fraction P-cell 1, and bound proteins were eluted first with buffer A adjusted to 400 mM NaCl (fraction P-cell 2), and then with buffer A containing 800 mM NaCl (fraction P-cell 3). These fractions were dialyzed against buffer A containing 25 mM NaCl and 20% (wt/vol) sucrose, flash frozen in liquid N_2, and stored at −80°C. The protein concentrations of P-cell 1 (20 mg/ml), P-cell 2 (3.5 mg/ml), and P-cell 3 (1.5 mg/ml) were determined by a Bio-Rad (Hercules, Calif.) protein assay with bovine serum albumin as the standard.

To isolate PIF, P-cell 1 was further fractionated (see Fig. 5A). P-cell 1 was loaded on to a Q-Sepharose column (3 ml of resin per liter of original cell culture) equilibrated in buffer A adjusted to 150 mM NaCl. The column was washed with the same buffer and then eluted with buffer A adjusted to 450 mM NaCl. The eluate was diluted threefold in buffer A minus NaCl, applied to a Zn^{2+}-chelate Sepharose column (1.5 ml of resin per liter of original cell culture), and washed with buffer A containing 450 mM NaCl, and bound proteins were eluted in buffer A adjusted to 150 mM NaCl plus 50 mM imidazole. The eluted fractions were dialyzed, frozen, and stored as described above. The eluate from the Zn^{2+}-chelate Sepharose column was further fractionated on an analytical scale by using an FPLC Mono-Q column (HR5/5) equilibrated in buffer A adjusted to 150 mM NaCl. Bound proteins were eluted with a 40-ml linear gradient in 2-ml fractions from 0.05 to 0.7 M NaCl in buffer A. The fractions were dialyzed and stored as described above. Selected Mono-Q fractions were pooled, concentrated 20-fold with a Centricon-30 concentrator (Amicon, Beverly, Mass.) to a volume of 100 μl, and further analyzed by FPLC gel filtration on a Superose 12 (HR 10/30) column equilibrated in buffer A adjusted to 100 mM NaCl. The resulting fractions were dialyzed and stored as described above.

Replication assays. Replication assays were carried out in the presence of optimized amounts of the various cell fractions or purified proteins and 1 to 2 μg of purified histidine-tagged NS1. Assay mixtures contained 25 mM HEPES-KOH (pH 7.8), 1 mM EGTA, 7 mM $MgCl_2$, 0.5 mM DTT, 4 mM ATP, 0.08 mM CTP, 0.08 mM GTP, 0.08 mM UTP, 0.1 mM dCTP, 0.1 mM dGTP, 0.1 mM dTTP, 0.04 mM dATP, 40 mM creatine phosphate, 10 μg of phosphocreatine kinase, 5 μCi of [α-^{32}P]dATP, and 100 ng of the appropriate DNA template in a total volume of 45 μl. After incubation at 37°C for 2 h, the reactions were terminated by addition of 100 μl of 10 mM Tris-HCl–10 mM EDTA–0.5% SDS and the mixtures were incubated at 50°C for 1 h with 50 μg of proteinase K. Samples were extracted with phenol-chloroform and separated from unincorporated nucleotides by G-50 Sephadex spin-column chromatography. DNA was linearized with *Hin*dIII before electrophoresis on 1% agarose gels. The gels were fixed in 7% trichloroacetic acid, dried, and exposed for autoradiography. If samples were subjected to immunoprecipitation analysis, no proteinase K was added and the samples were frozen if not analyzed immediately. Immunoprecipitations were carried out as previously described (12). Incorporation of [α-^{32}P]dAMP was monitored by spotting 10% of the reaction mixture on to Whatman DE81 paper, followed by three successive washes with 0.5 M Na_2HPO_4, one wash with H_2O, and one wash with ethanol, after which incorporation was quantitated by liquid scintillation counting.

Gel shift mobility and UV cross-linking assays. The oligonucleotide probe was prepared by 3′-end labeling a double-stranded oligonucleotide covering the region of the minimal replication origin containing the consensus ATF motif. This ATF oligonucleotide has the sequence [GATC]TTCACGTAAGTGACGT GATGA, where the 5′ [GATC] sequence is a 4-base 5′ overhang synthesized on each strand to facilitate labeling with Sequenase (U.S. Biochemical, Cleveland, Ohio). DNA binding reactions were carried out in 20 mM HEPES-KOH (pH 7.8)–100 mM NaCl–0.5 mM EDTA–1 mM DTT–10% glycerol–50 μg of bovine serum albumin per ml. The samples were incubated for 10 min at room temperature in the presence of the indicated nonspecific carrier DNA before addition of 50,000 cpm of radiolabeled probe.

For competition assays, unlabeled oligonucleotides were added to the reaction mixture at the same time as the labeled probe. The reaction mixtures were incubated for 30 min at room temperature and then loaded onto 6% native polyacrylamide gels containing 0.25× Tris-borate-EDTA (TBE) which had been preelectrophoresed for 30 min at 180 V. After electrophoresis for 3 h at 180 V, the gels were fixed in 10% acetic acid, dried, and exposed for autoradiography. The double-stranded oligonucleotides used as competitors had the following sequences: scramble oligonucleotide, GATCTAGAGAGTCGATGTATCTG CA; *tar* oligonucleotide, GTATGGTTACTACAAGCAAT; somatostatin CRE, CTAGCTCTCTGACGTCAGGCAATCTCT; and mutant somatostatin CRE, CTAGCTCTCTGACCACAGGCAATCTCT. The last two oligonucleotides were the generous gift of Francois Fuks, DKFZ, Heidelberg, Germany.

For UV cross-linking experiments, binding-reaction mixtures were scaled up twofold. After electrophoresis, the wet gel was subjected to UV irradiation for 15 min and exposed for autoradiography. Shifted complexes were localized, excised, boiled for 10 min in an equal volume of 2× SDS-PAGE loading buffer, and loaded directly into the wells of an SDS-PAGE gel (10% polyacrylamide). After electrophoresis, the gel was fixed, dried, and exposed for autoradiography.

Nicking assays. Nicking-assay mixtures contained 30 mM HEPES-KOH (pH 7.8), 100 mM NaCl, 7 mM $MgCl_2$, 5 mM DTT, 2.5 mM ATP, and 0.1% NP-40. A 50-ng sample of purified baculovirus NS1 was preincubated in sample buffer for 15 min on ice with 1, 2, or 4 μl of Mono-Q PIF concentrate in the presence of 50 μg of nonspecific scramble oligonucleotide competitor DNA per ml. A 1.5-ng sample of a 95-bp DNA fragment containing the active MVM 3′ origin from pL1-2 TC or the inactive origin from pL1-2 GAA, 3′-end labeled with ^{32}P-dATP, was added to give a final volume of 10 μl, and the reaction mixture was incubated at 37°C for 1 h. Reactions were stopped by the addition of an equal volume of 10 mM Tris-HCl (pH 7.5)–10 mM EDTA–1% SDS. Where indicated, aliquots of the products were incubated with 50 μg of proteinase K per ml at 60°C for 1 h.

For immunoprecipitation, aliquots were diluted into immunoprecipitation buffer (150 mM NaCl, 20 mM Tris-HCl [pH 8.0], 2 mM EDTA, 1% NP-40) containing 2% SDS and incubated at 60°C for 30 min. They were then diluted with immunoprecipitation buffer alone to a final SDS concentration of 0.25% and precipitated as previously described (12). All samples were diluted into a mixture containing 0.2% SDS, 10 mM EDTA, 10 mM DTT, 7% glycerol, and 0.01% bromophenol blue and heated to 60°C for 20 min prior to electrophoresis, and the samples to be melted were then heated to 100°C for 5 min. The samples were electrophoresed for 2 h at 200 V through 7% nondenaturing polyacrylamide gels (39:1 acrylamide/bisacrylamide ratio) which had been preelectrophoresed for 1 h at 200 V in 0.5× TBE containing 0.1% SDS.

RESULTS

Purified recombinant RPA and PCNA will substitute for fraction P-cell 1 in replication assays with a template from the MVM 5′ origin. NS1 initiates DNA replication from both the 3′ and 5′ origin sequences of MVM, nicking the DNA at a specific site and becoming covalently attached to the newly liberated 5′ nucleotide during the process. However, the structure and sequence of the 3′ and 5′ origins are very different, and while both have cognate NS1 binding sites, the ATF consensus sequence depicted in Fig. 1C, which is absolutely required for initiation at the 3′ origin, is replaced in its 5′ counterpart by a less well-defined sequence. A further striking difference between the two origins is the segregation of active and inactive forms of the 3′ origin sequence, outlined in Fig. 1A and B, which does not occur at the 5′ origin. We were interested in determining whether each of these origins had different biochemical requirements for function in vitro.

Using fractionation procedures initially developed for identifying cellular factors required for SV40 DNA replication (36, 42), we prepared cytosolic extracts from human 293 cells containing leached nuclear proteins and fractionated the extracts by chromatography on phosphocellulose. This generated the three fractions outlined in Figure 2; P-cell 1, which contained proteins and nucleic acids that were unable to bind to the

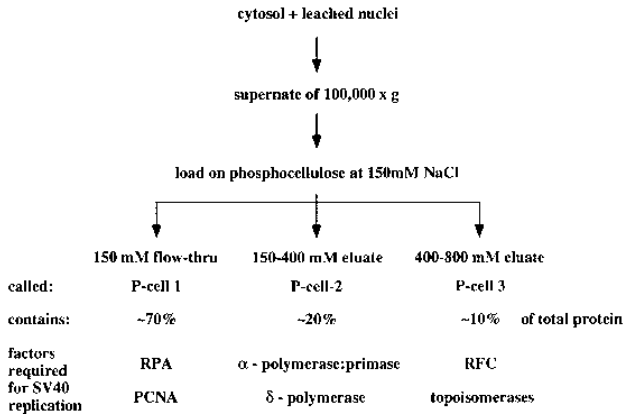

FIG. 2. Properties of phosphocellulose fractions from an S293 cell S100 extract. S100 extracts from S293 cells were separated into three fractions, designated P-cell 1, P-cell 2, and P-cell 3, by chromatography on phosphocellulose. The only factors in P-cell 1 which have been shown to be necessary for SV40 DNA replication in vitro are RPA and PCNA. P-cell 2 contains α- and δ-polymerases, and P-cell 3 contains RFC and topoisomerases (36).

column in 150 mM NaCl (approximately 70% of the total protein); P-cell 2, containing material which eluted with 400 mM NaCl (approximately 20% of the total protein); and P-cell 3, containing material which eluted with 800 mM NaCl (approximately 10% of the total protein). Fraction P-cell 1 is known to contain two factors which are essential for SV40 replication, RPA, a single-stranded DNA-binding protein, and PCNA, a processivity factor required by DNA polymerase δ. Fraction P-cell 2 is known to contain both α-polymerase:primase and δ-polymerase, while P-cell 3 contains topoisomerases and RFC (replication factor C), an elongation factor required for loading PCNA (37). To determine which of these fractions contained cellular factors necessary for initiating and supporting authentic MVM replication, we used combinations of different fractions, together with purified recombinant NS1, to reconstitute in vitro replication assays. Recombinant proteins were expressed in and purified from bacteria (PCNA) or insect (RPA) cells as described in Materials and Methods and titrated to fully complement fractions P-cell 2 and P-cell 3 in a recombinant T-antigen-driven SV40 in vitro replication assay with a plasmid, pUC-HS, which contains the SV40 origin of DNA replication, as the substrate (34).

Initially, we were interested to see whether PCNA and RPA could substitute for P-cell 1 in replication assays primed with a cloned form of the 5′ origin. Replication was quantitated by incorporation of [^{32}P]dAMP into DE81 paper-selectable high-molecular-weight DNA, and replication products were analyzed by gel electrophoresis following digestion with the restriction endonuclease *Hin*dIII, which cuts once in the template plasmid and thus reduces the products to the size of linear input plasmid. Figure 3, lanes 1 and 2, shows that the S100 extract supports abundant replication from the 5′ origin in a strictly NS1-dependent fashion. Furthermore, we found that a combination of fractions P-cell 1, P-cell 2, and P-cell 3 could reconstitute this NS1-driven replication (Fig. 3, lanes 3 and 4), and that each fraction was required (lanes 5 to 7). While neither PCNA nor RPA alone could support efficient replication in the absence of P-cell 1 (lanes 8 and 9), both of these recombinant factors together could effectively reconstitute the system (lane 10). The newly replicated DNA could also be immunoprecipitated efficiently with anti-NS1 serum after SDS denaturation (data not shown), strongly indicating

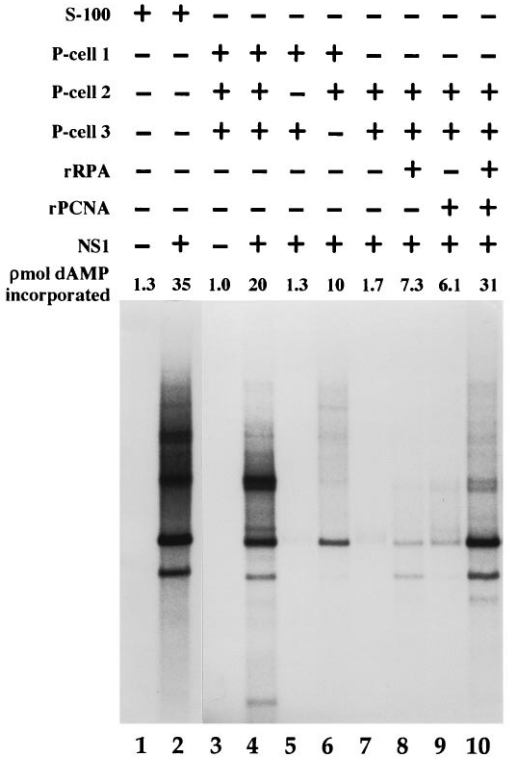

FIG. 3. Reconstitution of MVM replication from the viral 5′ origin. Replication reactions were assembled, as indicated at the top of the figure, with combinations of phosphocellulose fractions P-cell 1, P-cell 2, P-cell 3, purified recombinant RPA (rRPA), purified recombinant PCNA (rPCNA), and purified viral NS1. Plasmid p5′AGA, providing an active 5′ origin, was used as a replication substrate. S100 represents the replication in unfractionated extracts. Each of the phosphocellulose fractions and purified proteins were titrated to achieve optimal replication. Thus, the standard assay mixture consisted of 4 μl of the phosphocellulose fraction indicated, 0.5 μg of rRPA, and 0.1 μg of rPCNA. The protein concentrations of the phosphocellulose fractions, amounts of DNA substrate, viral NS1, and assay conditions are described in Materials and Methods. The number of picomoles of dAMP incorporated into the replication substrate is indicated.

that this was authentic NS1-mediated viral DNA replication and that it requires both RPA and PCNA.

Replication of the MVM 3′ origin requires RPA, PCNA, and an additional factor from fraction P-cell 1. For studies with the 3′ origin, an additional control for authenticity can be obtained by comparing DNA replication initiation on two different DNA substrates. One of these, plasmid pL1-2 TC, contains the minimal active 3′ origin, while the other, plasmid pL1-2 GAA, contains an MVM sequence which is almost identical to the origin but does not support NS1-dependent initiation. pL1-2 GAA is inactive because a triplet, GAA, replaces the dinucleotide, TC, in a critical spacer element called the bubble sequence, as depicted in Fig. 1. Again, replication occurring in each reaction was quantitated by measuring the incorporation of [^{32}P]dAMP by DE81 assay, and linearized replication products were analyzed by gel electrophoresis.

Replication from the 3′ origin could be reconstituted with the three P-cell fractions and appeared to be entirely NS1-dependent and specific for the active form of the origin (Fig. 4, lanes 1 to 3). In contrast to the results with the 5′ origin, however, P-cell 3 appears to be dispensable for this DNA replication (lane 4), while P-cell 1 and P-cell 2 were absolutely required (lanes 5 and 6). While this observation might mean that RFC, predominantly present in P-cell 3, is not required for

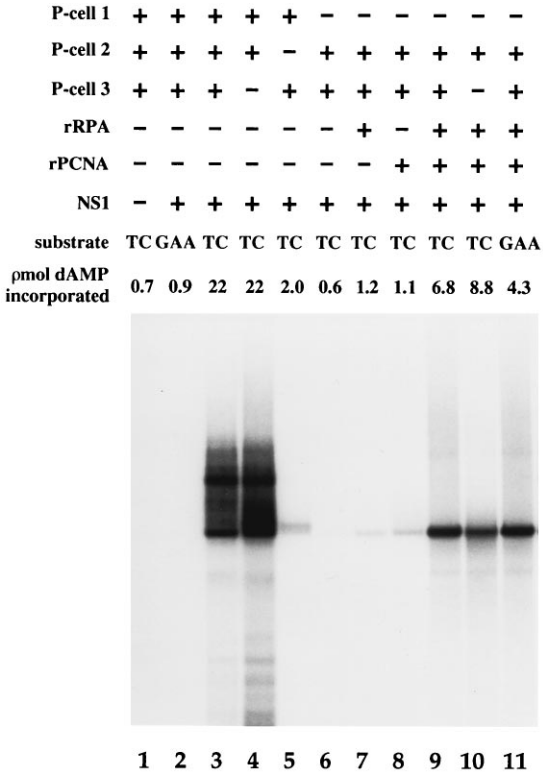

FIG. 4. Reconstitution of MVM replication from the viral 3′ origin. Replication reactions were assembled, as indicated at the top of the figure, with combinations of phosphocellulose fractions P-cell 1, P-cell 2, P-cell 3, purified recombinant RPA (rRPA), purified recombinant PCNA (rPCNA), and purified viral NS1. Plasmid pL1-2 TC, containing the minimal active 3′ origin, was used as the replication-positive template. Plasmid pL1-2 GAA, which is inactive due to the extra single-base insertion (described in the legend to Fig. 1), served as a negative control. The different phosphocellulose fractions and purified proteins were titrated to achieve optimal replication. A standard assay mixture consisted of 4 μl of the phosphocellulose fraction indicated, 0.5 μg of rRPA, and 0.1 μg of rPCNA. Protein concentrations of the phosphocellulose fractions, amounts of DNA substrate, viral NS1, and assay conditions are described in Materials and Methods. The number of picomoles of dAMP incorporated into the replication substrate is indicated.

MVM replication, it should be pointed out that parvovirus DNA replication is a leading-strand specific process, probably requiring δ-polymerase and all of its accessory factors. Thus, it is quite likely that either there is sufficient residual RFC in the other fractions to support efficient synthesis or RFC is not necessary under the relatively low-salt reaction conditions used here. This aspect of the enzymology of parvovirus DNA replication is the subject of further study. Since no replication of the inactive pL1-2 GAA plasmid was observed with any combination of these fractions (Fig. 4, lane 2, and data not shown), initiation in this system is origin specific and there appeared to be very little nonspecific DNA repair synthesis.

In further contrast to our findings with the 5′ origin, when recombinant RPA and PCNA were used to supplement fractions P-cell 2 and 3 in replication assays primed with the active 3′ origin, we observed only a limited increase in [^{32}P]dAMP incorporation when both were present (Fig. 4, lanes 7 to 9), and this was not origin specific, since the control plasmid, pL1-2 GAA, evoked as much synthetic activity as did the authentic substrate pL1-2 TC (compare lanes 9 and 11). Moreover, the labeled replication products synthesized in these reactions could not be immunoprecipitated with anti-NS1 sera, indicating that NS1 had not become covalently attached to the DNA during initiation (data not shown). This enhanced nonspecific DNA repair activity is also independent of P-cell 3 (lanes 9 and 10) and appears to be driven by the nonspecific nicking activity of contaminants present in the NS1 preparations which are normally suppressed in the presence of the complex P-cell 1 fraction.

We interpret these results to mean that fraction P-cell 1 contains a factor(s), in addition to RPA and PCNA, which is required for initiation from the 3′ origin and that this factor mediates the specificity of initiation from the active rather than the inactive form of this origin. Since the experiments illustrated in Fig. 3 showed that replication at the 5′ origin was reconstituted to the level obtained with unfractionated P-cell 1 by the addition of just rRPA and rPCNA, the P-cell 1 factor(s) required for replication from the 3′ origin appears not to be required for initiation at the 5′ site. The further characterization of this 3′ origin specificity factor(s) is the focus of the remainder of this paper.

Identification of a DNA binding activity in P-cell 1 which cofractionates with the factor required for replication of the MVM 3′ origin. Initially, we performed a number of pilot fractionation experiments testing different chromatography resins for their ability to retain and subsequently release a protein fraction which would complement recombinant NS1, RPA, PCNA, and P-cell 2 in the in vitro replication assay. Using this strategy, we eventually developed a purification scheme, outlined in Fig. 5A, which involves batch elution from Q Sepharose and zinc metal-chelate Sepharose followed by salt gradient elution from an FPLC Mono-Q column and gel filtration through an FPLC Superose 12 column. While these fractions did not increase dramatically in specific activity with purification, they maintained their specificity for the active versus inactive forms of the 3′ origin (Fig. 5B) and decreased in protein complexity (Fig. 5C).

The minimal 3′ replication origin of MVM depicted in Fig. 1 is a 50-bp sequence containing an NS1 nick site, an NS1 binding site, and a consensus ATF binding sequence which is separated from the NS1 binding site by a 2-base spacer element called the bubble sequence. Since mutating the ATF consensus or inserting any third base into this spacer severely impaired or eliminated initiation, we speculated that a cellular site-specific DNA binding protein was required to bind to the ATF consensus and activate the endonuclease function of NS1 by interacting with it in a sterically precise manner across the bubble sequence. Since this putative cellular DNA binding protein was an obvious candidate for the P-cell 1 factor required for replicating templates containing the 3′ origin, we screened column fractions both for their ability to support origin-specific replication (Fig. 5B) and for site-specific DNA binding activity directed toward the ATF consensus sequence from the viral origin (Fig. 6 and 7).

To screen for DNA binding activity directed toward the ATF consensus region, we tested fractions for their ability to bind a ^{32}P-end-labeled 25-bp oligonucleotide containing this sequence in a gel mobility shift assay. This oligonucleotide is referred to as the ATF oligomer. As a control, we used a second oligonucleotide, referred to as the scramble sequence, which had the same nucleotide composition and the same terminal sequences as the ATF oligomer but in which the internal sequence was randomized. Using these oligonucleotides, we found (data not shown) that fractions eluting from Q Sepharose and Zn^{2+}-Sepharose, which were able to complement the P-cell 2, RPA, PCNA, and NS1 mixture for origin-specific replication (Fig. 5 and data not shown), also contained several nonspecific DNA binding activities and one major site-specific activity which resisted competition with the scramble

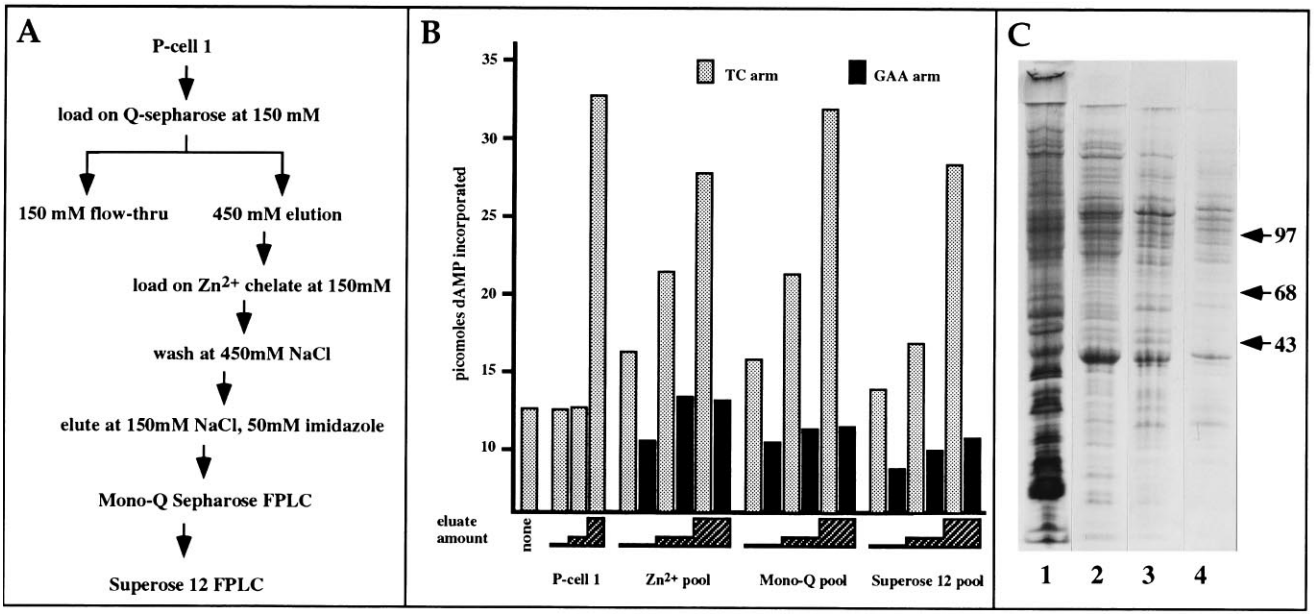

FIG. 5. Subfractionation of P-cell 1. (A) Fractionation scheme of cell extracts used for the identification of the parvovirus initiation factor (PIF) in P-cell 1 by purification on Q-Sepharose, Zn^{2+} metal-chelate Sepharose, Mono-Q FPLC, and gel filtration on Superose 12 FPLC. (B) Specificity of P-cell 1 subfractions for replication of pL1-2 TC versus pL1-2 GAA. Portions (0.5, 1.5, or 4.5 μl) of fraction P-cell 1 (20 mg of protein per ml), portions (1.5, 4.5, or 13.5 μl) of Zn^{2+} Sepharose eluate (6 mg per ml), portions (1.5, 4.5, or 13.5 μl) of Mono-Q pool (4 mg per ml), and portions (1.5, 4.5, or 13.5 μl) of Superose 12 replication complementing fractions (3 mg per ml) were titrated into replication assay mixtures containing a constant amount of P-cell 2 (8 μl), 0.5 μg of rRPA, 0.1 μg rPCNA, 1 μg of NS1, and pL1-2 TC as the substrate for replication. Parallel reactions containing pL1-2 GAA as a substrate served as controls for nonspecific replication. The protein concentrations of the different fractions were determined after SDS-PAGE gel analysis and silver staining followed by gel-scanning quantification. Assay conditions and measurement of incorporation were performed as described in Materials and Methods. (C) Silver-stained SDS-PAGE analysis of samples from the P-cell 1 fraction (lane 1), Zn^{2+} Sepharose eluate (lane 2), Mono-Q pool (lane 3), and Superose 12 pool (lane 4). The amounts of each fraction loaded were adjusted to be equivalent for replication activity, as described above for panel B.

oligonucleotide or with salmon sperm DNA but which could be inhibited effectively with poly(dI-dC).

The eluate from the Zn^{2+}-chelate Sepharose column was then fractionated by salt gradient chromatography on Mono-Q Sepharose. Fractions were analyzed for their ability to complement RPA, PCNA, P-cell 2, and NS1 in an origin-specific replication assay, and fractions from this peak region were then tested for their ability to bind the ATF oligonucleotide in the presence and absence of poly(dI-dC). In the presence of poly(dI-dC), a series of retarded bands were observed (Fig. 6A) which were predominantly nonspecific. It was subsequently shown that the major high-mobility band in this figure resulted from the binding of Ku antigen (human helicase II), which binds with high affinity to the blunt ends of all double-stranded oligonucleotides (38). However, in the absence of poly(dI-dC), an additional major band appeared, designated PIF in Fig. 6B. This activity was most apparent in fractions 7 through 12 from this elution, which were also enriched for the factor required to replicate plasmid pL1-2 TC specifically. Fractions showing the least contaminating Ku activity (fractions 9 through 12) were then pooled, concentrated by ultrafiltration, and titrated again for their ability to support origin-specific replication (Fig. 5B). This material was then applied to a Superose 12 FPLC gel filtration column, and fractions were again assayed for their ability to complement replication and to bind specifically to the ATF oligomer. Once again, these two specific activities were found to coelute (Fig. 7B and C). Replication assays require relatively high protein concentrations, so that by this stage in the fractionation procedure, individual fractions showed relatively little activity. However, after concentration by ultrafiltration, a pool of fractions 10 through 13 from Fig. 7 was able to replicate pL1-2TC specifically and in a concentration-dependent manner (Fig. 5B). Since the major poly(dI-dC)-resistant band-shifting activity (designated Ku) elutes later than the PIF peak, this suggests that the complex responsible for this nonspecific activity is slightly smaller than the MVM replication factor. Ku is a heterodimer of 156 kDa formed between polypeptide chains of approximately 70 and 86 kDa (30, 38).

DNA binding specificity of PIF. To explore the nature of the DNA binding activities directed against the ATF oligomer, we carried out competitive gel retardation assays with a variety of specific and nonspecific inhibitors. Competition with a 1,000-fold molar excess of various unrelated double-stranded oligonucleotides failed to inhibit the low-mobility PIF band-shifting activity observed when the ^{32}P-labeled ATF oligomer was used as a probe (Fig. 8A, lanes 1 and 2). However, binding was efficiently competed by the unlabeled ATF oligonucleotide and by poly(dI-dC) (lanes 3 and 4). Salmon sperm DNA also failed to inhibit binding (lane 5), and, using this as a nonspecific competitor, we were able to titrate the specific inhibitors. Both 10 and 50 ng of the cold ATF oligonucleotide effectively eliminated PIF binding activity (lanes 6 and 7), while equivalent amounts of poly(dI-dC) had a similar, albeit less profound, effect (lanes 8 and 9). Single-stranded M13 DNA failed to inhibit binding (lane 10).

In contrast, the faster-migrating DNA-protein complex, which we designate Ku in Fig. 6 through 8, did not bind in a sequence-specific fashion to the ATF probe. This is shown by the fact that its binding could be inhibited with a variety of unrelated double-stranded oligonucleotides, especially if these had blunt ends (Fig. 8A, compare lane 5 with lanes 1 and 2). Since recourse to the literature suggested that such activity might be due to Ku antigen, we explored the effects of poly-

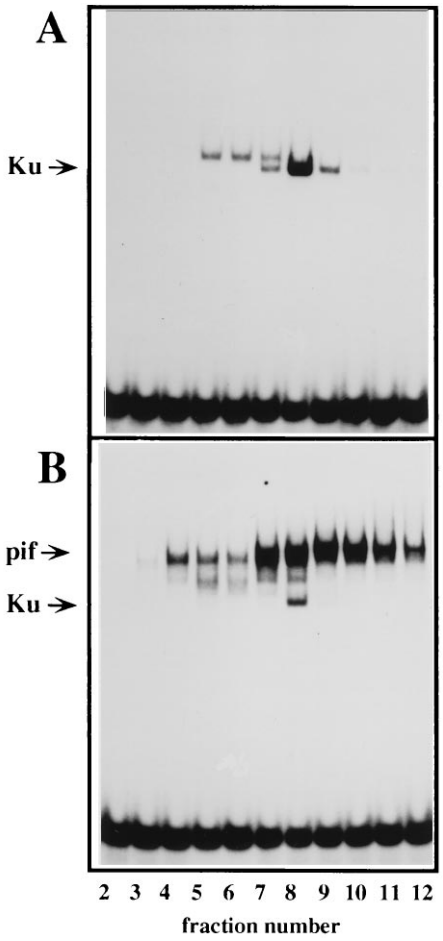

FIG. 6. Fractionation of Zn^{2+} Sepharose eluate by salt gradient elution on Mono-Q. A 2-ml portion of Zn^{2+} Sepharose eluate (6 mg of protein per ml) was fractionated by FPLC on Mono-Q, eluting with a linear gradient from 0.05 to 0.7 M NaCl. Fractions (2 ml) were collected and analyzed for DNA binding to ^{32}P-labeled MVM ATF double-stranded oligonucleotide by mobility shift analysis, in the presence (A) or absence (B) of 1 μg of poly(dI-dC). Band shifts due to the Ku autoantigen heterodimer are marked Ku.

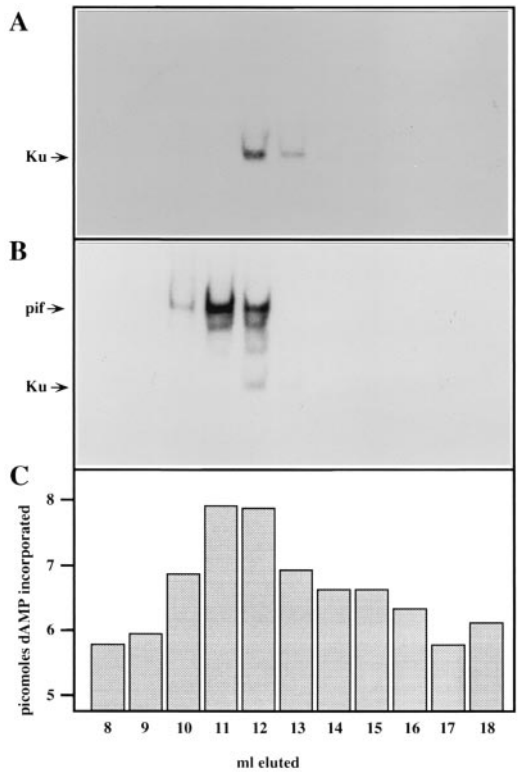

FIG. 7. Fractionation of selected Mono-Q fractions by gel filtration on Superose 12. (A and B) Mono-Q fractions 9 through 12 (from Fig. 6) were pooled, concentrated by ultrafiltration, and fractionated by FPLC on Superose 12. Fractions (1 ml) were collected and analyzed for DNA binding to ^{32}P-labeled MVM ATF double-stranded oligonucleotide by mobility shift analysis in the presence (A) or absence (B) of 1 μg of poly(dI-dC). (C) Fractions containing PIF activity were identified by their ability to complement DNA replication in the presence of NS1, RPA, PCNA, P-cell 2, and pL1-2 TC as a substrate for replication.

clonal sera directed against purified, bacterially expressed Ku 70 and 86 subunits. When added to preformed complexes, serum directed against the 70-kDa subunit was unable to influence the migration of the nonspecific DNA-protein complex but serum directed against the 86-kDa subunit was able to supershift a proportion of the product in a concentration-dependent manner (Fig. 8B, lanes 1 through 5). Recombinant Ku was then purified from insect cells coinfected with baculovirus vectors expressing the two separate polypeptide subunits. Recombinant Ku was able to band shift the ATF probe, giving a complex which comigrated with the putative Ku band in the presence of poly(dI-dC) (lanes 6 and 7). Ku antigen purified from 293 cells by DNA affinity chromatography (kindly provided by Edward Kuff, National Cancer Institute, Bethesda, Md.) gave a band of the same mobility with this probe (data not shown). When poly(dI-dC) was replaced by a specific oligonucleotide competitor (lane 8), recombinant Ku also gave a second band, which had a slightly higher mobility than the PIF complex. This band appears to be due to the binding of more than one Ku complex per oligonucleotide, since if all nonspecific competitor DNA was removed from the reaction mixture, this became the predominant band (not shown). When recombinant Ku complexes were exposed to polyclonal anti-Ku sera, antibodies directed against the 70-kDa subunit were once again unable to shift the complex, while sera directed against the 86-kDa polypeptide gave a supershifted band (lanes 9 and 10). Neither purified Ku from 293 cells nor recombinant Ku purified from insect cells was able to complement RPA, PCNA, P-cell 2, and NS1 in the replication assay, effectively ruling out the possibility that Ku itself is the P-cell 1 factor essential for MVM replication (data not shown).

We have previously shown by mutagenesis that an intact consensus ATF motif within the viral 3′ origin is absolutely required for DNA replication initiation. This suggested that PIF might well be a member of this family of transcription factors (21). To explore this possibility, we compared the binding of PIF and several recombinant ATFs to the MVM ATF motif. Although recombinant ATFs 1, 2, and 3 were able to shift the MVM probe, none of these complexes comigrated with the PIF product, and the PIF complex could not be supershifted by a variety of monoclonal and polyclonal antibodies which did shift the various ATF complexes (data not shown). Recombinant CREB 327 did give a band which approximately comigrated with PIF (Fig. 8C, lane 8), but antisera against CREB 327 did not supershift the PIF complex (data not shown), and when we compared the specificity of these two interactions by using competitor oligonucleotides containing wild-type or mutant cyclic AMP responsive element (CRE) sites, the results were substantially different. As expected, the

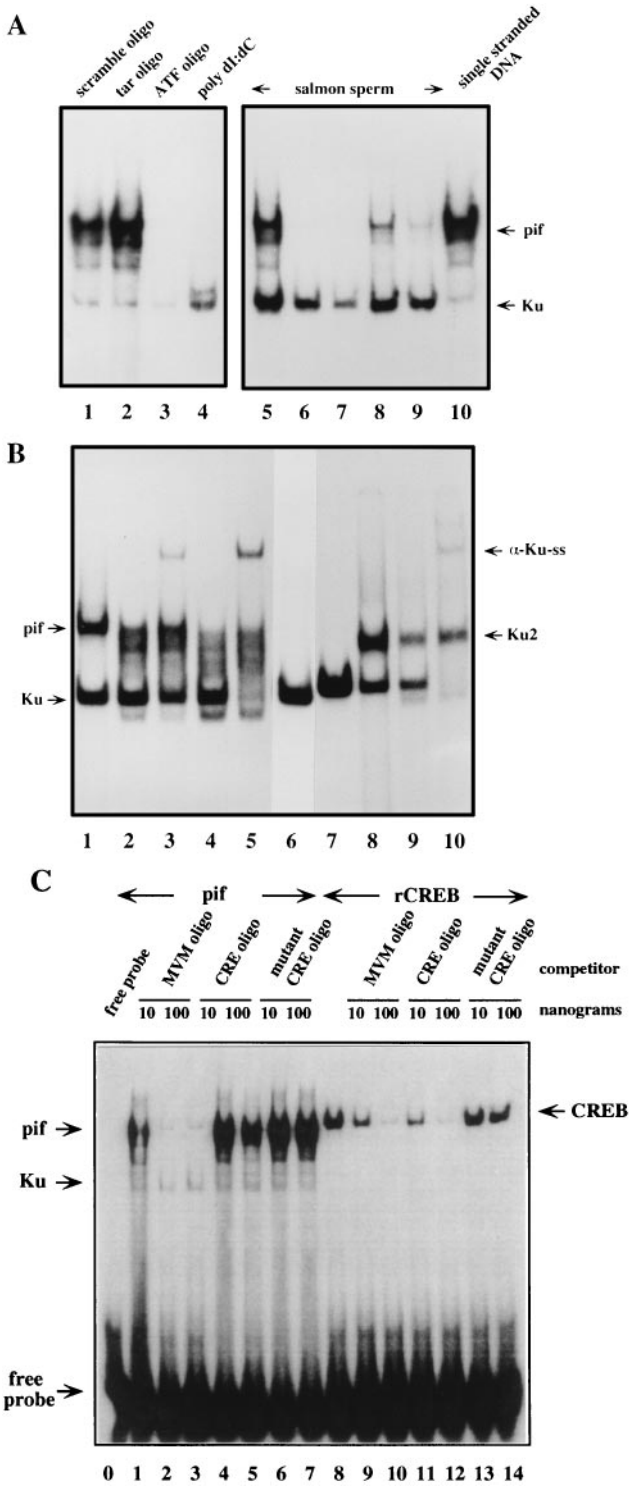

FIG. 8. Mobility shift analysis of PIF DNA binding. (A) DNA binding to ^{32}P-labeled MVM ATF double-stranded oligonucleotide was analyzed in the presence of 1 μg of the DNA indicated above each lane. Lanes: 1, PIF binding in the presence of the scramble double-stranded oligonucleotide; 2, PIF binding in the presence of the *tar* double-stranded oligonucleotide; 3, PIF binding in the presence of the MVM ATF double-stranded oligonucleotide; 4, PIF binding in the presence of poly(dI-dC); 5 through 9, PIF binding in the presence of 1 μg of salmon sperm DNA and the presence of 10 ng (lane 6) or 50 ng (lane 7) of MVM ATF double-stranded oligonucleotide or 10 ng (lane 8) or 50 ng (lane 9) of poly(dI-dC); 10, PIF binding in the presence of 1 μg of single-stranded M13 DNA. (B) Binding to the ^{32}P-labeled MVM ATF double-stranded oligonucleotide, supershift analysis of the observed nonspecific DNA binding activity, and comparison to recombinant Ku antigen. Lanes: 1, 4 μl of Zn^{2+} Sepharose eluate in the presence of 1 μg of *tar* double-stranded oligonucleotide; 2, 4 μl of Zn^{2+} Sepharose eluate in the presence of 1 μg of *tar* double-stranded oligonucleotide and 1 μl of polyclonal antibody to the p70 Ku subunit; 3, 4 μl of Zn^{2+} Sepharose eluate in the presence of 1 μg of *tar* double-stranded oligonucleotide and 1 μl of polyclonal antibody to the p86 Ku subunit; 4, 4 μl of Zn^{2+} Sepharose eluate in the presence of 1 μg of *tar* double-stranded oligonucleotide and 3 μl of polyclonal antibody to the p70 Ku subunit; 5, 4 μl of Zn^{2+} Sepharose eluate in the presence of 1 μg of *tar* double-stranded oligonucleotide and 3 μl of polyclonal antibody to the p86 Ku subunit; 6, 4 μl of Zn^{2+} Sepharose eluate in the presence of 1 μg of poly(dI-dC); 6, 50 ng of purified recombinant Ku in the presence of 1 μg of poly(dI-dC); 7, 50 ng of purified recombinant Ku in the presence of 1 μg of double-stranded *tar* oligonucleotide; 8, 50 ng of purified recombinant Ku in the presence of 1 μg of double-stranded *tar* oligonucleotide and 1 μl of polyclonal antibody to the p70 Ku subunit; 9, 50 ng of purified recombinant Ku in the presence of 1 μg of double-stranded *tar* oligonucleotide and 1 μl of polyclonal antibody to the p86 Ku subunit. Ku2 denotes the position of the oligo with a Ku heterodimer bound to both ends, and α-Ku-ss indicates the position of the supershifted Ku band. (C) Comparison of PIF and rCREB binding specificity by inhibition of binding to the ^{32}P-labeled MVM ATF double-stranded oligonucleotide by double-stranded oligonucleotides containing a consensus or mutant CRE site from the human somatostatin promoter. Lanes containing PIF or rCREB are indicated at the top of the figure, and the amount and type of competitor are indicated above each lane. The sequences of the unlabeled competitors are listed in Materials and Methods.

ATF oligomer from the MVM origin was able to compete for the binding of both CREB and PIF to a ^{32}P-labeled form of the same ATF probe (lanes 1 to 3 and 8 to 10). In contrast, neither wild-type or mutant CRE oligonucleotides were able to compete effectively for PIF activity (lanes 4 to 7), while the wild-type but not the mutant CRE oligonucleotide could compete for the binding of recombinant CREB 327 (lanes 11 to 14). Thus, CREB and PIF do not show the same DNA binding specificity, and authentic CRE sequences do not inevitably interact with PIF, suggesting that the latter is unlikely to be a member of the ATF family of DNA binding proteins.

Purified recombinant ATFs 1, 2, and 3 and CREB 327 also failed to complement RPA, PCNA, P-cell 2, and NS1 in in vitro replication assays, and when added to competent replication extracts, they actually inhibited, rather than potentiated, NS1-mediated replication (data not shown). Thus, it appears that the cellular factor required to support the replication of the MVM 3′ origin is not ATF 1, 2, or 3 or CREB 327, although these factors do indeed bind efficiently to the MVM 3′ origin.

UV cross-linking experiments suggest that PIF has a molecular mass of approximately 110 kDa. To help identify the proteins responsible for the various DNA binding activities, mobility shift assay gels were UV irradiated to induce covalent cross-linking of the ^{32}P-labeled probe to the associated protein. These cross-linked complexes were then excised from the gel and analyzed by SDS-PAGE (Fig. 9). Protein fractions from both the Zn^{2+}-Sepharose and Mono-Q peaks gave a major cross-linked band of approximately 110 kDa (Fig. 9, lanes 1 and 2) and an additional band at about 200 kDa, the position expected of a cross-linked dimer of the 110-kDa species. A further indistinct and somewhat variable band was observed at around 90 kDa, with no smaller species evident. These data support our contention that PIF is not a member of the ATF family, since all known members of this family have molecular masses of 55 kDa or less. If the bands on the mobility shift gels were allowed to migrate further prior to cross-linking, the 110-kDa band remained the predominant species, while the indistinct 90-kDa band was substantially diminished, suggesting that it is simply a contaminant or a breakdown product of the authentic molecule (data not shown).

The Ku band generated in the presence of poly(dI-dC) from the same Zn^{2+}-Sepharose and Mono-Q fractions gave a single

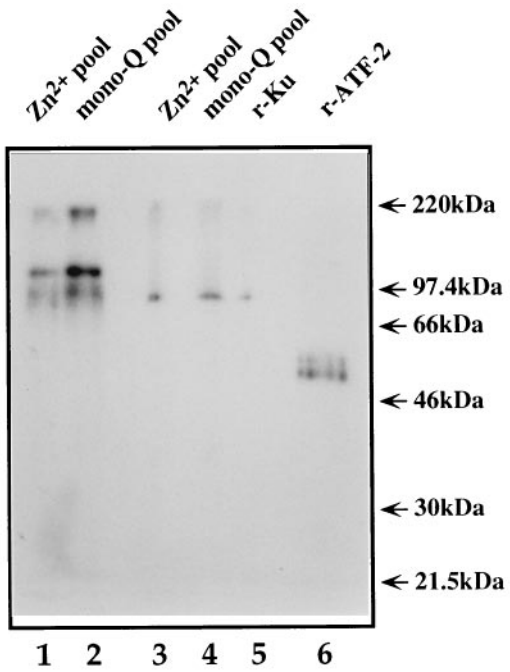

FIG. 9. UV cross-linking analysis of proteins binding to the MVM ATF oligonucleotide. The DNA binding activities observed for Mono-Q or Zn^{2+}-Sepharose fractions, in mobility shift assays performed in the presence of tar oligonucleotide (lanes 1 and 2) or poly(dI-dC) (lanes 3 through 6), were UV irradiated to induce covalent cross-linking of the ^{32}P-labeled probe to the associated protein. These cross-linked complexes were then excised from the gel and analyzed by SDS-PAGE. Lanes: 1, cross-linked PIF (Zn^{2+}-Sepharose fraction); 2, cross-linked PIF (Mono-Q pool); 3, cross-linked Ku (Zn^{2+}-Sepharose fraction); 4, cross-linked Ku (Mono-Q pool); 5, cross-linked recombinant Ku; 6, cross-linked recombinant ATF 2. Arrows on the right show the migration positions of rainbow molecular mass markers (Amersham, Arlington Heights, Ill.).

band of about 90 kDa, suggesting that the probe was cross-linked only to the larger 86-kDa subunit of Ku. This pattern was identical to that observed for recombinant Ku (Fig. 9, lanes 3 to 5). Recombinant human ATF 2, which is a homodimer of a 54-kDa polypeptide, was included as a control (lane 6) and gave a band of the expected size for the monomer polypeptide. Thus, proteins cross-linked to the 25-bp MVM probe used in this experiment subsequently migrated with approximately their correct molecular mass upon subsequent SDS-PAGE.

PIF is an obligatory cofactor for NS1-mediated nicking of the viral replication origin. The reconstituted MVM 3′ origin replication assay requires an ordered sequence of events involving a variety of different biochemical steps, any of which could be regulated by an unknown but essential cellular factor. However, if this factor operates by binding to the consensus ATF site in the 3′ origin, it probably acts at the level of initiation and might well be able to cooperate with NS1 to introduce a single-strand nick at the correct site in the active origin. We have previously shown that purified NS1 can nick only a very small percentage of such templates even under conditions of minimal ionic strength and that, by itself, it shows no preference for the active (TC) versus the mutant (GAA) template (28). To explore this possibility further, we incubated purified NS1 with aliquots of the selected Mono-Q eluate in the presence of ATP and a ^{32}P-end-labeled DNA fragment containing either the active or the inactive origin sequence. After incubation, the products were analyzed by electrophoresis in nondenaturing polyacrylamide gels. In this assay, we would expect NS1 to nick a single strand of the template at the authentic initiation site (Fig. 10A) and become covalently attached to the 5′ end of a 53-nucleotide single-stranded product. The other fragment generated by this nick, a 42-nucleotide single strand, would not be detected since it is unlabeled. The second strand of the 95-bp substrate duplex should remain intact and might or might not remain base paired to the nicked strand, depending on whether the helicase activity of NS1 was engaged. Under the reaction conditions used here (100 mM salt and in the presence of nonspecific competitor DNA), purified NS1 alone failed to modify the authentic DNA template (Fig. 10B, lane 1), but when increasing amounts of the Mono-Q peak were added, a low-mobility DNA complex was formed with the active origin (lanes 2 and 3) but not with the inactive control sequence (lane 4). Since these samples were heated for 10 min in 0.5% SDS at 60°C before electrophoresis, the presence of such a DNA complex suggested that they might contain covalently attached protein. When equivalent samples were boiled in 0.5% SDS to melt all double-stranded DNA prior to electrophoresis, the resulting complexes showed increased mobility (lanes 6 and 7), indicating that the initial product was indeed double stranded. Moreover, if parallel samples were first treated with proteinase K and then boiled, the associated protein was removed and the predicted 53-nucleotide single-stranded DNA fragment was liberated.

To confirm that the retarded bands in lanes 2 and 3 were formed by the covalent attachment of NS1 to the probe, reaction mixtures were heated to 60°C for 30 min in 2% SDS and then immunoprecipitated with an antibody directed against the amino terminus of NS1 (Fig. 10C, lanes 1 to 4). These immunoprecipitates also allowed analysis of the reaction products in the absence of free probe. Thus, denaturation of the specific immunoprecipitate by boiling not only gave a DNA complex with a somewhat higher mobility than the original band but also released an intact, NS1-free, 95-nucleotide single strand, confirming that the original product was a duplex (lane 5). Proteinase K treatment coupled with heat denaturation gave the predicted radiolabeled 95- and 53-nucleotide ^{32}P-labeled products (lane 7). Taken together, these observations demonstrate that the factor we have identified and dubbed PIF functions as an essential cofactor for the NS1-mediated, site-specific nicking reaction which initiates MVM DNA replication at the 3′ origin.

DISCUSSION

In this study, we have shown that a cellular protein binds in a site-specific manner to the region of the MVM 3′ replication origin which contains the consensus ATF binding site, and that this protein acts an obligatory and specific cofactor in the NS1-mediated nicking event which initiates viral DNA replication. DNA binding analyses and in vitro replication assays using recombinant ATF proteins both suggest that this host protein is not a member of the ATF family of transcription factors. Rather, it appears to be a hitherto unknown human sequence-specific DNA binding protein, with an apparent molecular mass of approximately 110 kDa, which recognizes a sequence that overlaps with the consensus ATF binding motif in the MVM origin but also shows a marked affinity for poly(dI-dC). Since this protein is required for the initiation of DNA replication from the MVM 3′ origin we have designated it PIF, for parvovirus initiation factor.

By using an in vitro replication assay in which the substrate is a cloned form of the viral 3′ origin and by supplying the viral initiator protein NS1 in *trans*, we have begun to dissect 293 cell replication extracts to ask which cellular proteins are required

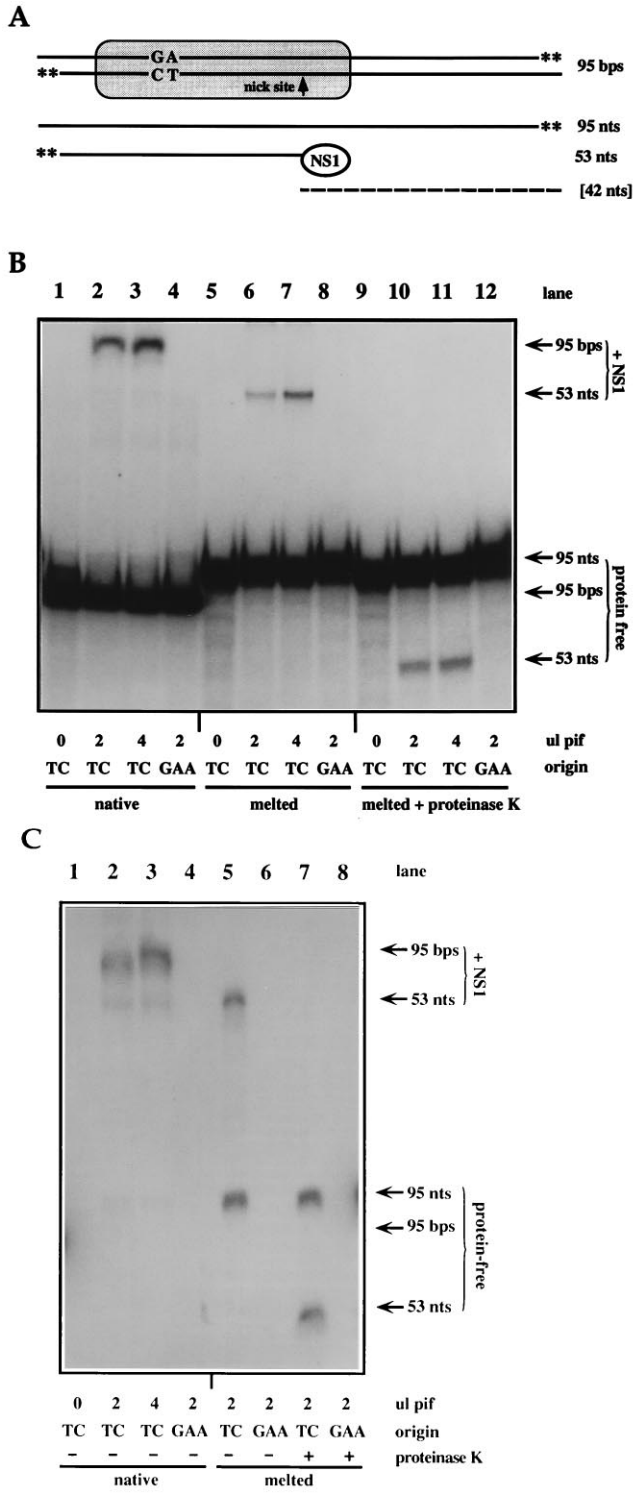

FIG. 10. (A) Diagram of substrate and products in the nicking reaction. Sequenase-labeled 3′ ends are marked **, the shaded area denotes the footprint of NS1 on the substrate, and the circled NS1 indicates the position of covalently linked NS1 at the end of the reaction. The dashed line indicates the predicted 42-nucleotide (nt) single strand released by denaturation of the product DNA, which is unlabeled. (B) NS1 requires PIF for nicking of the 3′ origin. The minimal active or inactive origin was isolated from pL1-2 TC or pL1-2 GAA after digestion with appropriate restriction enzymes, end labeled by filling in with [^{32}P]dATP, and incubated with NS1 in the presence or absence of PIF. The predicted DNA fragments produced by the NS1 nicking reaction are indicated to the right of the figure. The amounts of PIF (in microliters of Mono-Q pool) in the reaction mixture are listed below each lane, together with the nicking substrate. The reaction mixtures were all heated to 60°C in 0.5% SDS and, where indicated (below the lanes), melted by boiling or incubated with proteinase K and boiled prior to analysis by nondenaturing PAGE. (C) PIF is required for covalent attachment of NS1 to the 3′ origin. Reactions carried out as described for panel A were immunoprecipitated with a polyclonal antibody directed against the amino terminal of NS1 and analyzed by nondenaturing PAGE. The samples were all heated to 60°C in 0.5% SDS and, where indicated below the lanes, melted by boiling or incubated with proteinase K and boiled prior to electrophoresis.

for specific steps in parvovirus replication. Similar approaches have been used in several other viral systems to identify cellular factors which play essential roles in both the initiation and progression of viral replication forks, and such studies have been instrumental in identifying and allowing the biochemical characterization of many of the essential DNA replication proteins encoded by the host cell. However, in many cases, these studies have also focused attention on the unexpected roles played by known cellular DNA binding proteins, particularly transcription factors, in replication initiation (reviewed in reference 16). Thus, for example, early events in adenovirus replication require three host cell nuclear factors (NF), originally designated NFI, NFII, and NFIII, of which two have subsequently been shown to be known cellular transcription factors. NF-1 is a member of the CCAAT box binding protein family (6), and the requirement for NFIII can be satisfied by the POU domain of any of several octamer-binding transcription factor family members, such as OCT-1 (25, 29, 32). NF-II copurifies with topoisomerase I activity and is required only for leading-strand synthesis of sequences exceeding 50 bases (27).

The assay we have used involves NS1-mediated nicking and initiation at the MVM 3′ origin sequence, followed by the establishment and maintenance of a unidirectional and highly processive replication fork, so that this approach should theoretically allow the identification of cellular components working at several levels in the replication process. In this report, we explore the requirement for proteins from the P-cell 1 fraction of the replication extract and have demonstrated that three factors from this fraction, RPA, PCNA, and PIF, are absolutely required. As expected, P-cell 1 could be replaced by purified recombinant RPA and PCNA in our control in vitro SV40 replication assays (data not shown), indicating that our preparations of these proteins were fully active. Both of these components were also absolutely required for replication from the MVM 3′ origin in vitro in the presence of PIF. Similarly, both RPA and PCNA were required for replication of a template containing the MVM 5′ origin in vitro, even though PIF was not required, suggesting that the endonuclease function of NS1 must be activated by an alternate mechanism at this origin, perhaps, in this case, a factor(s) which fractionates in P-cell fraction 3. Although relatively little is known about the enzymology of parvovirus DNA replication, the requirement for RPA and PCNA was not surprising since these viruses employ a unidirectional, continuous, single-strand-specific mode of synthesis of the type generally associated with δ-polymerase and its accessory proteins (14). This suggestion is supported by the observation that neutralizing antibodies directed against the other host cell replicative polymerase, polymerase-α:primase, failed to inhibit replication from the MVM 3′ origin in vitro while severely impairing T-antigen-driven replication of an SV40 template in parallel reactions (data not shown).

Competition assays showed that the DNA binding activity of PIF was site specific for the left half of the MVM 3′ replication origin, which, as shown in Fig. 1C, contains a consensus ATF binding motif. Since we had previously shown that a mutation

in the core ACGT sequence in the ATF consensus motif severely reduced replication from this origin (13), it seemed possible that a member of the ATF family was involved, particularly since this family of transcription factors has been shown to interact with a number of other viral proteins, including adenovirus E1A and human T-cell leukemia virus (HTLV-1) *tax* (23, 43). However, while recombinant ATF 1, ATF 2, ATF 3, and CREB bound the origin avidly and site specifically, none were able to support replication in the reconstituted system, and they suppressed rather than stimulated replication in full S100 extracts. These data, taken together with results of competitive binding assays with wild-type and mutant ATF motifs from the human somatostatin promoter and UV cross-linking experiments which identified a protein with a molecular mass of around 110 kDa, make it very unlikely that PIF is an ATF family member. While PIF is quite likely to be a transcription factor, its binding site is unusual because, although it involves part of the ATF consensus, it can be competed effectively with poly(dI-dC). Possibly, such a factor escaped detection in most promoter regulation studies because of the routine use of poly(dI-dC) in DNA binding assays. Our recent mapping studies of the PIF site in the MVM origin by using oligonucleotide competition assays and a variety of footprinting techniques (9a) indicate that the site is bipartite, involving both the ACGT motif in the ATF consensus (Fig. 1C) and an identical tetranucleotide just inboard of this motif, separated from it by 5 bp, so that it immediately juxtaposes the bubble dinucleotide.

A similar activity was identified previously by Roberts et al. (33) in attempts to purify a transcription factor, which they named TRAC, involved in regulation of the human transferrin receptor promoter (19). The consensus DNA binding specificity of TRAC was determined to be A(A/T)GTGACG, which is a perfect match to the ATF-like sequence in the MVM 3′ origin. Binding of TRAC to its cognate sequence was also efficiently inhibited by poly(dI-dC). However, purification of the polypeptides thought to be responsible for the site-specific binding by combined DNA affinity and ion-exchange chromatography resulted in the apparent isolation of the two subunits of the Ku autoantigen. In retrospect, it seems quite unlikely that Ku was responsible for this binding. Ku is an abundant nonspecific DNA binding complex and is a frequent contaminant in purification schemes that use site-specific DNA affinity chromatography because of its affinity for duplex DNA ends. In the purification of TRAC and PIF, Ku causes a major problem because it copurifies with each of these site-specific factors through several different types of chromatography steps. While the cellular function and identity of PIF remain unknown, cross-competition in gel shift assays between PIF and the TRAC oligonucleotide suggest that PIF and TRAC are probably the same factor (9a). TRAC activity is known to be serum inducible, and the TRAC binding site in the promoter of the transferrin receptor appears to be involved in the mitogen-stimulated upregulation of this receptor, which is itself critical for cell cycle progression (19).

In this study, we have used reconstituted replication and direct nicking assays to demonstrate that PIF functions as an obligatory specific cofactor for NS1-mediated single-strand nicking. Although NS1 is known to have intrinsic helicase activity (28, 41), in our assays nicking did not lead directly to release of the duplex strand from the origin, perhaps suggesting that other components of the replication fork may be required to promote this activity. Although NS1 and PIF bind equally well to both the active TC origin sequence and its inactive GAA counterpart when assayed separately (reference 10 and data not shown), it will be of interest to see if they bind equally well to both sequences in the presence of the other protein. Since there is actually more space between the PIF and NS1 binding sites in the inactive sequence than in the active origin, it is unlikely that the inactivity of the GAA construct is due to direct competition for binding. However, the nature of the interactions between PIF and NS1, both on and off their DNA target, are at present unknown. It is also theoretically possible that the two do not interact directly at all but that PIF simply causes NS1 to pause at a critical point in its progress along the DNA, perhaps just as NS1 induces local melting of the two DNA strands at the nick site, allowing it time to recognize and cleave the nick site before progressing along the unwound DNA strands. According to this hypothesis, the insertion of a single extra nucleotide in the bubble sequence of the inactive origin might impair the alignment of the two proteins around the helix and thus allow the NS1 to move too far from the nick site before pausing.

The region containing the PIF binding site, including the ATF motif, is highly conserved in the 3′ hairpin of a number of other parvoviruses related to MVM, such as H1, LuIII, KRV and MPV, and is conserved to a lesser degree in several others including FPV, MEV, CPV, and PPV (1, 4). This suggests that PIF is likely to be an important factor in the asymmetric resolution of the 3′ replication origins in all these viruses and thus that PIF is likely to be expressed in a wide variety of cell types in many mammalian species. Since high-affinity NS1 binding sites and consensus NS1 nick sites are liberally distributed throughout these viral genomes, the requirement for an additional controlling interaction, such as an appropriately placed PIF complex, to activate the endonuclease function of NS1 allows this pleiotropic effector molecule to bind extensively throughout the duplex viral replication intermediates without nicking them. However, comparable PIF binding sites are not apparent in the 5′ telomeres of these viruses, suggesting that the virus must use an alternate mechanism or cellular cofactor at this location.

In this study, we have established that PIF acts as a cofactor for NS1 in replication initiation. However, since the 3′ origin overlaps several significant upstream elements in the initiating promoter of MVM (P4), PIF might also play an important role in transcriptional control during the viral life cycle. Possibly the interaction between PIF and NS1 has evolved as part of a mechanism to orchestrate optimal transcription and replication. The ATF motif and surrounding sequences are involved in modulating MVM P4 transcription in different cell lines, and transcriptional activity has been directly correlated to binding by members of the ATF family (17, 31). Binding activity that could be attributed to PIF was not reported in these studies, possibly due to the use of poly(dI-dC) as a presumed noncompetitive inhibitor, leaving open the possibility that PIF also plays a direct role in controlling transcription through this region of the origin and promoter. Purification of PIF and characterization of its function in viral and cellular replication and gene regulation are in progress.

ACKNOWLEDGMENTS

We are grateful for the generous contributions of reagents from J. Donald Capra, Francois Fuks, Tsonwin Hai, James Hoeffler, Tom Kelly, Edward Kuff, Nick Muzyczka, David Pintel, Natalie Salome, Bruce Stillman, and Chris Umbrecht, without which these studies would not have been possible.

This work was supported by Public Health Service grant AI26109 from the National Institute of Allergy and Infectious Diseases. J.C. was supported by grants from The Danish Agricultural and Veterinary Research Council.

REFERENCES

1. **Astell, C. R.** 1990. Terminal hairpins of parvovirus genomes and their role in DNA replication, p. 59–80. *In* P. Tijssen (ed.), Handbook of parvoviruses. CRC Press, Inc., Boca Raton, Fla.
2. **Astell, C. R., M. B. Chow, and D. C. Ward.** 1985. Sequence analysis of the termini of virion and replicative forms of minute virus of mice DNA suggests a modified rolling-hairpin model for autonomous parvovirus DNA replication. J. Virol. **54:**171–177.
3. **Baldauf, A., J. Rommelaere, and K. Willwand.** 1995. Distinct steps of MVM DNA replication reproduced in an in vitro system, abstr. Sb 4. Proceedings of the VIth Parvovirus Workshop, Montpellier, France.
4. **Berns, K. I., M. Bergoin, M. Bloom, M. Lederman, N. Muzyczka, G. Siegl, J. Tal, and P. Tattersall.** 1995. Parvoviridae. Arch. Virol. Suppl. **10:**169–178.
5. **Chen, C. Y., D. H. Bessesen, S. M. Jackson, and J. P. Hoeffler.** 1991. Overexpression and purification of transcriptionally competent CREB from a recombinant baculovirus. Protein Expression Purif. **2:**402–411.
6. **Chen, M., N. Mermod, and M. S. Horwitz.** 1990. Protein-protein interactions between adenovirus DNA polymerase and nuclear factor I mediate formation of the DNA replication preinitiation complex. J. Biol. Chem. **265:**18634–18642.
7. **Christensen, J., S. F. Cotmore, and P. Tattersall.** 1995. The minute virus of mice transcriptional activator protein, NS1, binds directly to the *trans*-activating region (*tar*) of the viral P38 promoter in a strictly ATP-dependent manner. J. Virol. **69:**5422–5430.
8. **Christensen, J., M. Pedersen, B. Aasted, and S. Alexandersen.** 1995. Purification and characterization of the major nonstructural protein (NS-1) of Aleutian mink disease virus. J. Virol. **69:**1802–1809.
9. **Christensen, J., T. Storgaard, B. Bloch, S. Alexandersen, and B. Aasted.** 1993. Expression of Aleutian mink disease parvovirus proteins in a baculovirus vector system. J. Virol. **67:**229–238.

9a. **Christensen, J., et al.** Unpublished data.

10. **Cotmore, S. F., J. Christensen, J. P. F. Nüesch, and P. Tattersall.** 1995. The NS1 polypeptide of the murine parvovirus minute virus of mice binds to DNA sequences containing the motif $[ACCA]_{2-3}$. J. Virol. **69:**1652–1660.
11. **Cotmore, S. F., J. P. Nüesch, and P. Tattersall.** 1993. Asymmetric resolution of a parvovirus palindrome in vitro. J. Virol. **67:**1579–1589.
12. **Cotmore, S. F., and P. Tattersall.** 1988. The NS-1 polypeptide of minute virus of mice is covalently attached to the 5′ termini of duplex replicative-form DNA and progeny single strands. J. Virol. **62:**851–860.
13. **Cotmore, S. F., and P. Tattersall.** 1994. An asymmetric nucleotide in the parvoviral 3′ hairpin directs segregation of a single active origin of DNA replication. EMBO J. **13:**4145–4152.
14. **Cotmore, S. F., and P. Tattersall.** 1995. DNA replication in the autonomous parvoviruses. Semin. Virol. **6:**271–281.
15. **Cotmore, S. F., and P. Tattersall.** 1996. Parvovirus DNA replication, p. 799–813. *In* M. DePamphilis (ed.), DNA replication in eukaryotic cells. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

15a. **Cotmore, S. F., and P. Tattersall.** Unpublished observations.

16. **DePamphilis, M. (ed.).** 1996. DNA replication in eukaryotic cells. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. **Faisst, S., M. Perros, L. Deleu, N. Spruyt, and J. Rommelaere.** 1994. Mapping of upstream regulatory elements in the P4 promoter of parvovirus minute virus of mice. Virology **202:**466–470.
18. **Fien, K., and B. Stillman.** 1992. Identification of replication factor C from *Saccharomyces cerevisiae*: a component of the leading-strand DNA replication complex. Mol. Cell. Biol. **12:**155–163.
19. **Hirsch, S., and W. K. Miskimins.** 1995. Mitogen induction of nuclear factors that interact with a delayed responsive region of the transferrin receptor gene promoter. Cell Growth Differ. **6:**719–726.
20. **Holscher, C., J. A. Kleinschmidt, and A. Burkle.** 1995. High-level expression of adeno-associated virus (AAV) Rep78 or Rep68 protein is sufficient for infectious-particle formation by a rep-negative AAV mutant. J. Virol. **69:**6880–6885.
21. **Karin, M., and T. Smeal.** 1992. Control of transcription factors by signal transduction pathways: the beginning of the end. Trends Biochem. Sci. **17:**418–422.
22. **Kornberg, A., and T. A. Baker.** 1991. DNA replication, 2nd ed. W. H. Freeman & Co., New York, N.Y.
23. **Liu, F., and M. R. Green.** 1990. A specific member of the ATF transcription factor family can mediate transcription activation by the adenovirus E1a protein. Cell **61:**1217–1224.
24. **Liu, Q., C. B. Yong, and C. R. Astell.** 1994. In vitro resolution of the dimer bridge of the minute virus of mice (MVM) genome supports the modified rolling hairpin model for MVM replication. Virology **201:**251–262.
25. **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.
26. **Naeger, L. K., J. Cater, and D. J. Pintel.** 1990. The small nonstructural protein (NS2) of the parvovirus minute virus of mice is required for efficient DNA replication and infectious virus production in a cell-type-specific manner. J. Virol. **64:**6166–6175.
27. **Nagata, K., R. A. Guggenheimer, and J. Hurwitz.** 1983. Adenovirus DNA replication in vitro: synthesis of full-length DNA with purified proteins. Proc. Natl. Acad. Sci. USA **80:**4266–4270.
28. **Nüesch, J. P. F., S. F. Cotmore, and P. Tattersall.** 1995. Sequence motifs in the replicator protein of parvovirus MVM essential for nicking and covalent attachment to the viral origin: identification of the linking tyrosine. Virology **209:**122–135.
29. **O'Neill, E. A., C. Fletcher, C. R. Burrow, N. Heintz, R. G. Roeder, and T. J. Kelly.** 1988. Transcription factor OTF-1 is functionally identical to the DNA replication factor NF-III. Science **241:**1210–1213.
30. **Ono, M., P. W. Tucker, and J. D. Capra.** 1994. Production and characterization of recombinant human Ku antigen. Nucleic Acids Res. **22:**3918–3924.
31. **Perros, M., L. Deleu, J. M. Vanacker, Z. Kherrouche, N. Spruyt, S. Faisst, and J. Rommelaere.** 1995. Upstream CREs participate in the basal activity of minute virus of mice promoter P4 and in its stimulation in *ras*-transformed cells. J. Virol. **69:**5506–5515.
32. **Pruijn, G. J., R. T. van Miltenburg, J. A. Claessens, and P. C. Van der Vliet.** 1988. Interaction between the octamer-binding protein nuclear factor III and the adenovirus origin of DNA replication. J. Virol. **62:**3092–3102.
33. **Roberts, M. R., Y. Han, A. Fienberg, L. Hunihan, and F. H. Ruddle.** 1994. A DNA-binding activity, TRAC, specific for the TRA element of the transferrin receptor gene copurifies with the Ku autoantigen. Proc. Natl. Acad. Sci. USA **91:**6354–6358.
34. **Traut, W., and E. Fanning.** 1988. Sequence-specific interactions between a cellular DNA-binding protein and the simian virus 40 origin of DNA replication. Mol. Cell. Biol. **8:**903–911.
35. **Tsurimoto, T., T. Melendy, and B. Stillman.** 1990. Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin. Nature **346:**534–539.
36. **Tsurimoto, T., and B. Stillman.** 1989. Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication in vitro. Mol. Cell. Biol. **9:**609–619.
37. **Tsurimoto, T., and B. Stillman.** 1991. Replication factors required for SV40 DNA replication in vitro. I. DNA structure-specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory proteins. J. Biol. Chem. **266:**1950–1960.
38. **Tuteja, N., R. Tuteja, A. Ochem, P. Taneja, N. W. Huang, A. Simoncsits, S. Susic, K. Rahman, L. Marusic, J. Chen, J. Zhang, S. Wang, S. Pongor, and A. Falaschi.** 1994. Human DNA helicase II: a novel DNA unwinding enzyme identified as the Ku autoantigen. EMBO J. **13:**4991–5001.
39. **Umbricht, C. B., L. F. Erdile, E. W. Jabs, and T. J. Kelly.** 1993. Cloning, overexpression, and genomic mapping of the 14-kDa subunit of human replication protein A. J. Biol. Chem. **268:**6131–6138.
40. **Umbricht, C. B., C. A. Griffin, A. L. Hawkins, K. H. Grzeschik, P. O'Connell, R. Leach, E. D. Green, and T. J. Kelly.** 1994. High-resolution genomic mapping of the three human replication protein A genes (RPA1, RPA2, and RPA3). Genomics **20:**249–257.
41. **Wilson, G. M., H. K. Jindal, D. E. Yeung, W. Chen, and C. R. Astell.** 1991. Expression of minute virus of mice major nonstructural protein insect cells: purification and identification of ATPase and helicase activities. Virology **185:**90–98.
42. **Wold, M. S., D. H. Weinberg, D. M. Virshup, J. J. Li, and T. J. Kelly.** 1989. Identification of cellular proteins required for simian virus 40 DNA replication. J. Biol. Chem. **264:**2801–2819.
43. **Yoshida, M., T. Suzuki, J. Fujisawa, and H. Hirai.** 1995. HTLV-1 oncoprotein tax and cellular transcription factors. Curr. Top. Microbiol. Immunol. **193:**79–89.