MOLECULAR AND CELLULAR BIOLOGY, Feb. 1987, p. 875–886
0270-7306/87/020875-12$02.00/0

Vol. 7, No. 2

Sequence-Specific Interactions between Cellular DNA-Binding Proteins and the Adenovirus Origin of DNA Replication

PHILIP J. ROSENFELD, EDWARD A. O'NEILL, RONALD J. WIDES, AND THOMAS J. KELLY*

Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received 8 August 1986/Accepted 10 November 1986

The adenovirus origin of DNA replication contains three functionally distinct sequence domains (A, B, and C) that are essential for initiation of DNA synthesis. Previous studies have shown that domain B contains the recognition site for nuclear factor I (NF-I), a cellular protein that is required for optimal initiation. In the studies reported here, we used highly purified NF-I, prepared by DNA recognition site affinity chromatography (P. J. Rosenfeld and T. J. Kelly, Jr., J. Biol. Chem. 261:1398–1408, 1986), to investigate the cellular protein requirements for initiation of viral DNA replication. Our data demonstrate that while NF-I is essential for efficient initiation in vitro, other cellular factors are required as well. A fraction derived from HeLa cell nuclear extract (BR-FT fraction) was shown to contain all the additional cellular proteins required for the complete reconstitution of the initiation reaction. Analysis of this complementing fraction by a gel electrophoresis DNA-binding assay revealed the presence of two site-specific DNA-binding proteins, ORP-A and ORP-C, that recognized sequences in domains A and C, respectively, of the viral origin. Both proteins were purified by DNA recognition site affinity chromatography, and the boundaries of their binding sites were defined by DNase I footprint analysis. Additional chracterization of the recognition sequences of ORP-A, NF-I, and ORP-C was accomplished by determining the affinity of the proteins for viral origins containing deletion and base substitution mutations. ORP-C recognized a sequence between nucleotides 41 and 51 of the adenovirus genome, and analysis of mutant origins indicated that efficient initiation of replication is dependent on the presence of a high-affinity ORP-C-binding site. The ORP-A recognition site was localized to the first 12 base pairs of the viral genome within the minimal origin of replication. These data provide evidence that the initiation of adenovirus DNA replication involves multiple protein-DNA interactions at the origin.

Initiation of adenovirus DNA replication takes place at either terminus of the linear viral genome (for a review, see reference 23). The initiation reaction involves the formation of a covalent linkage between dCMP, the first nucleotide of the nascent DNA chain, and a virus-encoded primer protein, referred to as the 80-kilodalton (kDa) preterminal protein (pTP) (7, 9, 28, 36, 45). The protein-nucleotide bond formed during this reaction has been identified as an ester linkage between the α-phosphoryl group of dCMP and the β-OH of a serine residue in the pTP (11). The synthesis of a new viral DNA strand takes place by extension from the free 3′ hydroxyl group present in the pTP-dCMP initiation complex.

Analysis of the replication of mutant viral genomes in vitro has demonstrated that the initiation of viral DNA replication is dependent on the presence of specific nucleotide sequence domains within the terminal region of the viral DNA (10, 12, 21, 25, 26, 38, 45–47, 49). These required sequence domains constitute the adenovirus origin of DNA replication. In the accompanying paper, we have presented evidence that the origin of DNA replication is wholly contained within the first 51 base pairs of the viral genome (49). Moreover, our data indicate that the origin is made up of at least three functionally distinct domains. Domain A consists of the first 18 base pairs of the viral genome and represents the minimal origin of DNA replication. The presence of domain A is absolutely required for the initiation reaction, but in the absence of the other domains the efficiency of the reaction is only about 3% of the optimal level. The addition of the DNA segments between nucleotides 19 and 39 (domain B) and between nucleotides 40 and 51 (domain C) increases the efficiency of initiation 10-fold and 3-fold, respectively.

Both virus-encoded and cell-encoded proteins are required for initiation of adenovirus DNA replication in vitro. The viral initiation proteins include the 80-kDa pTP, described above, and the 140-kDa adenovirus DNA polymerase (14, 27, 34, 43, 44). However, these two proteins are incapable of supporting a significant level of initiation unless they are supplemented with a nuclear extract from human (HeLa) cells. One stimulatory cellular activity present in such nuclear extracts, nuclear factor I (NF-I), was identified in previous studies (31, 38). Analysis of partially purified preparations of NF-I demonstrated that the protein binds with high affinity to domain B of the adenovirus origin of DNA replication (12, 26, 32, 38, 40). In addition, studies with mutant viral templates showed that binding of NF-I to the replication origin is essential for its stimulatory activity (12, 21, 26, 38). However, since these functional studies were performed with relatively crude preparations of NF-I it was not possible to rule out the possibility that additional cellular proteins were also required for efficient initiation of adenovirus DNA replication.

We have recently developed a method for the purification of NF-I that is based on the high-affinity interaction between the protein and its recognition site (39). This method (DNA recognition site affinity chromatography) has made it possible to obtain highly purified NF-I with good yield. In this report we describe studies in which we made use of affinity-purified NF-I to reinvestigate the requirement for cellular proteins in the initiation of adenovirus DNA replication. Our data demonstrate that while NF-I is essential for efficient initiation in vitro, other cellular factors are required as well.

* Corresponding author.

We identified a fraction derived from uninfected HeLa cell extracts (BR-FT fraction) that contains these additional factors. The complete reconstitution of the initiation reaction was obtained with purified HeLa NF-I and the BR-FT fraction, together with the purified viral initiation proteins pTP and adenovirus DNA polymerase (Ad pol). Analysis of the complementing BR-FT fraction revealed the presence of two site-specific DNA-binding proteins, ORP-A and ORP-C, that interact specifically with domains A and C, respectively, of the adenovirus origin of DNA replication. Both of these origin-specific binding proteins were purified by DNA recognition site affinity chromatography, and the boundaries of their recognition sites were defined by DNase I footprint analysis. These findings, together with those presented in the accompanying paper (49), suggest that the initiation of adenovirus DNA replication requires the specific interaction of multiple cellular proteins with the viral origin of DNA replication.

MATERIALS AND METHODS

Plasmid DNAs. Plasmid DNAs containing both wild-type and mutant adenovirus DNA sequences were constructed as described by Wides et al. (49) with the exception of plasmids pU*dl*19-67 and pU*dl*-38-67. Plasmid pU*dl*19-67 was constructed from plasmid pU*dl*67I10 (49). pU*dl*67I10 was digested with *Bgl*II and *Bam*HI, and the fragment containing the adenovirus origin sequences from positions 19 to 67 was subcloned into the *Bam*HI site of pUC9. Plasmid pU*dl*38-67 was constructed from the point mutant pUpm37 (49). The mutation at position 37 of pUpm37 created an *Nde*I restriction site within the adenovirus terminus. pUpm37 was cleaved with *Nde*I, and the large fragment containing the adenovirus DNA sequence from positions 38 to 67 was circularized to generate the plasmid pM*dl*38-67. Cleavage of the plasmid pM*dl*38-67 with *Hgi*AI and *Pst*I released a fragment that contained the adenovirus DNA sequence which was subcloned into the *Pst*I site of pUC9 to construct the plasmid pU*dl*38-67.

pTP-dCMP complex formation in vitro. The standard in vitro initiation reaction was performed as previously described with only minor modifications (39). Before incubation at 37°C, the initiation assays were adjusted to a final concentration of 30 mM NaCl and 10% glycerol. The amount of pTP-dCMP complex formed in these assays was determined after the reaction products were electrophoresed and visualized by autoradiography. The gel segment containing the initiation complex was excised, and the radioactivity contained in the gel slice was quantitated by liquid scintillation counting in Betafluor (National Diagnostics). An adjacent gel segment of identical size served as a control for background radioactivity in each lane. The molar amount of initiation complex was calculated from a standard curve that was prepared with gel slices that contained known molar amounts of radioactive nucleotide.

Gel electrophoresis DNA-binding assay. The standard gel electrophoresis DNA-binding assay was a modification of previously published protocols (18, 19, 42). Plasmid DNAs were linearized with either *Bam*HI or *Hin*dIII and labeled at their 3′ termini (2.0×10^4 cpm/fmol) by incubation with [α-^{32}P]dATP and [α-^{32}P]dTTP (3,000 Ci/mmol) in the presence of *Micrococcus luteus* polymerase (Midland Certified Reagent Co., Midland, Tex.) (29). The radioactive DNAs were cleaved with either *Pvu*II or *Eco*RI, and the appropriate fragments were isolated by electroelution after electrophoresis through an 8% polyacrylamide gel. The DNA fragments used in this assay contained similar pUC9 sequences but differed in the sequences derived from the adenovirus origin of DNA replication. The standard binding assay (25 μl) contained 25 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 7.5), 50 mM NaCl, 10% glycerol, 0.05% Nonidet P-40, 5 μg of bovine serum albumin, 1 mM dithiothreitol (DTT), 1 mM EDTA, 1 μg of poly(dI-dC)-poly(dI-dC) (Pharmacia P-L Biochemicals), and 0.1 nM [^{32}P]DNA fragment. The assays were incubated at 25°C for 30 min and electrophoresed at 10 V/cm through a 6% polyacrylamide gel (30% acrylamide–0.8% bisacrylamide) that had been preelectrophoresed for 2 h at 10 V/cm. The gel buffer, 12 mM Tris acetate (pH 7.5)–1 mM EDTA–0.01% Nonidet P-40, was continually recirculated to maintain a constant pH during the electrophoretic separation. After 90 min, the gel was dried on DE-81 chromatography paper (Whatman, Inc., Clifton, N.J.), and radioactivity was detected by exposure to Kodak XAR-5 film with a Du Pont Cronex Lightning-Plus intensifying screen at −70°C.

DNase I footprint analysis. The plasmid pU*dl*67 was linearized with either *Bam*HI or *Pvu*II, and the 5′ termini were radioactively labeled (1.3×10^4 cpm/fmol) by incubation with [γ-^{32}P]ATP (5,000 Ci/mmol) in the presence of T4 polynucleotide kinase after treatment with calf intestinal alkaline phosphatase (29). The radioactive DNAs were subsequently cleaved with *Pvu*II or *Bam*HI, respectively, and the 157-base-pair fragment containing the adenovirus origin of DNA replication was isolated by electroelution from an 8% polyacrylamide gel. Various amounts of ORP-A, NF-I, and ORP-C were incubated with the radioactive DNA fragments (5 fmol) at 25°C for 30 min in a reaction mixture (50 μl) containing 25 mM HEPES (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, and 1 μg of poly(dI-dC)-poly(dI-dC). DNase I (0.4 U) was added, and after 30 s the digestion was terminated by adjusting the reaction mixture to a final concentration of 0.3 M sodium acetate (pH 5.0)–10 mM EDTA–0.5% sodium dodecyl sulfate (SDS) to 30 μg of sheared calf thymus DNA per ml in a final volume of 75 μl. After extraction with phenol-chloroform, the DNA was ethanol precipitated and suspended in sample buffer containing 85% formamide, 15 mM NaOH, 1 mM EDTA, 0.01% bromphenol blue, and 0.01% xylene cyanol. Samples were incubated at 100°C for 5 min and electrophoresed through an 8% and a 12% polyacrylamide–8 M urea gel as described by Maxam and Gilbert (30). G+A and C+A sequencing reactions were performed by a modification of the Maxam and Gilbert protocol (3). Radioactivity was detected by exposure of the wet gel to Kodak XAR-5 film with a Du Pont Cronex Lightning-Plus intensifying screen at −70°C.

Competition nitrocellulose filter binding assay. The nitrocellulose filter binding assay was performed as previously described (39) with only minor modifications. The modified reaction mixture contained 0.625 ng of NF-I purified by DNA recognition site affinity chromatography (39), 40 pM radioactive DNA, and no sheared *Escherichia coli* DNA. The concentration of competitor DNA was determined by optical A_{260} and by the diphenylamine method (6).

Purification of virus-encoded replication proteins. The Ad pol and preterminal protein (pTP) were detected during the purification by using the in vitro initiation assay previously described (39). Fractions containing the pTP-Ad pol activities were identified by their ability to support formation of pTP-dCMP initiation complexes when incubated in the presence of purified NF-I (10 ng) and the Bio-Rex flowthrough fraction (BR-FT fraction) from the NF-I purification (7.5

μg). The aphidicolin-resistant DNA polymerase activity associated with the pTP-Ad pol complex was quantitated with activated calf thymus DNA as the template in the presence and absence of 100 μM aphidicolin (2). A unit of polymerase activity (pol unit) is defined as the amount of polymerase required to incorporate 1 nmol of dTMP into acid-insoluble material in 20 min at 30°C. The 72-kDa single-stranded (ss) DBP was detected by electrophoresis of protein fractions through an 8% SDS-polyacrylamide gel (24) followed by staining with Coomassie blue.

The adenovirus-encoded replication proteins were purified by a modification of the procedure described by Enomoto et al. (14). Cytoplasmic extract was prepared from adenovirus-infected HeLa cells (1.5×10^{10}; 30 g) as previously described (8), followed by centrifugation at 105,000 × *g* for 1 h. The supernatant was adjusted to 150 mM NaCl–1 mM DTT–1 mM EDTA (Ad-pro fraction 1; 201 ml; 2.2 g; 169 pol units) and loaded onto a phosphocellulose column (2.5 by 13.5 cm; Whatman P-11) that was preequilibrated with buffer B containing 150 mM NaCl. The column was washed with buffer B containing 150 mM NaCl (100 ml), and proteins were eluted with a linear gradient from 150 mM to 1 M NaCl in buffer B. The peak of pTP-Ad pol activity eluted at 360 mM NaCl. Fractions were pooled (Ad-pro fraction 2; 80 ml; 128 mg; 114 pol units), adjusted to 160 mM NaCl by the addition of buffer C, and loaded onto an ss calf thymus DNA-cellulose column (2.5 by 6.5 cm; Sigma Chemical Co., St. Louis, Mo.) that was preequilibrated with buffer C containing 150 mM NaCl. The matrix was washed with 150 mM NaCl in buffer C (48 ml), and the pTP-Ad pol was eluted with a linear gradient from 150 to 560 mM NaCl in buffer C (125 ml). The peak of pTP-Ad pol activity eluted at 300 mM NaCl (Ad pro fraction 3; 21 ml; 8.1 mg; 29 pol units). The 72-kDa ss DBP was recovered from the ss DNA-cellulose matrix by a step elution of 2 M NaCl in buffer C (10 ml; 1.2 mg). To concentrate the pTP-Ad pol activity, we adjusted Ad-pro fraction 3 to 150 mM NaCl with buffer B and loaded the fraction onto a phosphocellulose column (1.0 by 1.5 cm) preequilibrated with buffer B containing 150 mM NaCl. Concentrated pTP-Ad pol was obtained by a step elution with 500 mM NaCl in buffer B (Ad-pro fraction 4; 0.4 ml; 2.3 mg; 15 pol units). Fraction 4 was used in all experiments containing pTP-Ad pol.

The composition of buffer B was 10 mM sodium phosphate (pH 6.0), 10% sucrose, 10% glycerol, 0.01% Nonidet P-40, 1 mM DTT, 1 mM EDTA, and 0.1 mM phenylmethylsulfonyl fluoride. Buffer C was 25 mM sodium phosphate (pH 6.0), 10% glycerol, 0.01% Nonidet P-40, 1 mM DTT, 1 mM EDTA, and 0.1 mM phenylmethylsulfonyl fluoride.

Purification of ORP-A and ORP-C. ORP-A and ORP-C were detected during the purification by a modification of the gel electrophoresis DNA-binding assay described previously. Assays (25 μl) contained 10 μl of a column fraction diluted (1:10) in buffer S, and column fractions were assayed with the radioactive fragments derived from pUC9, pU*dl*67, pU*dl*18, or pU*dl*19-67 as described above. The binding assays were electrophoresed through a 2% agarose gel containing 12 mM Tris acetate (pH 7.5), 1 mM EDTA, and 0.05% Nonidet P-40. The gel was electrophoresed at 10 V/cm for 90 min with continuous recirculation. After the gel was dried and the DNA was visualized by autoradiography, fractions containing ORP-A and ORP-C were identified by comparing the elution profiles of the binding activities specific for each of the DNA fragments. The specific binding activity (SBA) for each fraction was defined as the amount of specific [^{32}P]DNA that was bound per milligram of protein. The SBAs were calculated from the linear portion of the binding curves that were generated at each stage of the purification. The binding curves were obtained by incubating various concentrations of a given protein fraction with a fixed DNA concentration followed by separation of the bound and unbound DNA by gel electrophoresis. The band that corresponded to the unbound DNA was excised from the dried gel, and the amount of DNA was quantitated by liquid scintillation counting (Betafluor). The amount of DNA bound was calculated from the difference between the amount of input DNA and the amount of DNA that remained unbound at each protein concentration.

Nuclear extract was prepared from S-3 HeLa suspension culture (2.5×10^{10} cells; 60 g) as previously described (39). After centrifugation at 10,000 × *g* and dilution with buffer D, the nuclear extract (450 ml; 3,825 mg) was loaded onto a Bio-Rex 70 column (200/400 mesh; 200 ml; 5 by 10 cm; Bio-Rad Laboratories, Richmond, Calif.) that was preequilibrated with 200 mM NaCl in buffer S. The flowthrough from the Bio-Rex 70 column (BR-FT; 450 ml; 675 mg) contained both ORP-A- and ORP-C-binding activities (SBA: ORP-A, 8.4×10^3; ORP-C, 42×10^3). The BR-FT was diluted with an equal volume of buffer S and loaded onto a calf thymus DNA-cellulose column (60 ml; 2.5 by 12 cm) that was preequilibrated with 100 mM NaCl in buffer S. The DNA-cellulose matrix was prepared by the method of Alberts and Herrick (1). The column was eluted with a linear gradient from 100 to 500 mM NaCl in buffer S (350 ml), and the peak of ORP-A activity eluted at 250 mM NaCl (55 ml; 12.1 mg; SBA, 116×10^3). The ORP-C activity was subsequently recovered by a 2 M NaCl step elution (25 ml; 3.5 mg; SBA, $1,616 \times 10^3$). Further purification of ORP-A was achieved by DNA recognition site affinity chromatography on a pKB67-88 DNA-cellulose column prepared by the method of Rosenfeld and Kelly (39). The pool of ORP-A activity was adjusted to 100 mM NaCl by the addition of buffer S and loaded onto the specific DNA-cellulose column (8.5 ml; 1.5 by 4.8 cm) that was preequilibrated with 100 mM NaCl in buffer S. The column was developed with a linear gradient between 100 and 500 mM NaCl (60 ml) and the peak of ORP-A activity eluted at 300 mM NaCl (5 ml; 0.74 mg; SBA, $2,026 \times 10^3$). The ORP-A pool was adjusted to 100 mM NaCl with buffer S and chromatographed a second time on the pKB67-88 DNA-cellulose column with an identical gradient elution (5.6 ml; 0.45 mg; SBA, $2,950 \times 10^3$). Both ORP-A and ORP-C were dialyzed against 50 mM NaCl in buffer S and stored at −70°C.

The buffers used in the purification of ORP-A and ORP-C contained the protease inhibitors pepstatin, chymostatin, and antipain (Sigma) at a concentration of 1 μg/ml, and phenylmethylsulfonyl fluoride (Aldrich Chemical Co., Inc., Milwaukee, Wis.) at a concentration of 0.1 mM. The composition of buffer D was 25 mM HEPES (pH 7.5), 40% glycerol, 0.01% Nonidet P-40, 1 mM DTT, and 2 mM EDTA. Buffer S was 25 mM HEPES (pH 7.5), 20% glycerol, 0.01% Nonidet P-40, 1 mM DTT, and 1 mM EDTA.

Protein assay. Protein concentrations were determined by the method of Bradford (5). Bovine γ-globulin was used as the standard.

RESULTS

The initiation of adenovirus DNA replication involves the formation of a covalent linkage between the preterminal protein (pTP) and a dCMP residue. The initiation reaction can be monitored in vitro by measuring the incorporation of

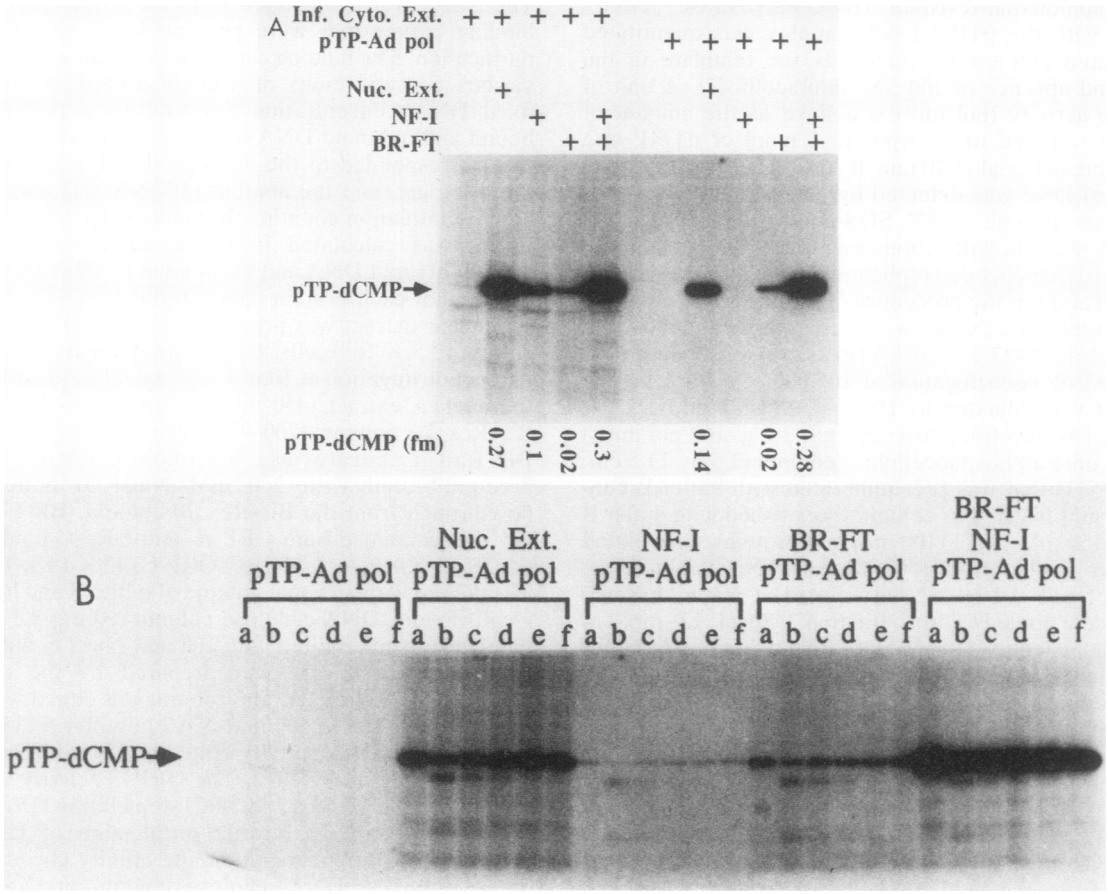


FIG. 1. Cellular and viral protein requirements for the initiation of adenovirus DNA replication. (A) Formation of the pTP-dCMP initiation complex was assayed as described in the Materials and Methods. The pTP-dCMP complexes formed in the presence of [α-^{32}P]dCTP were analyzed by SDS-polyacrylamide gel electrophoresis followed by autoradiography. The viral proteins required for initiation (pTP-Ad pol) were added either as a crude cytoplasmic extract from adenovirus type 5-infected HeLa cells (Inf. Cyto. Ext.; 11 μg; 0.001 pol units) or as a purified protein complex (pTP-Ad pol; 0.6 μg; 0.004 pol units). The following amounts of cellular proteins were added as indicated: HeLa nuclear extract (Nuc. Ext.) 30 μg; purified NF-I, 15 ng; Bio-Rex 70 flowthrough fraction (BR-FT), 7.5 μg. (B) The extent of pTP-dCMP complex formation was assayed in either the absence (lanes a) or presence (lanes b to f) of the virus-encoded 72-kDa ss DBP. Lanes b, 2.4 μg of ss DBP; lanes c, 1.2 μg of ss DBP; lanes d, 0.6 μg of ss DBP; lanes e, 0.3 μg of ss DBP; lanes f, 0.15 μg of ss DBP.

[α-^{32}P]dCTP into SDS-resistant pTP-dCMP complexes. Using this assay, several laboratories have shown that initiation requires a complex of two virally encoded proteins, the preterminal protein and Ad pol (pTP-Ad pol) (14, 27, 34, 36). However, the efficiency of the reaction is very low unless the viral replication proteins are complemented with a nuclear extract prepared from permissive (e.g., HeLa) cells. One stimulatory activity present in HeLa nuclear extracts, NF-I, has been characterized previously (12, 13, 26, 31, 32, 38, 39). Partially purified preparations of NF-I were shown to contain a origin-specific binding activity required for the initiation of replication, but the possibility remained that these preparations contained more than one essential initiation protein. To investigate this possibility, we purified NF-I to near homogeneity (39) and used the purified material to reinvestigate the protein requirements for the initiation of adenovirus DNA replication in vitro.

Protein requirements for the initiation of adenovirus DNA replication. As an approach to identifying the cellular proteins required for the initiation of adenovirus DNA replication, we carried out a series of reconstitution experiments with purified NF-I and pTP-Ad pol complex (Fig. 1A). The NF-I preparation used in these studies was purified from HeLa nuclear extract by DNA recognition site affinity chromatography (39), using an assay based solely on specific DNA-binding activity. The viral replication proteins alone (pTP-Ad pol) did not support the formation of detectable amounts of pTP-dCMP initiation complexes in vitro (Fig. 1A). As expected, the addition of crude nuclear extract from HeLa cells resulted in efficient reconstitution of the initiation reaction. However, when the viral proteins were incubated with affinity-purified NF-I alone, no detectable initiation complex was observed (Fig. 1A). Since comparable amounts of NF-I-specific binding activity were added to both assays, these results suggested that additional cellular proteins are required for initiation of adenovirus DNA replication.

To identify the additional cellular initiation proteins, we assayed side fractions from the purification of NF-I for their ability to complement NF-I in reconstitution assays. We found that the flowthrough fraction from the Bio-Rex 70 column (BR-FT), the first column used in the NF-I purification (39), completely reconstituted the initiation reaction when added to the viral proteins together with NF-I (Fig. 1A). The BR-FT fraction was free of detectable NF-I-specific DNA-binding activity (data not shown), and in the presence of a fixed amount of NF-I, the extent of pTP-dCMP

complex formation was directly proportional to the BR-FT concentration in the reaction mixture. Although the addition of the BR-FT fraction alone slightly stimulated initiation in vitro, complete reconstitution of the reaction required both BR-FT and NF-I. Identical results were obtained when the same protein fractions were assayed for their ability to support the complete adenovirus DNA replication reaction (unpublished data). We conclude from these data that the BR-FT fraction contains a novel cellular activity required for the initiation of viral DNA replication.

We have previously reported that affinity-purified NF-I is capable of stimulating initiation when added to cytoplasmic extract from adenovirus-infected cells in the absence of any additional source of nuclear proteins (39). Although this result was confirmed in the present study, it is clear from the data shown in Fig. 1A that the stimulation of initiation by NF-I is much less than that observed with crude nuclear extract. It seems likely, therefore, that the stimulatory effect of NF-I in this context is due to the presence of low levels of BR-FT activity in cytoplasmic extracts from infected cells. In any case, the addition of BR-FT to cytoplasmic extract supplemented with NF-I markedly enhanced the efficiency of initiation, confirming the conclusion that both NF-I and an activity(s) present in BR-FT are required for optimal initiation.

It has been reported that the adenovirus 72-kDa ss DBP enhances the efficiency of initiation in reactions containing the viral proteins (pTP-Ad pol) and NF-I (31). In our experiments the adenovirus ss DBP had no effect on the extent of initiation either in the presence or in the absence of NF-I (Fig. 1B). On the other hand, the 72-kDa ss DBP was absolutely required for efficient chain elongation in assays that measured the complete replication reaction (unpublished data). It is possible that the stimulatory effect on initiation previously attributed to the ss DBP was a result of contamination of ss DBP preparations with the activity(s) associated with the BR-FT fraction.

Recognition of adenovirus origin of replication by cellular DNA-binding proteins. Genetic studies described in the accompanying paper (49) indicate that the adenovirus origin of replication is entirely contained within the first 51 base pairs of the viral genome and consists of three functionally distinct domains designated A, B, and C. Domain A (nucleotides 1 to 18) contains the minimal sequence sufficient for origin function. Domains B (nucleotides 19 to 40) and C (nucleotides 41 to 51) contain accessory sequences that significantly increase the activity of the minimal origin. Previous studies have demonstrated that domain B contains the recognition site for NF-I (12, 26, 32, 38, 40). To determine whether domains B and C contain recognition sites for other initiation proteins, we analyzed the BR-FT fraction for the presence of sequence-specific DNA-binding proteins.

The method chosen to identify sequence-specific DNA-binding activities in the BR-FT was the gel electrophoresis DNA-binding assay (18, 19, 42). In this method radioactive DNA fragments are incubated with protein fractions and then subjected to gel electrophoresis. Specific protein-DNA complexes are detected by observing changes in the normal electrophoretic mobilities of the DNA fragments. In the present case we used three DNA fragments containing different segments of the adenovirus origin of replication (Fig. 2). The first fragment contained the entire adenovirus origin of replication (nucleotides 1 to 67), the second contained only domain A (nucleotides 1 to 18), and the third contained domains B and C (nucleotides 19 to 67). In each case the segment of adenovirus DNA was located adjacent

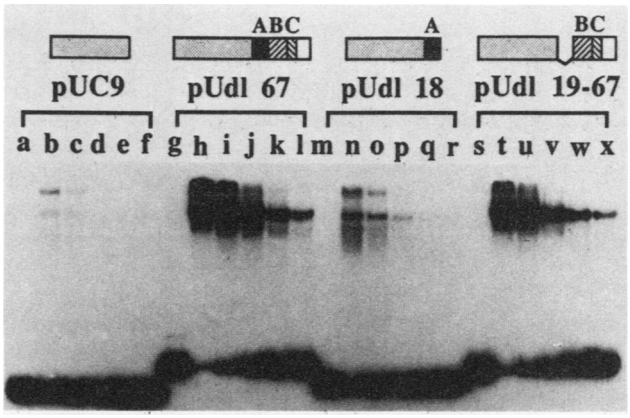


FIG. 2. Sequence-specific interactions between the BR-FT fraction and the adenovirus origin of DNA replication. The gel electrophoresis DNA-binding assay was used to detect sequence-specific DNA-protein interactions as described in Materials and Methods. Radioactive DNA fragments were incubated with various amounts of the BR-FT fraction and electrophoresed through a native polyacrylamide gel. The gel was dried, and the radioactive DNA was visualized by autoradiography. The DNA fragments were prepared by *Bam*HI-*Pvu*II cleavage of the plasmids pUC9 (101 base pairs), pU*dl*67 (157 base pairs), pU*dl*18 (108 base pairs), and pU*dl*19-67 (139 base pairs). These fragments (diagrammed at the top of the figure) contain the same 90-base-pair segment of pUC9 DNA, but differ in the sequences derived from the adenovirus origin of replication. The letters A, B, and C refer to the domains of the adenovirus origin. Lanes a, g, m, s, No added BR-FT; lanes b, h, n, t, 10 μg of BR-FT; lanes c, i, o, u, 5 μg of BR-FT; lanes d, j, p, v, 2.5 μg of BR-FT; lanes e, k, q, w, 1.25 μg of BR-FT. Lanes f, l, r, x, 0.625 μg of BR-FT.

to a 90-base-pair segment of pUC9 DNA. The same 90-base-pair segment lacking adenovirus sequences served as a control for nonspecific DNA binding. Serial twofold dilutions of the BR-FT fraction were incubated with each of the radioactively labeled DNA fragments and electrophoresed through a native polyacrylamide gel (Fig. 2). A low level of DNA-binding activity was observed with the control pUC9 fragment, but only at the highest concentration of BR-FT that was tested. A high level of specific DNA-binding activity, evident at all protein concentrations tested, was observed with the fragments that contained either the entire origin or only domains B and C. Since the BR-FT fraction did not contain detectable NF-I, it seemed likely that the major DNA-binding activity was directed against domain C. This sequence-specific binding activity was designated origin recognition protein C (ORP-C). A somewhat lower level of specific DNA-binding activity was observed with the fragment that contained only domain A. This activity was designated origin recognition protein A (ORP-A).

Purification and characterization of ORP-A and ORP-C. The BR-FT fraction was used as the source for the purification since it contained more than 95% of the ORP-A and ORP-C binding activities in HeLa cell nuclear extract. During the course of the purification, specific DNA-binding activities were detected by means of the gel electrophoresis binding assay as described above. Column fractions were assayed with radioactive DNA fragments that contained domain A alone, domains B and C, or only pUC9 DNA. By comparing the elution profiles obtained with the three fragments, it was easily possible to identify the peaks of ORP-A and ORP-C activity.

Initially, the BR-FT fraction was chromatographed on a

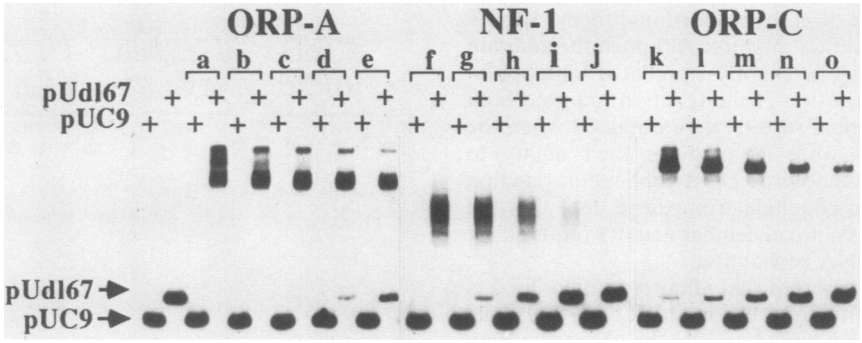


FIG. 3. Specific recognition of the adenovirus origin of DNA replication by ORP-A, NF-I, and ORP-C. The specific interactions between the adenovirus origin and purified recognition proteins were demonstrated by the gel electrophoresis DNA-binding assay described in Materials and Methods. ORP-A and ORP-C were purified as described in Materials and Methods, and NF-I was purified as previously described (39). Various amounts of ORP-A, NF-I, and ORP-C were incubated with radioactive DNA fragments prepared from pUC9 (101 base pairs) or pU*dl*67 (157 base pairs). The reaction mixtures were electrophoresed through a native polyacrylamide gel, and the radioactive DNA was visualized by autoradiography. ORP-A titration: lane a, 500 ng; lane b, 250 ng; lane c, 125 ng; lane d, 62.5 ng; lane e, 31.25 ng. NF-I titration: lane f, 20 ng; lane g, 10 ng; lane h, 5 ng; lane i, 2.5 ng; lane j, 1.25 ng. ORP-C titration: lane k, 400 ng; lane l, 200 ng; lane m, 100 ng; lane n, 50 ng; lane o, 25 ng.

calf thymus DNA-cellulose column which was developed with a linear gradient of NaCl (100 to 500 mM) followed by a 2 M NaCl step elution. Most of the ORP-A activity eluted at 250 mM NaCl. Most of the ORP-C activity was recovered in the 2 M NaCl elution, although a small fraction coeluted with ORP-A. The ORP-A activity was purified further by DNA recognition site affinity chromatography with the same affinity matrix previously used for the purification of NF-I (39). The matrix was prepared from a plasmid construct (pKB67-88) that contained 88 copies of the terminal 67 base pairs of the adenovirus genome. Elution of the affinity column with a linear gradient of NaCl effectively separated ORP-A from the residual ORP-C activity. Overall, ORP-A was purified more than 2,000-fold relative to nuclear extract, and ORP-C was enriched more than 200-fold. The DNA-binding properties of the purified ORP-A, ORP-C, and NF-I proteins were analyzed by the gel electrophoresis binding assay. Dilutions of the protein fractions were incubated with DNA fragments containing either nonspecific DNA (pUC9) or the complete adenovirus origin of replication (pU*dl*67), and the protein-DNA complexes were subjected to polyacrylamide gel electrophoresis as before (Fig. 3). As expected, the interaction between each protein and the viral origin was highly specific. In addition, each protein produced a characteristic change in the mobility of the origin fragment that differed from that produced by the others. This characteristic mobility change provides a simple alternative to the use of several different DNA fragments for distinguishing between the various origin recognition proteins during purification.

The recognition sites for ORP-A and ORP-C were characterized further by DNase I footprint analysis. Various concentrations of ORP-A, NF-I, and ORP-C were incubated with a radioactively labeled fragment from pU*dl*67 that contained the entire origin sequence, and the resulting protein-DNA complexes were digested briefly with DNase I. Each protein protected a single region of the origin located in the domain predicted from the results of the gel electrophoresis DNA binding assays (Fig. 4). The precise protection limits on both strands, summarized in Fig. 4C, were determined by examining several 8 and 12% polyacrylamide–urea gels that were electrophoresed for various periods. ORP-A protected a region that coincided with domain A of the adenovirus origin of replication. ORP-C protected a region that was largely contained within domain C. Although the segment protected by NF-I (base pairs 18 to 43) overlaps that protected by ORP-C, both proteins can be bound to the origin simultaneously as determined by both the gel electrophoresis gel-binding assay and the DNase I protection assay without any apparent changes in the protection patterns (unpublished data).

Analysis of NF-I recognition site. The essential base pairs required for NF-I recognition were identified by measuring the affinity of the protein for binding sites that contain base substitutions. The mutations were targeted to the region of the adenovirus origin that is protected from DNase I cleavage by NF-I. In this assay, fixed quantities of NF-I and a radioactive plasmid DNA (pKR67) containing the wild-type adenovirus origin were incubated with various concentrations of unlabeled competitor DNA. At each competitor concentration the fraction of the radioactive DNA bound to NF-I was determined, and a competition curve was generated (Fig. 5A). By comparing the amounts of two competitor DNAs required to decrease binding of the radioactive DNA to a given level, one can obtain a good measure of the relative affinities of NF-I for the two DNAs. The various mutant DNAs yielded several different types of competition curves (Fig. 5). A number of mutant DNAs (e.g., pUpm17, pUpm18, pUpm19, pUpm20, pUpm21, pUpm30, pUpm31, pUpm32, pUpm33, pUpm40, pUpm42) competed for binding to NF-I with an efficiency identical to that of wild-type DNA (pU*dl*67). Thus, the base substitutions present in these mutants do not alter the affinity of NF-I for its recognition site. Another class of mutants displayed competition curves identical to the curve observed with nonspecific DNA (pUC9). The base substitutions in these mutants (nucleotides 26, 35, and 36 of the adenovirus genome) abolish detectable specific NF-I binding, and thus, they must affect important contacts for the protein. Most of the remaining mutants (e.g., pUpm24, pUpm25, pUpm27, pUpm29, pUpm34, pUpm35, pUpm36, pUpm37, pUpm38) competed for NF-I binding with efficiencies intermediate between those of wild-type DNA and nonspecific DNA. The base substitutions in these mutants also define base pairs involved in specific binding by NF-I, although the mutations did not result in the complete loss of specificity. Finally, one mutant (pUpm28) bound NF-I with an affinity greater than that of the wild-type adenovirus origin sequence. It is of interest that the base

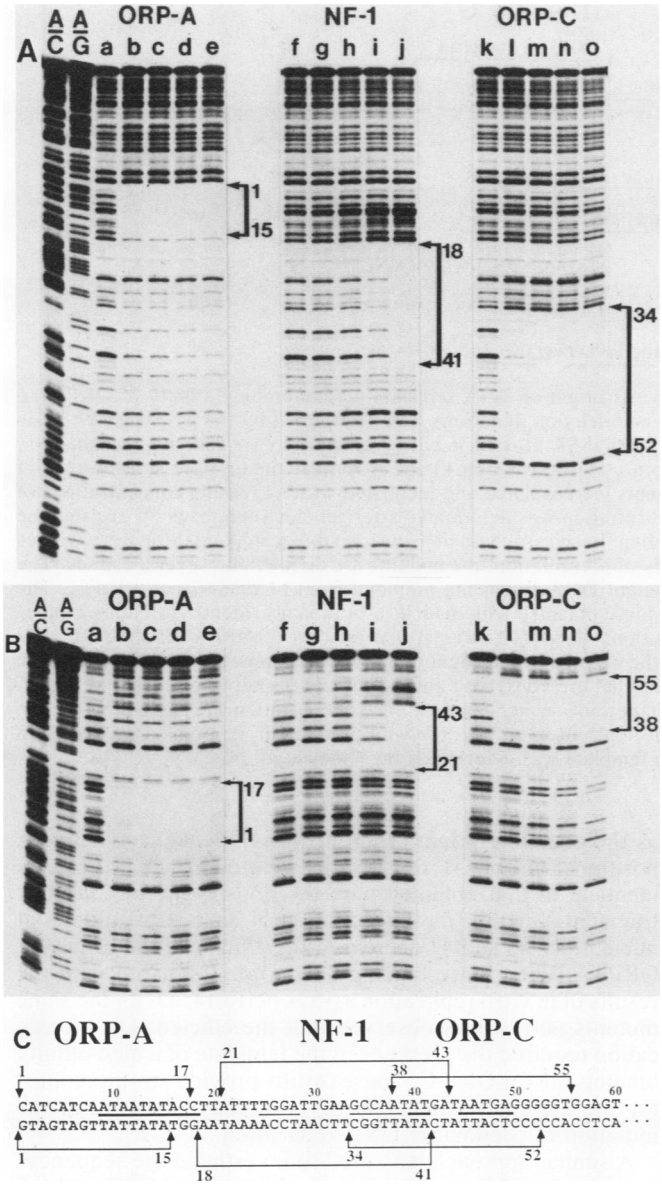


FIG. 4. DNase I footprint analysis of ORP-A, NF-I, and ORP-C interactions with the adenovirus origin of DNA replication. Footprinting reactions were performed with the 157-base-pair *Bam*HI to *Pvu*II fragment from the plasmid pU*dl*67, which contains nucleotides 1 to 67 of the adenovirus type 2 genome. The r strand (A) and the l strand (B) were labeled at their 5′ termini with ^{32}P as described in Materials and Methods. The labeled fragments were incubated in the absence of protein (lane a, f, and k) or with increasing amounts of ORP-A (lanes b to e), NF-I (lanes g to j), and ORP-C (lanes l to o). After a brief digestion with DNase I, the fragments were electrophoresed on a 12% (A) or an 8% (B) polyacrylamide-urea sequencing gel. A+C and A+G sequencing reactions were electrophoresed in parallel lanes. ORP-A titration: lane a, no added protein; lane b, 50 ng; lane c, 100 ng; lane d, 200 ng; lane e, 400 ng. NF-I titration: lane f, no added protein; lane g, 0.8 ng; lane h, 3.1 ng; lane i, 12.5 ng; lane j, 50 ng. ORP-C titration: lane k, no added protein; lane l, 400 ng; lane m, 800 ng; lane n, 1.6 μg; lane o, 3.2 μg. (C) Summary of DNase I footprint analysis. The regions protected by ORP-A, NF-I, and ORP-C on each strand are indicated by brackets. The underlined segments represent sequences that are conserved among most human adenovirus serotypes (23).

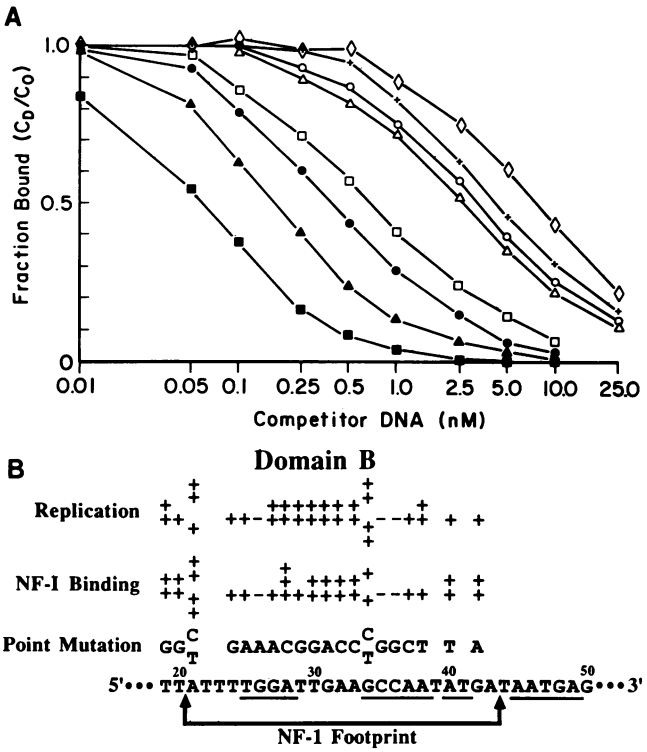


FIG. 5. Relative affinity of NF-I for DNA molecules containing point mutations within the adenovirus origin of DNA replication. Fixed concentrations of purified NF-I and ^{32}P-labeled pKB67 DNA, which contains a wild-type origin of DNA replication, were incubated with various concentrations of competitor DNA, and the extent of binding was determined by a quantitative nitrocellulose filter binding assay. (A) The concentration of radioactive DNA bound to NF-I at each competitor concentration (C_D) was normalized to the concentration of radioactive DNA bound to protein in the absence of competitor (C_0). For simplicity of presentation the mutants are divided into several classes. The wild-type DNA competition profile (▲) was identical to the profile observed with the point mutants pUpm17T, pUpm18T, pUpm19G, pUpm20G, pUpm21C, pUpm30G, pUpm31A, pUpm32C, pUpm33C, pUpm40T, and pUpm42A. The nonspecific DNA competition profile (◇) (pUC9) was identical to the profile observed with the point mutants pUpm26A, pUpm35G, and pUpm36G. Identical competition curves were obtained with point mutants pUpm24G and pUpm25A (×) and with point mutants pUpm27A, pUpm34C, and pUpm37C (○). The remaining competition profiles are as follows: pUpm28C (■), pUpm29G (●), pUpm34T (□), and pUmp38T (△). Each competition assay was performed at least twice, and the average value for each competitor concentration was plotted. The error associated with each competitor concentration was less than 5% of the average value. (B) Correlation between NF-I binding and replication efficiency of mutant DNA molecules. The boundaries of the NF-I footprint (bracket) on the terminal sequence of the adenovirus type 2 genome are shown. Sequence elements that are conserved among adenovirus serotypes are underlined (23). Each base substitution is shown above the wild-type sequence with an indication of its effect on the binding affinity of NF-I and the efficiency of replication (49). Two plus signs indicate that the mutation had no effect on the binding affinity or no effect on replication efficiency. A minus sign indicates that the mutation abolished specific NF-I binding or reduced replication efficiency 10-fold. A single plus sign indicates that NF-I recognized the mutant sequence with reduced affinity or that replication efficiency was diminished approximately fourfold. Three plus signs indicate that NF-I recognized the mutant DNA with a fourfold-greater affinity.

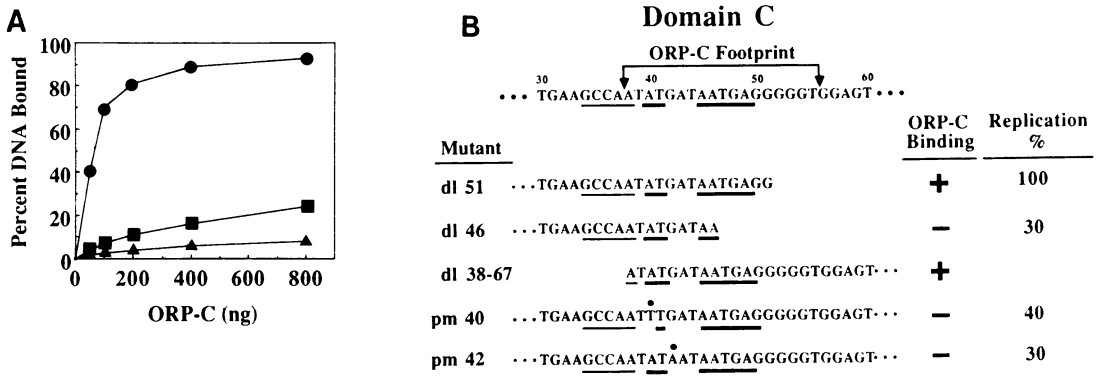


FIG. 6. Sequence requirements for ORP-C recognition of the adenovirus origin of DNA replication. The binding affinity of ORP-C to adenovirus type 2 origin sequences containing various base substitution and deletion mutations was determined by the gel electrophoresis DNA-binding assay. The plasmids pU*dl*67, pU*dl*51, and pU*dl*46 contain the first 67, 51, and 46 base pairs of the viral genome, respectively, and the plasmid pU*dl*38-67 contains the segment from nucleotides 38 to 67. Plasmids pUpm40 and pUpm42 contain base substitutions at nucleotides 40 and 42, respectively. Radioactive adenovirus DNA fragments were isolated and incubated with increasing concentrations of ORP-C. The percentage of the input fragment bound at each ORP-C concentration was determined after gel electrophoresis by excising the appropriate band and quantitating radioactivity by liquid scintillation counting. Each assay was performed in duplicate, and the average values were plotted. Duplicate samples varied by less than 5%. (A) For simplicity of presentation, the mutants are divided into several classes. The wild-type DNA binding profile (pU*dl*67) (●) was also observed with mutant DNA fragments prepared from pU*dl*38-67 and pU*dl*51. The nonspecific DNA binding profile (pUC9) (▲) was identical to the binding curve obtained with pU*dl*46. Additionally, identical binding profiles were observed with DNA fragments containing point mutations at positions 40 and 42 (■). (B) Correlation of ORP-C binding with the replication efficiency of the mutant templates. The terminal sequence of the r strand of the adenovirus type 2 genome from position 30 to 60 is shown with the boundaries of the ORP-C footprint. Sequence elements that are conserved among adenovirus serotypes are underlined. Deletion mutations and base substitutions that were used to identify important sequences for ORP-C recognition are shown below the wild-type sequence. A plus sign reflects an affinity for ORP-C identical to that observed with the wild-type origin. A minus sign signifies a significant decrease in binding affinity. The replication efficiency of each template was determined by Wides et al. (49).

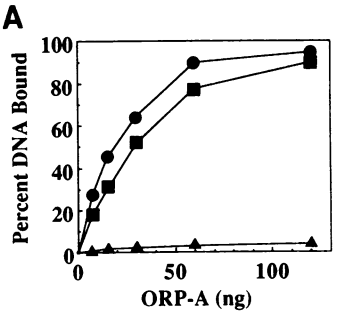
substitution at position 28 increases the overall symmetry of the NF-I recognition site.

The results of the binding studies with the mutant templates are summarized in Fig. 5B. Mutations that significantly affect the affinity of NF-I for the viral origin of replication lie in two clusters of five base pairs (half sites) separated by a spacer of five base pairs. Mutations in the spacer segment had minimal (approximately twofold) effect on NF-I binding. The two half sites that are required for recognition are about one turn of the helix apart and display approximate dyad symmetry. As described above, the affinity of NF-I binding can be increased by increasing the symmetry of the site. Thus, the highest-affinity binding site identified in this study contains the sequence TTGGCN$_5$GC CAA.

In the accompanying paper we have described the effects of the same set of base substitution mutations on the initiation of adenovirus DNA replication in vitro (49). For purposes of comparison, the results of these replication studies are also summarized in Fig. 5B. The data clearly indicate that mutations within the NF-I recognition site that significantly decrease the efficiency of DNA replication also decrease the affinity of NF-I for the site. These results provide additional evidence that binding of NF-I to the origin is essential for NF-I function.

Analysis of recognition sites for ORP-A and ORP-C. Mutations within domain C of the adenovirus origin were studied to identify the base pairs required for specific binding of ORP-C. The mutations were constructed as described by Wides et al. (49), and the affinity of ORP-C for the mutant DNAs was quantitated by the gel electrophoresis DNA-binding assay. The various mutant DNAs yielded three types of binding profiles (Fig. 6). Deletion mutants that preserved sequences between position 38 and 51 of the adenovirus origin were capable of binding ORP-C with the same affinity as the complete origin. The deletion of sequences between positions 46 and 51 resulted in a binding curve that was identical to that obtained with the nonspecific pUC9 DNA fragment. Finally, the introduction of base substitutions at nucleotides 40 and 42 decreased the affinity of the origin for ORP-C. When these binding data were compared with the results of in vitro replication assays performed with the same mutants (49), it was observed that the efficient DNA replication required the presence in the template of a high-affinity binding site for ORP-C. These results provide strong support for the hypothesis that ORP-C is directly involved in the initiation of adenovirus DNA replication.

A similar approach was used to investigate the sequences required for ORP-A recognition and the role of ORP-A in the replication of adenovirus DNA. The ORP-A recognition site was localized to within the terminal 12 base pairs of the adenovirus genome (Fig. 7A). A deletion mutant containing the terminal 12 base pairs bound ORP-A with the same affinity as the complete origin, while a deletion mutant that contained the terminal 7 base pairs did not bind ORP-A to any significant extent. In these constructs the segment of adenovirus DNA was preceded by a 90-base-pair segment of pUC9 DNA. To be certain that nucleotides in the pUC9 segment did not contribute to ORP-A binding, we also studied a DNA fragment in which domain A was positioned at the molecular terminus, identical to its location in the adenovirus genome. This construct bound ORP-A with an affinity equivalent to that of constructs containing pUC9 sequences (data not shown). We conclude from these results that the recognition site for ORP-A is located between nucleotide 1 and nucleotide 12 of the adenovirus genome. Since the terminal 12 base pairs of adenovirus DNA are not sufficient for minimal origin function (10, 25, 38, 46, 49), we have not yet been able to correlate a loss of ORP-A binding with an effect on replication.



B

Domain A

ORP-A Footprint

5' CATCATCAATAATATACCTT · · ·
10 20

Mutant		ORP-A Binding	Replication %
dl 18	CATCATCAATAATATACC	+	3
dl 12	CATCATCAATAA	+	0
dl 7	CATCATC	−	0
pm 4	CATTATCAATAATATACCTT · · ·	+	100
pm 17	CATCATCAATAATATATCTT · · ·	+	3
pm 18	CATCATCAATAATATACTTT · · ·	+	3

FIG. 7. Sequence requirements for ORP-A binding. The affinity of ORP-A for various base substitution and deletion mutants was determined by the gel electrophoresis DNA-binding assay. Plasmids pU*dl*67, pU*dl*18, pU*dl*12, and pU*dl*7 contain the first 67, 18, 12, and 7 base pairs of the adenovirus genome, respectively. Plasmids pUpm4, pUpm17, and pUpm18 contain base substitutions at nucleotides 4, 17, and 18, respectively. Radioactive adenovirus DNA fragments were isolated and incubated with various amounts of ORP-A. The percentage of input DNA bound by ORP-A was determined by quantitation of the unbound radioactive DNA fragment after gel electrophoresis. Each assay was performed in duplicate, and the average values are plotted. Duplicate samples varied by less than 5%. (A) For simplicity of presentation, the mutants are divided into several classes. The wild-type DNA binding profile (pU*dl*67) (●) was observed with fragments prepared from pUpm4, pUpm17, and pUpm18. Identical binding curves were observed with fragments containing deletion mutations prepared from the plasmids pU*dl*18 and pU*dl*12 (■). The nonspecific DNA binding profile (pUC9) (▲) was identical to the binding curve obtained with pU*dl*7. (B) Summary of ORP-A binding affinities and replication efficiencies of mutant templates. The terminal sequence of the r strand of the adenovirus type 2 genome from position 1 to 20 is shown together with the boundaries of the ORP-A footprint. The underlined sequence is conserved among all adenovirus serotypes. The sequences of relevant mutant DNAs are shown below the wild-type sequence. The ORP-A binding and replication efficiencies of these mutant templates are compared. A plus represents a high-affinity interaction between ORP-A and the template, while a minus represents a loss of specific binding. The replication efficiencies were determined by Wides et al. (49).

Genetic studies of adenovirus DNA replication in vitro have revealed that sequences between nucleotides 12 and 18 of domain A are essential for minimal origin function (10, 25, 38, 46, 49). For example, point mutations at positions 17 and 18 (Fig. 7B) and at positions 20 and 21 (Fig. 5B) result in decreased replication efficiency. Since these mutations map outside of the region required for ORP-A binding, it is possible that an additional sequence-specific binding activity that recognizes the sequence between the ORP-A and NF-I binding sites is required for the replication of adenovirus DNA. Although we have no direct evidence for the existance of such a factor, it is worth pointing out that we have so far been unable to reconstitute efficient initiation using the purified ORP-A, ORP-C, and NF-I (Fig. 8). The amounts of ORP-A and ORP-C added to the initiation assays were comparable to the amounts of these proteins in the BR-FT fraction. However, when added separately or together, ORP-A and ORP-C could not substitute for the BR-FT fraction. Although ORP-A and ORP-C did stimulate a low level of pTP-dCMP formation in the presence of NF-I, increasing amounts of these proteins had no additional effect on the initiation signal. We are currently investigating prior steps in the purification to identify fractions that are capable of fully reconstituting the initiation of adenovirus DNA replication in vitro.

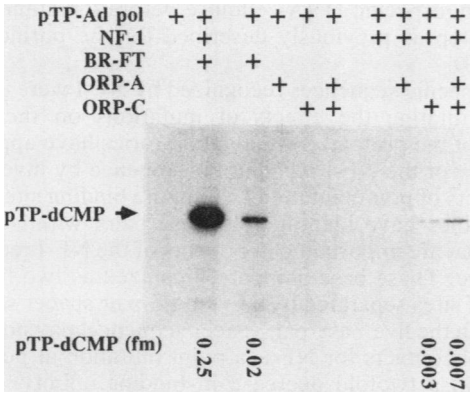


FIG. 8. Effects of ORP-A and ORP-C on the initiation of adenovirus DNA replication in vitro. The ability of ORP-A and ORP-C to substitute for the BR-FT fraction in the formation of the pTP-dCMP initiation complex was assayed as described in Materials and Methods. The viral proteins required for initiation (pTP-Ad pol) were added as the purified protein complex (0.6 μg; 0.004 pol units). The following amounts of cellular proteins were added where indicated: purified NF-I, 20 ng; Bio-Rex 70 flowthrough fraction (BR-FT), 7.5 μg; ORP-A, 20 ng; ORP-C, 200 ng. The amounts of ORP-A and ORP-C binding activity used for reconstitution were similar to the amounts of binding activity present in the BR-FT fraction. fm, 10^{-15} moles.

DISCUSSION

We identified two proteins from HeLa cells, distinct from the previously characterized NF-I, that are capable of site-specific recognition of the adenovirus origin of DNA replication. These proteins were discovered during the course of an investigation of the role of NF-I in the initiation of adenovirus DNA replication. NF-I was originally identified as a protein fraction capable of restoring efficient replication activity when added to the purified virus-encoded replication proteins (pTP and Ad pol) (31). It was subsequently shown that this fraction contained a protein that recognized a specific sequence element in the adenovirus origin of replication (12, 26, 32, 38, 40). We have exploited the specificity of the interaction between NF-I and the origin to develop a practical affinity purification method (39). In this report we showed that highly purified NF-I obtained by this affinity purification scheme is not sufficient to reconstitute the initiation reaction when incubated with the viral proteins. We also identified a fraction derived from HeLa cell nuclear

extract (the Bio-Rex 70 flowthrough fraction or BR-FT) that contains the additional cellular factors required for the complete reconstitution of the initiation reaction in vitro. Our initial studies of the latter fraction revealed the existence of the two new origin recognition proteins, which were subsequently purified by affinity chromatography and characterized. These proteins and NF-I are still not sufficient for the reconstitution of efficient initiation, so it seems likely that the BR-FT fraction harbors additional factors that are required for the reaction.

In contrast to our results, Diffley and Stillman (13) have reported that affinity-purified NF-I was sufficient to reconstitute the initiation reaction in the presence of pTP-Ad pol and the 72-kDa ss DBP. The reaction conditions used in the latter study were similar to our conditions except for the inclusion of the ss DBP in the reconstitution assays. Previously, Nagata et al. (31) reported that the addition of the 72-kDa ss DBP stimulated the initiation of adenovirus DNA replication. Our data indicate that ss DBP has no effect on efficiency of the initiation reaction. This finding is consistent with studies of temperature-sensitive mutations in the ss DBP gene done by several laboratories (9, 17, 37, 48). It was shown that these mutations decrease the efficiency of DNA synthesis both in vivo and in vitro but have no effect on the initiation of viral DNA replication in vitro. In retrospect, the most likely explanation for the stimulation of initiation observed by Nagata et al. upon addition of ss DBP was that the ss DBP preparation contained BR-FT activity. Similarly, the results obtained by Diffley and Stillman (13) may be explained by the presence of contaminating BR-FT activity in one of their protein preparations.

To further study the BR-FT fraction, we focused on the sequences within the adenovirus origin of replication that are required for optimal initiation of adenovirus DNA replication. In the accompanying paper (49) we present genetic evidence that the origin is composed of at least three functionally distinct domains, A, B, and C. Domain A (nucleotides 1 to 18) contains the minimal sequence sufficient for origin function. Domains B (nucleotides 19 to 40) and C (nucleotides 41 to 51) contain accessory elements that significantly increase the activity of the minimal origin. NF-I was previously shown to recognize an essential sequence element in domain B, and the importance of this interaction for the initiation of replication was confirmed by demonstrating that mutations within domain B have similar effects on binding affinity and replication efficiency (12, 26, 32, 38). Using the same strategy, we tested the BR-FT fraction for activities capable of recognizing the remaining two domains of the adenovirus origin, domain A and domain C. Initially, we searched for additional sequence-specific binding activities using the nitrocellulose filter binding assay and DNase I protection analysis. Both methods yielded equivocal results for technical reasons (see below). However, the gel electrophoresis DNA-binding assay (18, 19, 42) provided direct evidence that the BR-FT fraction contained sequence-specific proteins that recognize the adenovirus origin. As predicted, the DNA sequences recognized by the BR-FT fraction included domain A and domain C, and thus these origin recognition proteins were designated ORP-A and ORP-C, respectively. Our ability to detect these binding proteins was dependent on the addition of poly(dI-dC)-poly(dI-dC) as a competitor for nonspecific DNA-binding activity (42). Moreover, the addition of Nonidet P-40 to both the sample buffer and the polyacrylamide gel greatly enhanced the apparent affinity of these proteins for their recognition sequence.

The gel electrophoresis DNA-binding assay has distinct advantages over the nitrocellulose filter binding assay and DNase I protection analysis for the detection of sequence-specific binding proteins. Most notably, the utility of DNase I protection analysis is limited by the low concentration of the binding proteins in many preparations, since the assay requires nearly complete occupancy of the recognition site. While the nitrocellulose filter binding assay is useful at less than saturating protein concentrations, the technique is limited by the inability of some proteins to simultaneously bind DNA and nitrocellulose. Moreover, both methods are relatively sensitive to the presence of contaminating nonspecific DNA-binding proteins. However, the gel electrophoresis DNA-binding assay appears to be generally applicable, especially for the detection of low-abundance recognition proteins. Moreover, the binding proteins can be identified not only on the basis of specificity of binding, but also by the characteristic mobility of the protein-DNA complex. In particular, the multiple banding pattern observed with NF-I was characteristic for the protein and reminiscent of the multiple polypeptides observed by silver staining after SDS-polyacrylamide gel electrophoresis of the purified protein (39). Thus, the gel binding assay may be useful to assess the structural homogeneity of a given recognition protein.

The purification of ORP-A and ORP-C was accomplished by DNA recognition site affinity chromatography. ORP-A was purified over 2,000-fold from nuclear extract by applying the same strategy used for the purification of NF-I (39). Recently, ORP-C has been purified over 8,000-fold by an identical approach (E. O'Neill and T. Kelly, unpublished data). Thus, three different sequence-specific DNA-binding proteins have now been successfully purified by DNA recognition site affinity chromatography, suggesting that the method is generally applicable to the rapid and efficient purification of such proteins. It should be noted that the elution profile of a particular DNA-binding protein from the specific affinity matrix depends on the affinity of the protein for its recognition site. Therefore, to obtain optimal purification with this approach, it is important to remove contaminating nonspecific DNA-binding proteins by chromatography on nonspecific DNA-cellulose before the affinity purification step as previously described for the purification of NF-I (39).

The specific sequences recognized by NF-I were identified by quantitating the effects of mutations on the binding affinity of the protein. Several laboratories have approached the study of the NF-I recognition sequence by investigating the effects of point mutations within the binding site (12, 40). Our studies have identified 10 base pairs within the viral origin that are important components of the NF-I recognition sequence. These base pairs are organized as two five-base-pair half sites separated by a five-base-pair spacer sequence. Although the five-base-pair spacer sequence may not contain essential contacts for NF-I, a point mutation at position 29 results in a twofold decrease in binding affinity. Interestingly, this particular base change is not observed in any naturally occurring NF-I site (20, 33, 38, 39, 41). The five-base-pair half sites contain sequences that are conserved among most adenovirus serotypes. These conserved sequences reflect essential contacts within domain B that are necessary for NF-I recognition and optimal replication efficiency of the template. These results are consistent with the consensus sequence (TGG$^{A}_{C}$NNNNNGCCAA) obtained by comparing NF-I-binding sites identified in viral and cellular DNAs (4, 20, 33, 38, 39, 41). The two half sites that are essential for NF-I recognition are related by twofold rota-

tional symmetry, although in the case of the adenovirus origin the symmetry is not perfect. The observation that a mutation that increases the symmetry of the NF-I recognition site also increases binding affinity strongly suggests that NF-I binds DNA as a dimer. It has also been shown by DNase I footprint analysis that NF-I can recognize a half site (unpublished data). This observation has prompted an investigation into the similarities between NF-I and the CCAAT box transcription factor (22). The two proteins have been shown to recognize identical sequences and to stimulate both transcription and adenovirus DNA replication in vitro (K. Jones, P. Rosenfeld, T. Kelly, and R. Tjian, Cell, in press). Moreover, the two proteins exhibit bands of identical mobility in the gel electrophoresis DNA-binding assay and by SDS-polyacrylamide gel analysis.

The ORP-C-binding domain contains two blocks of sequences that are conserved among all adenovirus serotypes (ATNNNAATGA). The importance of these conserved sequences for recognition by ORP-C was demonstrated by quantitating the effects of deletion and point mutations within domain C. Moreover, the correlation between the binding affinity and the replication efficiency of these mutant templates strongly suggests that ORP-C is required for optimal replication of adenovirus DNA. The inability of ORP-C to bind a point mutant at position 40 demonstrates that only one base pair separates the recognition sequences of NF-I and ORP-C. Although the footprints of the two proteins overlap, both NF-I and ORP-C are capable of binding the DNA simultaneously (unpublished data).

The ORP-C recognition site closely resembles the octamer sequence that has been implicated in the regulation of transcription of several genes including the immunoglobulin (15, 16, 35) and histone H2B (H. L. Sive and R. G. Roeder, Proc. Natl. Acad. Sci. USA, in press) genes. It seems possible that ORP-C or a closely related protein is responsible for the octamer-binding activities that have been reported (42; Sive and Roeder, in press). Using the gel electrophoresis DNA-binding assay, we have recently obtained evidence that purified ORP-C binds specifically to the octamer sequence adjacent to the histone H2B gene (E. O'Neill and T. Kelly, unpublished data). ORP-C bound the recognition site in adenovirus type 2 DNA (ATGATAAT) and the canonical octamer sequence (ATGCAAAT) with similar affinity, but we have not determined whether these sequences are functionally interchangeable for replication and transcription. Interestingly, the origins of other adenovirus serotypes, such as adenovirus types 4, 9, and 10, contain exact copies of the octamer sequence but do not contain an NF-I site (23).

The conserved sequences contained within domain A were shown to be essential for the replication of the adenovirus genome. However, these conserved sequences are not entirely required for ORP-A recognition. Deletion of domain A from position 13 to 18 had no effect on ORP-A binding, but resulted in the loss of minimal origin function. Thus, with the currently available mutants we were not able to correlate the loss of ORP-A recognition with an effect on replication. We are constructing point mutations within the ORP-A-binding site to investigate the role of ORP-A in the replication of adenovirus DNA.

The inability of ORP-A, NF-I, and ORP-C to fully substitute for nuclear extract in a reconstituted initiation assay suggests that additional cellular factors are required. Furthermore, optimal replication efficiency requires conserved origin sequences that are not recognized by ORP-A, NF-I, or ORP-C, suggesting that at least one of the additional cellular factors is a sequence-specific binding protein. The recent identification of NF-I as a CCAAT transcription factor and the possibility that ORP-C is the octamer-binding protein suggest that the recognition proteins required for the initiation of adenovirus DNA replication serve a transcriptional role in vivo. While it is possible that these cellular transcriptional proteins are subverted by the virus for the replication of its DNA, an alternative possibility is that these proteins function in a dual role as initiators of cellular replication and transcription. Future comparisons between the initiation mechanisms of transcription and replication will depend on the in vitro reconstitution of these cellular processes with purified proteins.

ACKNOWLEDGMENTS

We are indebted to our colleagues Joachim Li, James Sherley, Jordan Kriedberg, Marc Wold, David Weinberg, and Lorne Erdile for helpful discussions and many useful suggestions. We also thank Brenda Athey for assistance in preparing the manuscript.

This study was supported by Public Health Service grant CA16519 from the National Cancer Institute. P.J.R. is a trainee of the Medical Scientist Training Program (GM07309). E.A.O. is a Monsanto Co. postdoctoral fellow.

ADDENDUM IN PROOF

Pruijn et al. (G. J. M. Pruijn, W. van Driel, and P. C. van der Vliet, Nature [London] **322**:656–659, 1986) have recently reported the identification of a HeLa nuclear protein designated Nuclear Factor III which has properties similar to those of ORP-C.

LITERATURE CITED

1. **Alberts, B., and G. Herrick.** 1971. DNA-cellulose chromatography. Methods Enzymol. **21**:198–217.
2. **Aposhian, H. V., and A. Kornberg.** 1962. Enzymatic synthesis of deoxyribonucleic acid. J. Biol. Chem. **237**:519–525.
3. **Bencini, D. A., G. A. O'Donovan, and J. R. Wild.** 1984. Rapid chemical degradation sequencing. Biotechniques **2**:4–5.
4. **Borgmeyer, U., J. Nowock, and A. E. Sippel.** 1984. The TGGCA-binding protein: a eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded linear DNA. Nucleic Acids Res. **12**:4295–4311.
5. **Bradford, M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
6. **Burton, K. A.** 1956. A study for the conditions and mechanisms of the diphenylamine reaction for the colorimetric estimation of deoxyribonucleic acid. Biochem. J. **62**:315–321.
7. **Challberg, M. D., S. V. Desiderio, and T. J. Kelly, Jr.** 1980. Adenovirus DNA replication *in vitro*: characterization of a protein covalently linked to nascent DNA strands. Proc. Natl. Acad. Sci. USA **77**:5105.
8. **Challberg, M. D., and T. J. Kelly, Jr.** 1979. Adenovirus DNA replication *in vitro*. Proc. Natl. Acad. Sci. USA **76**:655–659.
9. **Challberg, M. D., J. M. Ostrove, and T. J. Kelly, Jr.** 1982. Initiation of adenovirus DNA replication: detection of covalent complexes between nucleotide and the 80-kilodalton terminal protein. J. Virol. **41**:265–270.
10. **Challberg, M. D., and D. R. Rawlins.** 1984. Template requirements for the initiation of adenovirus DNA replication. Proc. Natl. Acad. Sci. USA **81**:100–104.
11. **Desiderio, S. V., and T. J. Kelly, Jr.** 1981. The structure of the linkage between adenovirus DNA and the 55,000 dalton terminal protein. J. Mol. Biol. **145**:319–337.
12. **de Vries, E., W. van Driel, M. Tromp, J. van Boom, and P. C. van der Vliet.** 1985. Adenovirus DNA replication in vitro: site-directed mutagenesis of the nuclear factor I binding site of the Ad2 origin. Nucleic Acids Res. **13**:4935–4952.
13. **Diffley, J. F. X., and B. Stillman.** 1986. Purification of a cellular,

double-stranded DNA-binding protein required for initiation of adenovirus DNA replication by using a rapid filter-binding assay. Mol. Cell. Biol. **6**:1363–1373.
14. **Enomoto, T., J. H. Lichy, J.-E. Ikeda, and J. Hurwitz.** 1981. Adenovirus DNA replication in vitro: purification of the terminal protein in a functional form. Proc. Natl. Acad. Sci. USA **78**:6779–6783.
15. **Ephrussi, A., G. M. Church, S. Tonegawa, and W. Gilbert.** 1985. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors *in vivo*. Science **227**:134–140.
16. **Falkner, F. G., and H. G. Zachan.** 1984. Correct transcription of an immunoglobulin K gene requires an upstream fragment containing conserved sequence elements. Nature (London) **310**: 71–74.
17. **Freifeld, B. R., M. D. Krevolin, and M. S. Horwitz.** 1983. Effects of the adenovirus H5ts 125 and H5ts 107 DNA binding proteins on DNA replication *in vitro*. Virology **124**:380–389.
18. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
19. **Garner, M. M., and A. Revzin.** 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. **9**:3047–3060.
20. **Gronostajski, R. M., S. Adhya, K. Nagata, R. A. Guggenheimer, and J. Hurwitz.** 1985. Site-specific DNA binding of nuclear factor I: analyses of cellular binding sites. Mol. Cell. Biol. **5**:964–971.
21. **Guggenheimer, R. A., B. W. Stillman, K. Nagata, F. Tamanoi, and J. Hurwitz.** 1984. DNA sequences required for the in vitro replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA **81**:3069–3073.
22. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell **42**:559–572.
23. **Kelly, T. J., Jr.** 1984. Adenovirus DNA replication, p. 271–308. *In* H. S. Ginsberg (ed.), The adenoviruses, vol. 19. Plenum Publishing Corp., New York.
24. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
25. **Lally, C., T. Dorper, W. Groger, G. Antoine, and E.-L. Winnacker.** 1984. A size analysis of the adenovirus replicon. EMBO J. **3**:333–337.
26. **Leegwater, P. A., W. van Driel, and P. C. van der Vliet.** 1985. Recognition site of nuclear factor I, a sequence-specific DNA-binding protein from HeLa cells that stimulates adenovirus DNA replication. EMBO J. **4**:1515–1521.
27. **Lichy, J. H., J. Field, M. S. Horwitz, and J. Hurwitz.** 1982. Separation of the adenovirus terminal protein precursor from its associated DNA polymerase: role of both proteins in the initiation of adenovirus DNA replication. Proc. Natl. Acad. Sci. USA **79**:5225–5229.
28. **Lichy, J. H., M. S. Horwitz, and J. Hurwitz.** 1981. Formation of a covalent complex between the 80,000-dalton adenovirus terminal protein and 5′-dCMP in vitro. Proc. Natl. Acad. Sci. USA **78**:2678–2682.
29. **Maniatis, T., E. F. Fritsch, and J. Sambrook (ed.).** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
31. **Nagata, K., R. A. Guggenheimer, T. Enomoto, J. H. Lichy, and J. Hurwitz.** 1982. Adenovirus DNA replication in vitro: identification of a host factor that stimulates synthesis of the preterminal protein-dCMP complex. Proc. Natl. Acad. Sci. USA **79**:6438–6442.
32. **Nagata, K., R. A. Guggenheimer, and J. Hurwitz.** 1983. Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA **80**:6177–6181.
33. **Nowock, J., U. Borgmeyer, A. N. Puschel, R. A. W. Rupp, and A. E. Sippel.** 1985. TGGCA protein binds to the MMTV-LTR, the adenovirus origin of replication, and the BK virus enhancer. Nucleic Acids Res. **13**:2045–2061.
34. **Ostrove, J. U., P. Rosenfeld, J. Williams, and T. J. Kelly, Jr.** 1983. In vitro complementation as an assay for purification of adenovirus DNA replication proteins. Proc. Natl. Acad. Sci. USA **80**:935–939.
35. **Parslow, T. G., D. L. Blair, W. J. Murphy, and D. K. Granner.** 1984. Structure of the 5′ ends of immunoglobulin genes: a novel conserved sequence. Proc. Natl. Acad. Sci. USA **81**:2650–2654.
36. **Pincus, S., W. Robertson, and D. M. K. Rekosh.** 1981. Characterization of the effect of aphidicolin on adenovirus DNA replication: evidence in support of a protein primer model of initiation. Nucleic Acids Res. **9**:4919–4938.
37. **Prelich, G., and B. W. Stillman.** 1986. Functional characterization of thermolabile DNA-binding proteins that affect adenovirus DNA replication. J. Virol. **57**:883–892.
38. **Rawlins, D. R., P. J. Rosenfeld, R. J. Wides, M. D. Challberg, and T. J. Kelly, Jr.** 1984. Structure and function of the adenovirus origin of replication. Cell **37**:309–319.
39. **Rosenfeld, P. J., and T. J. Kelly, Jr.** 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. J. Biol. Chem. **261**:1398–1408.
40. **Schneider, R., I. Gander, U. Muller, R. Metz, and E.-L. Winnacker.** 1986. A sensitive and rapid gel retention assay for nuclear factor I and other DNA-binding proteins in crude nuclear extracts. Nucleic Acids Res. **14**:1303–1317.
41. **Siebenlist, U., L. Hennighausen, J. Battey, and P. Leder.** 1984. Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. Cell **37**:381–391.
42. **Singh, H., R. Sen, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. 1986. Nature (London) **319**:154–158.
43. **Stillman, B. W.** 1981. Adenovirus DNA replication in vitro: a protein linked to the 5′ end of the nascent DNA strands. J. Virol. **37**:139.
44. **Stillman, B. W., F. Tamanoi, and M. B. Mathews.** 1982. Purification of an adenovirus-coded DNA polymerase that is required for initiation of DNA replication. Cell **31**:613–623.
45. **Tamanoi, F., and B. W. Stillman.** 1982. Function of adenovirus terminal protein in the initiation of DNA replication. Proc. Natl. Acad. Sci. USA **79**:2221–2225.
46. **Tamanoi, F., and B. W. Stillman.** 1983. Initiation of adenovirus DNA replication in vitro requires a specific DNA sequence. Proc. Natl. Acad. Sci. USA **80**:6446–6450.
47. **van Bergen, B. G. M., P. A. van der Ley, W. van Driel, A. D. M. van Mansfeld, and P. C. van der Vliet.** 1983. Replication of origin containing adenovirus DNA fragments that do not carry the terminal protein. Nucleic Acids Res. **11**:1975–1989.
48. **van Bergen, B. G. M., and P. C. van der Vliet.** 1983. Temperature-sensitive initiation and elongation of adenovirus DNA replication in vitro with nuclear extracts from H5ts36-, H5ts149-, and H5ts125-infected HeLa cells. J. Virol. **46**:642–648.
49. **Wides, R. J., M. D. Challberg, D. R. Rawlings, and T. J. Kelly.** 1986. Adenovirus origin of DNA replication: sequence requirements for replication in vitro. Mol. Cell. Biol. **7**:864–874.