Conserved Sequences at the Origin of Adenovirus DNA Replication

BRUCE W. STILLMAN,* WILLIAM C. TOPP, AND JEFFREY A. ENGLER

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724

Received 23 April 1982/Accepted 13 July 1982

The origin of adenovirus DNA replication lies within an inverted sequence repetition at either end of the linear, double-stranded viral DNA. Initiation of DNA replication is primed by a deoxynucleoside that is covalently linked to a protein, which remains bound to the newly synthesized DNA. We demonstrate that virion-derived DNA-protein complexes from five human adenovirus serological subgroups (A to E) can act as a template for both the initiation and the elongation of DNA replication in vitro, using nuclear extracts from adenovirus type 2 (Ad2)-infected HeLa cells. The heterologous template DNA-protein complexes were not as active as the homologous Ad2 DNA, most probably due to inefficient initiation by Ad2 replication factors. In an attempt to identify common features which may permit this replication, we have also sequenced the inverted terminal repeated DNA from human adenovirus serotypes Ad4 (group E), Ad9 and Ad10 (group D), and Ad31 (group A), and we have compared these to previously determined sequences from Ad2 and Ad5 (group C), Ad7 (group B), and Ad12 and Ad18 (group A) DNA. In all cases, the sequence around the origin of DNA replication can be divided into two structural domains: a proximal A · Trich region which is partially conserved among these serotypes, and a distal $G \cdot C$ rich region which is less well conserved. The $G \cdot C$ -rich region contains sequences similar to sequences present in papovavirus replication origins. The two domains may reflect a dual mechanism for initiation of DNA replication: adenovirusspecific protein priming of replication, and subsequent utilization of this primer by host replication factors for completion of DNA synthesis.

Adenovirus DNA replicates by a strand displacement mechanism from origins of replication that are located within an inverted repeated sequence at each end of the linear doublestranded DNA (see the review by Winnacker [37]). Replication is initiated at either origin, and elongation of DNA chains proceeds toward the opposite end of the molecule, displacing the nontemplate DNA strand (type I replication [16]). The displaced strand is then replicated as a result of a separate but analogous initiation event (type II replication).

The viral DNA has a 55,000-dalton protein (terminal protein, TP) covalently linked to each 5' terminus (21–23) via a phosphodiester bond between the β -OH of a serine residue in the protein and the 5'-OH of the terminal deoxycytidine residue (2, 9). The terminal protein is synthesized as an 80,000- to 87,000-dalton precursor protein (pTP) (2, 31) which is cleaved to TP late in infection (6). The pTP is encoded by early region E2b on the adenovirus genome, which forms part of a complex transcription unit encoding other replication proteins (31). Intracellular DNA (6, 8, 14, 29, 30, 36) and nascent

DNA strands replicated in vitro (2, 28) also contain protein linked to each 5' terminus of DNA, and this protein has been identified as pTP (2).

Recent developments have demonstrated that pTP functions in the initiation of DNA replication by covalently binding the first deoxynucleotide residue, which then provides a 3'-hydroxyl group for subsequent chain elongation (7, 11, 17, 20, 33). This mechanism is broadly consistent with a model originally proposed by Rekosh et al. (21). A DNA polymerase activity is associated with a protein complex consisting of pTP and a 140,000-dalton protein (11). The parental TP that is covalently attached to the template DNA is not required for the formation of the pTPdCMP complex; rather, a specific DNA sequence has to be located at the end of a linear molecule (33).

The inverted terminal repetitions (ITRs) of a number of adenovirus serotype DNAs have been sequenced and include human adenoviruses type 2 (Ad2) (1, 25) and Ad5 (27), both group C viruses; Ad12 and Ad18 (26, 32, 35; C. F. Garon, R. P. Parr, R. K. Padmanabhan, I. Vol. 44, 1982

Roninson, J. W. Garrison, and J. A. Rose, Virology, in press), both group A viruses; and Ad3 and Ad7 (10, 35), both group B viruses. The sequence of the ITRs of simian adenovirus SA7 (35), mouse adenovirus FL (34), and the avian CELO virus (P. Aleström, A. Stenlund, P. Li, and U. Pettersson, Gene, in press) have also been determined.

Challberg and Kelly (5) have described a cellfree system that is capable of initiation and completion of one round of type I replication on an exogenously added template DNA. Since this system faithfully mimics the mechanism of type I replication in vivo, it provides a convenient means for studying the sequence requirements for the initiation of DNA replication. We compared the ability of the DNA-protein complex from representative adenovirus serotype groups A to E to initiate and elongate DNA replication in heterologous adenovirus type 2-infected HeLa cell nuclear extracts, and we found that all DNA-protein complexes tested did replicate to some extent. The replication of Ad7 DNA-protein complex in an Ad2 nuclear extract was reported previously (28). As a first attempt to determine the DNA sequences required for the initiation of adenovirus DNA replication, we extended the analysis of ITR sequences to include representatives of group D (Ad9 and Ad10) and group E (Ad4) human adenoviruses, as well as another member of group A (Ad31). These combined data have enabled us to identify the DNA sequences that are common to different human adenovirus serotype DNAs which may allow DNA replication in vitro.

MATERIALS AND METHODS

Cells and virus. HeLa S3 cells were grown in suspension in Joklik modified Eagle Spinner medium (GIBCO Laboratories) containing 5% calf serum (Flow Laboratories, Inc.). A549 cells, derived from a human lung carcinoma, were obtained from Walter A. Nelson-Rees (Naval Biosciences Laboratory, Naval Supply Center, Oakland, Calif.) and grown as mono-layer cultures in Dulbecco modified Eagle medium (GIBCO) containing 10% calf serum. Ad2 was grown in both HeLa and A549 cells. Adenovirus serotypes Ad4, Ad7, Ad9, Ad10, and Ad31 were obtained from the American Type Culture Collection, plaque purified, and grown in A549 cells.

Purification of virus and DNA. Infected cells showing extensive cytopathic effect were collected by lowspeed centrifugation, frozen and thawed, and suspended in cell lysis buffer (0.5% Nonidet P-40-10 mM MgCl₂-10 mM Tris [pH 7.9]) at about 5×10^7 cells per ml. After 10 min on ice, the cell debris was removed by centrifugation, and virus was purified by two successive bandings in CsCl as described previously (18). Viral DNA was purified by digestion of purified virus by previously self-digested pronase (1 mg/ml) in the presence of 0.5% sodium dodecyl sulfate (SDS), followed by phenol and CH₃Cl-isoamyl alcohol (24:1) extractions and ethanol precipitation. DNA-protein complex was prepared as described previously (28).

Labeling and DNA sequencing of adenovirus termini. The 3' labeling of the termini of Ad4, Ad9, Ad10, and Ad31 DNA was performed as described by Challberg and Englund (3), except that the Klenow fragment of DNA polymerase I (New England Biolabs) was used instead of T4 DNA polymerase. DNA (30 to 50 µg) isolated from phenol-extracted virions was used in each labeling experiment. $[\alpha^{-32}P]dGTP$ (400 Ci/mmol, Radiochemical Centre, Amersham, England) was used for labeling. In addition, each reaction contained 0.3 mM dATP, 0.3 mM TTP, and 1 µM dGTP. Labeling was at 12°C for 3 h. After labeling, DNA samples were digested with restriction endonuclease RsaI (New England Biolabs) before being loaded on an 8% polyacrylamide gel (40:1, acrylamide-bisacrylamide) in TBE (135 mM Tris · OH-45 mM boric acid-2.5 mM disodium EDTA). Purification of labeled DNA from unincorporated nucleotide triphosphates, extraction of DNA from polyacrylamide gels, and nucleotide sequencing have been described by Maxam and Gilbert (19).

DNA replication in vitro. Nuclear extracts were prepared from Ad2-infected HeLa S3 cells as described previously (5). Reaction conditions have been described previously (28); each $50-\mu$ l reaction contained 80 ng of undigested or *Hin*dIII (New England Biolabs)-digested DNA-protein complex and was incubated for 1 h at 30°C. The reaction was stopped by the addition of EDTA to 15 mM and of pronase to 1 mg/ml and further incubated for 1 h at 37°C. The DNA products were then analyzed by either neutral or alkaline agarose gel electrophoresis.

Formation of pTP-dCMP complexes. A partially purified protein fraction containing active pTP was prepared by column chromatography through denatured DNA cellulose as described previously (33). Reaction conditions have also been described previously; each reaction, which contained $40 \ \mu g$ of DNA-protein complex, was incubated for 1 h at 30°C. The products were analyzed on 10% SDS-polyacrylamide gels as described previously (33).

RESULTS

DNA sequencing of terminal repetitions of Ad4, Ad9, Ad10, and Ad31. The sequences of the termini, determined by the chemical degradation procedures of Maxam and Gilbert (19), are shown in Fig. 1B. We infer that the first nucleotide of each ITR is dC, since preliminary experiments showed that $[\alpha^{-32}P]dGTP$ gave the highest incorporation of radioactivity into adenovirus DNA incubated with the Klenow fragment of DNA polymerase I (data not shown). Where determined, every other human adenovirus DNA had a dC residue at the 5' end.

The ITRs of Ad4 (group E), Ad9 (group D), and Ad31 (group A) are 116, 158, and 148 nucleotides long, respectively. The ITR of Ad10 is identical to that of Ad9, at least to nucleotide 135. The ITRs of the group C (Ad2 and Ad5) and group B adenoviruses (Ad3 and Ad7) have been shown to be 103 and 137 nucleotides long,



۷

		L			07	
₩P4	ŝ	(с)тстстстат	AATATACCTT	ATTTTTTG	TCTCAGTTAA	TATGCAAAT
Ad9	5	(C)TATCTATAT	AATATACCCĊ	ACAAGTAAA	CAAAGTTAA	TATGCAAAT
Ad 10	5	(C)TTCATCAAT	AATATACCCĊ	ACAAGTAAÅ	CAAAGTTAA	TATGCAAAT
Ad31	5	(C)ATCATCAAT	AATATACCTT	ACACTGGACT	TGAGCCAATA	TTAAATGA

· < · U · U · <

TGATAATGAG	GAAGCCAATA	ATTTGGATT	AATATACCTT	CATCATCAAT	5	2 PV
ATGTAAATGA	TGGTGCCAAC	ATAGATGGAA	AATATACCTT	CTCTCTATAT	5	Ad7
TTAAATGAÅ	AGTGCCAATA	ATACTGGACT	AATATACCTT	CATCATCAAT	5	Ad 12

	Ē	- m		3		
	gacatg	aacccg		ĩ		
160	AGGCGGGCtg	<ct< td=""><td></td><td>CCGGTG<u>TGAC</u></td><td></td><td></td></ct<>		CCGGTG <u>TGAC</u>		
•	GTCGCCGCGG	-	ATG <u>TGACG</u> aa <cg< th=""><th>AGCCGGCGCC</th><th>3.</th><th></th></cg<>	AGCCGGCGCC	3.	
at 3'	ACGCCTAACG	ACGGC/ 3	200000000	ccccccaccc.	TGACGT <	
tgacgttttg	<u>GTCACGACGC</u>	<u>GT</u> CACGACGĊ	GTGGGAGCCG	GAAGTGGGGC	CGTCGGGAAA	
120 GCGGGGcagg	TGCAAATGAC	TGCAAA <u>TGA</u> Ċ	TG <u>TCACC</u> CGT	GGCGGATGAG	ອດດອດດອດ	
TTCGTTAGGG	GAGAAGACGA	GAGAAGACGA	CGAAGCCGGÅ	TGTAAGTTTG	ACTAACATGG	<u>T</u> AG< 3'
100 GACAGCGGCG	GATTGGACGA	GATTGGACGA	GCAAGTTTGĊ	TTGGCGTGCT	GGGTGAATC	GCGGCTCACC
GATTGGCTGT	AGCCAACGCT	AGCCAACGCT	бесетестті	AGCTGTGGCT	GATTGGCTGT	TGGCAACGGG
80 0 <u>66650</u> 6667	TTCGGGGGGGG	TTCGGGGGGGG	ATTGACCGTA	CATTGGGTGG	GCTGTGTGGT	ecececció
ATTTGGGGGAT	TTTTAACGGT	TTTTACGGT	TGAATAGTTÅ	TGTGTAATTT	AAAGTGCGC	TTGTGACGTG
6C AGGCGTGAAA	AGCTTTTGAA	AGCTTTTGAA	GTCGCCGGAG	GTGGGCGTAG	GGTAATTTAA	GGGGTGGAGT
۹p	6 P	d 10	d31	d 1 2	42	d2

FIG. 1. Nucleotide sequence of the inverted terminal repetitions of human adenovirus DNAs. (A) The upper sequence represents a consensus sequence for the A . T-rich region of all human adenoviruses that have been sequenced. The first eight nucleotides are related to either the Ad2 or the Ad7 sequence, and all other serotype sequences can be derived from each by single base changes as shown, or by deletion of one base (triangle). The bracket emphasizes a highly conserved region within the A · T-rich region. Nucleotide sequences were previously determined for Ad5 (27, 35), Ad7 (26), Ad4 (Tokunaga et al., in press), and Ad18 (Garon et al., in press). Where more than one sequence is shown for each serotype DNA, this can be attributed to strain differences. (B) Middle and bottom show nucleotides 1 to 50 (A · T-rich region) and 60 to the 3' end of ITR (G · C-rich region), respectively, of servitypes Ad4, Ad9, Ad10, and Ad31 (determined in the present study) and of servitypes Ad12 (32, 35), Ad7 (10), and Ad2 (1, 25). The sequences outside the ITR are shown in small letters. The boxed sequences are those homologous to the sequence GGGxGGAG that is present at the origin of replication of papovavirus DNA (24).

J. VIROL.

 \mathbf{m}

Vol. 44, 1982

respectively. Two other group A adenovirus DNA termini have been sequenced; the ITRs of Ad12 and Ad18 are both 165 nucleotides long and are quite homologous. The ITR of Ad31, on the other hand, is shorter and exhibits imperfect homology with Ad12 and Ad18 in only two sections: nucleotides 1 through 68 and 120 through 148 of the Ad31 sequence. For comparison, the DNA sequences of the ITRs from Ad12, Ad7, and Ad2 are shown in Fig. 1B. It is noteworthy that whereas Ad12 and Ad18 grow the least well of the 20 virus serotypes we have propagated on A549 cells, Ad31 grows as rapidly as the viruses of subgroup B, though not as well as the groups C, D, and E (data not shown).

DNA replication in vitro. Since adenovirus DNA replication starts at the ends of the DNA, only the terminal fragments of restriction enzyme-digested adenovirus DNA-protein complex replicate specifically in vitro (13). We have used this assay to determine whether DNA-

protein complex from Ad4, Ad9, and Ad31 will replicate in vitro, using nuclear extract from Ad2-infected HeLa cells. First, the terminal restriction fragments were identified by their inability to enter agarose gels without prior digestion with pronase. DNA-protein complexes were digested with HindIII, and samples were subjected to electrophoresis in an agarose gel before and after digestion with pronase (Fig. 2A). The terminal fragments bound to TP were retarded in the gel and only migrated in the correct position after pronase digestion, thus identifying the terminal HindIII fragments as B and C of Ad4, A and D of Ad9, and C and G of Ad31. The terminal HindIII fragments of Ad2 are G and K. HindIII-digested DNA-protein complex from either Ad2, Ad4, Ad9, or Ad31 was incubated with Ad2 nuclear extract under standard reaction conditions, digested with pronase, and subjected to electrophoresis on an agarose gel (Fig. 2B). In each case, the terminal



FIG. 2. Replication of heterologous DNA-protein complexes in vitro. (A) DNA-protein complex derived from either Ad2, Ad4, Ad9, or Ad31 virions was isolated, digested with *Hind*III, and subjected to agarose gel electrophoresis either before (-) or after (+) digestion with pronase (1 mg/ml). Ad9 fragment F and Ad31 fragments J to S (Y. Sawada, T. Yamashita, and K. Fujinaga, personal communication) were not visible. The gel was stained with ethidium bromide and photographed. (B) DNA-protein complex from either Ad2, Ad4, Ad9, or Ad31 was incubated with extract from Ad2-infected HeLa cells. The products were then digested with pronase and subjected to electrophoresis through a 0.8% agarose gel. The gel was dried and autoradiographed. The positions of the terminal fragments are shown; the *Hind*IIII digest for both Ad2 and Ad31 was partial.

restriction fragments were preferentially labeled, indicating that each DNA-protein complex had the capacity to act as template for replication in vitro. The *Hin*dIII digests of Ad2 and Ad31 were both partial, and replication of the terminal and subterminal fragments is indicated (Fig. 2B). Ad7 DNA-protein complex was previously shown to replicate by using an Ad2 nuclear extract (28).

To determine whether replication in vitro resulted in genome-length DNA molecules, uncut DNA protein complex from either Ad2, Ad4, Ad7, Ad9, or Ad31 was used as template in an Ad2-infected HeLa cell nuclear extract. The products were digested with pronase and analyzed by electrophoresis on an alkaline agarose gel (Fig. 3). In each case, full-length product was observed, but the amount of product of each heterologous template was reduced relative to the homologous Ad2 template. In a number of experiments with different preparations of DNA-protein complex from each serotype, different amounts of product were observed. This has also been observed with various preparations of Ad2 complex and reflects the amount of proteolytic degradation and the concentration of the stock of each DNA-protein complex (unpub-



FIG. 3. Alkaline agarose gel electrophoresis of the replication reaction products. Undigested DNA-protein complex from either Ad2, Ad4, Ad7, Ad9, or Ad31 was incubated in a reaction mixture containing nuclear extract from Ad2-infected HeLa cells for 1 h at 37°C. The reaction products were digested with pronase and subjected to electrophoresis through a 0.7% alkaline agarose gel. The gel was dried and autoradiographed. J. VIROL.



FIG. 4. Formation of a pTP-dCMP complex with various DNA-protein complexes as template. Undigested DNA-protein complex from either Ad2, Ad4, Ad7, Ad9, or Ad31 was incubated in a reaction mixture containing $[\alpha^{-32}P]$ dCTP and a partially purified extract from Ad2-infected HeLa cell nuclei (33) for 1 h at 30°C. The reaction was stopped by the addition of EDTA to 10 mM and then was subjected to electrophoresis through a 10% SDS-polyacrylamide gel as described by Laemmli (15). Marker proteins were Ad2 virion proteins, and the molecular weights are shown. The gel was dried and autoradiographed with an intensifying screen.

lished data). However, Ad4, Ad9, and Ad31 DNA-protein complexes isolated concurrently with Ad2 complex consistently incorporated approximately 10-fold fewer picomoles of deoxynucleosides than did Ad2 DNA-protein complex, whereas Ad7 template showed 65% of the level of incorporation seen in Ad2 (28).

Initiation of DNA replication. It has been proposed (21) that the first step in the initiation of adenovirus DNA replication is the formation of complex between pTP and dCMP, the 5' terminal nucleotide of Ad2 DNA (1, 25). Such a covalent complex was found when extracts of adenovirus-infected HeLa cells were incubated with $[\alpha^{-32}P]dCTP$ and template DNA (7, 17, 33). We detected the formation of a complex between Ad2 pTP and dCMP when heterologous DNA-protein complexes were used as template. DNA-protein complexes derived from Ad2, Ad4, Ad7, Ad9, and Ad31 virions were incubated with the partially purified Ad2 nuclear extract and $[\alpha^{-32}P]dCTP$, and the reaction products were analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 4). Each DNA-protein complex facilitated the formation of an 87,000-dalton

Vol. 44, 1982

pTP-dCMP complex; however, the homologous Ad2 complex was by far the most active. Complex was not formed when either $[\alpha^{-32}P]dATP$, dGTP, or TTP was used (data not shown). The relative efficiency of pTP-dCMP complex formation for each heterologous template was reflected in the relative efficiency of replication as assayed by specific replication of the terminal restriction fragments (Fig. 2B) and by the production of full-length product (Fig. 3), suggesting that pTP-dCMP complex formation may be a rate-limiting step in the replication of viral DNA.

DISCUSSION

Challberg and Kelly (5) described the isolation of a soluble nuclear extract from adenovirusinfected HeLa cells that was capable of replication of exogenous DNA. Replication was dependent upon the template DNA having the terminal protein bound to each 5' end, since template DNA that had been digested with protease was inactive (4). Residual peptides present on pronase-digested DNA template probably inhibit the initiation reaction (33). It was therefore suggested that both the TP and DNA sequences at the origin of replication facilitated the initiation of DNA synthesis, with the TP interacting with one or more of the proteins in the soluble nuclear extract. The apparently strict requirement for the template DNA to contain covalently bound protein made it impractical to determine DNA sequence requirements at the origin by in vitro mutagenesis. However, the availability of a number of adenovirus serotypes, each having different DNA sequences at the origin of DNA replication, suggested a way to overcome this problem of analyzing DNA sequence requirements for the initiation of replication. We therefore tested the ability of heterologous DNA-protein complexes from a number of human adenoviruses, representing five serological groups, to act as template for both the initiation and the elongation of replication in nuclear extracts derived from Ad2-infected cells. All heterologous DNA-protein complexes were active for both initiation and elongation of replication in Ad2 nuclear extracts, but they were always less active than the homologous template. Template DNAs from Ad4, Ad9, and Ad31 formed a pTPdCMP complex with low efficiency, but the complex was sufficient to prime specific replication from both termini of these DNAs, as assayed by the replication of the terminal HindIII restriction enzyme fragments. Thus, we conclude that DNA sequences within the ITR that are conserved between these viruses and Ad2 may play an important role in the initiation reaction. It was not practical to reverse the experiment and use Ad2 DNA-protein complex as a template in a heterologous adenovirus nuclear extract, since all such viruses we have used grow poorly in our HeLa cells.

A problem posed by these results is the possible interaction of the DNA-bound heterologous TP with the Ad2 pTP that is present in the nuclear extracts. We do not know to what degree this will affect the initiation reaction, and thus we draw no conclusions as to the efficiency with which specific sequences within the ITRs of the different adenovirus DNAs support DNA replication in the Ad2 system. We conclude only that there is sufficient sequence homology to correctly initiate replication at some frequency. However, during the course of these experiments it became clear that the TP is not required for initiation of replication in a partially fractionated replication extract, since a plasmid DNA containing the ITR of Ad5 was active for pTPdCMP complex formation (33). The plasmid DNA was only active when the origin was located at the end of the linearized DNA. In addition, the plasmid DNA was active when it was heat denatured, again provided that the origin was located at the end of the molecule. Thus, interactions between pTP and the TP on the parental DNA are not obligatory, but they still may affect the efficiency of initiation.

The human adenovirus serotypes 1 to 38 have been subdivided into five groups on the basis of DNA homology that was derived from solution hybridization studies with virion DNA (12). DNA sequences of the ITRs from group A (Ad12 and Ad18), group B (Ad3 and Ad7), and group C (Ad2 and Ad5) have been determined, and regions of homology have been noted (26, 35). The replication studies suggest that sequences in the ITR which are important for DNA replication have been conserved; thus, we have extended the DNA sequence analysis to include the ITRs from group D (Ad9 and Ad10) and group E (Ad4) human adenoviruses and the ITR of an additional group A virus (Ad31). Recently, the sequences of the ITRs for Ad4 (O. Tokunaga, M. Shinagawa, and R. Padmanabhan, Gene, in press) and Ad18 (Garon et al., in press) were determined. Sequences of the ITRs for some of the known human adenoviruses are shown in Fig. 1B.

The ITR for each human adenovirus DNA can be divided into an $A \cdot T$ -rich region of 50 to 52 base pairs, which occurs at the termini of each DNA molecule, and a $G \cdot C$ -rich region that varies in length between 50 and 110 base pairs. Indeed, $A \cdot T$ - and $G \cdot C$ -rich sequences are also found in other eucaryotic origins of replication, most notably those in the papovaviruses (reviewed by Seif et al. [24]). Within the ITR, DNA sequences in the first 50- to 52-base pair $A \cdot T$ -rich region are the most highly conserved among all human adenoviruses (Fig. 1B). This region is also partially conserved in the adenoviruses of the lower primates (35), rodents (34), and birds (Aleström et al., in press). A consensus sequence for this region can be derived that conforms with the known sequences from all of the human adenoviruses (Fig. 1A).

Within this $A \cdot T$ -rich region, there is a perfectly conserved sequence of 10 base pairs (ATAATATACC), and all or part of this sequence is also present in the ITR of nonhuman adenoviruses. In addition, where it has been determined, all human adenoviruses contain a 5'-terminal dC, which is the base to which the TP is covalently attached. The sequence of 6 to 7 bases that lies between the terminal dC and the conserved 10-base sequence is one of two types: either related to the Ad2 sequence or related to the Ad7 sequence. Within this region, all serotype DNAs can be derived from either the Ad2 or Ad7 sequences by base substitutions or deletion (Fig. 1A). In fact, different isolates of serotypes Ad2, Ad4, and Ad7 show sequence heterogeneity within this region (Fig. 1A). At the other end of the 10-base, perfectly conserved sequence is a 9- to 11-base region that has diverged, which is followed by a region that is partially conserved among all human serotype DNAs.

The $G \cdot C$ -rich region, on the other hand, is not conserved and is also variable in length. However, two features within this region can be identified that are common to all of the human adenoviruses. The first is the occurrence of one or more copies of a sequence very similar to the sequence GGGxGGAG, which is present twice at the origin of replication of human BK virus (24) and which is also found at the simian virus 40 and polyoma virus (mouse) origins of replication (24). All human adenoviruses and the simian virus SA7 (35) have the sequence GGGCGG at least once in the ITR. As previously suggested (24), this papovavirus consensus sequence might play an important function as a recognition signal for a protein involved in DNA replication.

A second feature of the $G \cdot C$ -rich region is the occurrence of the sequence TGACG, which occurs at or near the end of all human adenovirus ITRs; related sequences occur in murine (FL) and avian (CELO) adenoviruses (35; Aleström et al., in press). It is noteworthy that this sequence is not always at the end of the ITR, as previously observed (26); it can be found within the ITR (Ad9, Ad10, and Ad31) or just outside the ITR (Ad4). A function has not been ascribed to this conserved sequence.

It is likely that these conserved features within the ITR play separate roles in the initiation of DNA replication. One possible role for the highly conserved $A \cdot T$ -rich region would be to serve as an origin for the adenovirus-specific protein priming of DNA synthesis. The complex between pTP and the 140-kilodalton protein (11) may bind to DNA sequences within this region. The conserved sequences within the $G \cdot C$ -rich region are similar to those found in other eucaryotic origins and may serve as cellular DNA polymerase recognition sites. Thus, the two structural domains may reflect two functional sites that initiate DNA replication. The recent demonstration that a linearized plasmid DNA containing the ITR at the end of the molecule can function for initiation of DNA replication in vitro (33) will enable us to define more precisely the adenovirus origin of replication by sitedirected mutagenesis.

ACKNOWLEDGMENTS

We thank John L. Hierholzer for verifying the serotypes of viruses used in this study and J. A. Rose and R. Padmanabhan for communicating sequence data before publication. We also thank Michael Mathews and Fuyuhiko Tamanoi for discussions and critical reading of the manuscript and Elizabeth Woodruff, Diane Smith, and Mark Hoppe for technical assistance.

This research was supported by grants from the National Cancer Institute (CA-13106) and from the American Cancer Society (RD-126). J.A.E. is a Special Fellow of the Leukemia Society of America.

LITERATURE CITED

- 1. Arrand, J. R., and R. J. Roberts. 1979. The nucleotide sequences at the termini of Ad2 DNA. J. Mol. Biol. 128:577-594.
- 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. U.S.A. 77:5105-5109.
- 3. Challberg, M. D., and P. T. Englund. 1980. Specific labeling of 3' termini with T4 DNA polymerase. Methods Enzymol. 65:39-43.
- 4. Challberg, M. D., and T. J. Kelly, Jr. 1979. Adenovirus DNA replication *in vitro*: origin and direction of daughter strand synthesis. J. Mol. Biol. 135:999–1012.
- Challberg, M. D., and T. J. Kelly, Jr. 1979. Adenovirus DNA replication *in vitro*. Proc. Natl. Acad. Sci. U.S.A. 76:655-659.
- Challberg, M. D., and T. J. Kelly, Jr. 1981. Processing of the adenovirus terminal protein. J. Virol. 38:272-277.
- 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.
- Coombs, D. H., A. J. Robinson, J. W. Bodnar, C. J. Jones, and G. D. Pearson. 1978. Detection of DNA-protein complexes: the adenovirus DNA-terminal protein and Hela DNA-protein complexes. Cold Spring Harbor Symp. Quant. Biol. 43:741-753.
- Desiderio, S. V., and T. J. Kelly, Jr. 1980. Structure of the linkage between the 55,000 molecular weight terminal protein. J. Mol. Biol. 145:319-337.
- 10. Dijkema, R., and B. M. M. Dekker. 1979. The inverted terminal repetition of the DNA of weakly oncogenic adenovirus type 7. Gene 8:7-15.
- 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. U.S.A. 78:6779–6783.
- 12. Green, M., J. K. Mackey, W. S. M. Wold, and P. Rigden.

1979. Thirty-one human adenovirus serotypes (Adl-Ad31) form five groups (A-E) based upon DNA genome homologies. Virology 93:481-492.

- Horwitz, M. S., and H. Ariga. 1981. Multiple rounds of adenovirus DNA replication in vitro. Proc. Natl. Acad. Sci. U.S.A. 78:1476-1477.
- Kelly, T. J., Jr., and R. L. Lechner. 1978. The structure of replicating adenovirus DNA molecules: characterization of DNA-protein complexes from infected cells. Cold Spring Harbor Symp. Quant. Biol. 43:721-728.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680-685.
- Lechner, R. L., and T. J. Kelly, Jr. 1977. The structure of replicating adenovirus 2 DNA molecules. Cell 12:1007– 1020.
- 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. U.S.A. 78:2678-2682.
- Lonberg-Holm, K., and L. Philipson. 1969. Early events of virus-cell interaction in an adenovirus system. J. Virol. 4:323-338.
- Maxam, A., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499-560.
- 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.
- Rekosh, D. M. K., W. C. Russell, A. J. D. Bellett, and A. J. Robinson. 1977. Identification of a protein linked to the ends of adenovirus DNA. Cell 11:283-295.
- Robinson, A. J., and A. J. D. Bellett. 1974. A circular DNA-protein complex from adenoviruses and its possible role in DNA replication. Cold Spring Harbor Symp. Quant. Biol. 39:523-531.
- Robinson, A. J., H. B. Younghusband, and A. J. D. Bellett. 1973. A circular DNA-protein complex from adenoviruses. Virology 56:54-69.
- Seif, I., G. Khoury, and R. Dhar. 1979. The genome of human papovavirus BKV. Cell 18:963-977.
- 25. Shinagawa, M., and R. Padmanabhan. 1979. Nucleotide sequences at the inverted terminal repetition of Ad2

DNA. Biochem. Biophys. Res. Commun. 87:671-678.

- Shinagawa, M., and R. Padmanabhan. 1980. Comparative sequence analysis of the inverted terminal repetitions from different adenoviruses. Proc. Natl. Acad. Sci. U.S.A. 77:3831-3835.
- Steenbergh, P. H., J. Maat, H. van Ormondt, and J. S. Sussenbach. 1977. The nucleotide sequence at the termini of adenovirus type 5 DNA. Nucleic Acids Res. 4:4371– 4389.
- Stillman, B. W. 1981. Adenovirus DNA replication in vitro: a protein linked to the 5' end of nascent DNA strands. J. Virol. 37:139-147.
- Stillman, B. W., and A. J. D. Bellett. 1978. Replication of DNA in adenovirus-infected cells. Cold Spring Harbor Symp. Quant. Biol. 43:729-739.
- Stillman, B. W., and A. J. D. Bellett. 1979. An adenovirus protein associated with the ends of replicating DNA molecules. Virology 93:69-79.
- Stillman, B. W., J. B. Lewis, L. T. Chow, M. B. Mathews, and J. E. Smart. 1981. Identification of the gene and mRNA for the adenovirus terminal protein precursor. Cell 23:497-508.
- Sugisaki, H., K. Sugimoto, M. Takanami, K. Shiroli, I. Saito, Y. Shimojo, Y. Sawada, Y. Uemizu, S.-I. Uesugi, and K. Fujinaga. 1980. Structure and gene organization of the transforming *Hind*III-G fragment of Ad12. Cell 20:777-786.
- Tamanoi, F., and B. W. Stillman. 1982. Function of the adenovirus terminal protein in the initiation of DNA replication. Proc. Natl. Acad. Sci. U.S.A. 79:2221-2225.
- 34. Temple, M., G. Antoine, H. Delius, S. Stahl, and E.-L. Winnacker. 1981. Replication of mouse adenovirus strain FL DNA. Virology 109:1-12.
- Tolun, A., P. Aleström, and U. Pettersson. 1979. Sequence of inverted terminal repetitions from different adenoviruses: demonstration of conserved sequences and homology between SA 7 termini and SV40 DNA. Cell 17:705– 713.
- Van Weilink, P. S., N. Naaktgeboren, and J. S. Sussenbach. 1979. Presence of protein at the termini of intracellular adenovirus type 5 DNA. Biochim. Biophys. Acta 563:89-99.
- Winnacker, E. L. 1978. Adenovirus DNA: structure and function of a novel replicon. Cell 14:761-773.