Identification of a Repeated Sequence Element Required for Efficient Encapsidation of the Adenovirus Type 5 Chromosome

PATRICK HEARING,¹ RICHARD J. SAMULSKI,^{2†} WILLIAM L. WISHART,² AND THOMAS SHENK^{2*}

*Department of Microbiology, Health Sciences Center, State University of New York, Stony Brook, New York 11794,¹ and
Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544²*

Received 29 January 1987/Accepted 6 May 1987

Adenovirus type 5 deletion mutants that lack portions of their *cis*-acting DNA encapsidation signal synthesized nearly normal levels of viral DNA and late polypeptides but failed to efficiently package the DNA into virus particles. A series of mutant viruses carrying small deletions were produced and used to identify a repeated element (AGTAAATTTGGGC and AGTAAGATTTGGCC) as a key component of the packaging signal. One copy of the repeat was sufficient to signal efficient packaging. The packaging domain could function near either end of the viral chromosome but was no longer active when moved several hundred base pairs toward the interior of the DNA molecule.

It is well established that adenoviruses carry a *cis*-acting packaging sequence located near the left end of the conventional adenovirus map. Tibbetts (14) described a series of adenovirus type 7 (Ad7) incomplete particles that contained fragments of the viral chromosome. Since all of the DNA fragments included the left end of the Ad7 genome and extended variable distances toward the right, he proposed that the viral DNA was packaged in a polar fashion from left to right. Hammariskjold and Winberg (3) demonstrated that polar packaging was signaled by a specific left-end sequence. They isolated an Ad16 variant in which 390 base pairs (bp) from the left end was translocated to the right end of the genome. Whereas examination of incomplete virus particles indicated that wild-type Ad16 packaged its DNA with a left-end polarity, the variant appeared to package from either end.

We first localized the Ad5 *cis*-acting packaging element to the left end of its chromosome during analysis of the E1A transcriptional control region (4). The E1A unit is the leftmost gene on the viral chromosome. Some mutants with deletions within the E1A upstream domain between sequence positions 194 and 358 grew poorly in 293 cells (which constitutively express a resident E1A gene; 2), and we were unable to propagate a variant lacking the entire interval. The large deletion mutant could be propagated, however, in an *in340* background. This virus carries a left-end segment (0 to 353 bp) in place of its normal right-end 103-bp repeat, mimicking the Ad16 variant (3) and presumably transferring the packaging signal to the right end of the viral DNA. The mutant which lacks the segment from positions 194 to 358 is designated *in340-2*. It grows nearly as well as wild-type virus in 293 cells, and we assume its DNA is packaged in a right-end polar fashion.

Here we report a more detailed characterization of the Ad5 packaging signal. First, we demonstrated that the domain from sequence positions 194 to 358 does indeed function to facilitate encapsidation. Next, we identified a repeated sequence motif which is critical for function of the packaging element. Finally, we determined that the packag-

ing sequence must reside near an end of the chromosome to function.

MATERIALS AND METHODS

Viruses and cells. Mutant *dl309* was selected as an Ad5 variant that contains only one *Xba*I cleavage site, at 3.8 map units (7). It is phenotypically wild type and is the parent of all other viruses used in this study. *in340* is identical to *dl309* except that it also contains the *cis*-acting Ad5 packaging sequence (left-end 353 bp) at the right terminus of its genome (4). *dl309-A5* and *dl309-B1* lack DNA segments between sequence positions 194 and 273 and positions 270 and 358, respectively, in a *dl309* background; *in340-A5* and *in340-B1* lack the same segments in an *in340* background; *in340-2* lacks the segment from positions 194 to 358 in an *in340* background. Mutants carrying small deletions within the packaging domain were generated by initially subjecting a recombinant plasmid, pE1A-WT (4), to a D-loop mutagenesis procedure (5), determining the exact endpoint of each deletion by nucleotide sequence analysis (11), and rebuilding mutations of interest into a *dl309* background by the procedure of Stow (13). Variants carrying the packaging domain in novel orientations and locations were generated by rearranging the segment from sequence positions 194 to 353 in a derivative of pE1A-WT initially lacking the domain and then rebuilding suitably modified variants into a *dl309* background (13).

Viruses were propagated and studied in 293 cells. These cells contain and express the E1A and E1B genes of Ad5 (2) and were grown in medium containing 10% calf serum.

DNA and protein analysis. High-molecular-weight DNA was isolated from infected cells (1), and viral DNA was prepared from virions purified by equilibrium density centrifugation in cesium chloride gradients; both were analyzed by blot hybridization as described by Pilder et al. (8).

The synthesis of late polypeptides was monitored by labeling 293 cells with [³⁵S]methionine (50 μCi/ml; 1,100 Ci/mmol) in medium lacking methionine and supplemented with 2% calf serum from 21 to 23 h after infection. Preparation of cellular extracts and sodium dodecyl sulfate-polyacrylamide gel electrophoresis were done as described by Sarnow et al. (12).

* Corresponding author.

† Present address: Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15217.

TABLE 1. Effect of deletion mutations within the packaging domain on virus yield in 293 cells^a

Virus	Packaging domain deletion (sequence positions)	Yield (PFU/ml)
<i>dl309</i>	None	2.0×10^9
<i>dl309-2</i>	194–358	Not viable ^b
<i>dl309-A5</i>	194–273	8.1×10^7
<i>dl309-B1</i>	270–358	7.2×10^7
<i>in340</i> ^c	None	2.0×10^9
<i>in340-2</i>	194–358	5.2×10^8
<i>in340-A5</i>	194–273	6.4×10^8
<i>in340-B1</i>	270–358	5.3×10^8

^a 293 cells were infected at a multiplicity of 5 PFU per cell, infected cultures were harvested 72 h later and subjected to three freeze-thaw cycles, and virus titers were determined on 293 cells.

^b Not viable indicates that virus was never successfully generated from a plasmid carrying the indicated mutation.

^c *in340* contains the normal left-end 353-bp segment at its right terminus.

RESULTS

Definitive identification of a *cis*-acting packaging sequence.

To conclusively demonstrate that the interval between sequence positions 194 and 358 on the Ad5 chromosome carries a *cis*-acting packaging element, the physiological defect of variants carrying mutations in this region was examined. Since it was not possible to propagate a virus lacking the entire interval (4), we analyzed two partially defective viruses that lack substantial portions of the interval. *dl309-A5* and *dl309-B1* are missing the sequence between positions 194 and 273 and positions 270 and 358, respectively. Both mutants produced yields reduced about

25-fold compared with that of their wild-type parent, *dl309*, in 293 cells (Table 1). The A5 and B1 mutations have also been built into an *in340* background (4). This variant carries the left-end packaging domain inserted at the right end of the viral chromosome. *in340-A5*, *in340-B1*, and *in340-2* (lacks both the A5 and B1 interval) generated yields in 293 cells reduced only fourfold compared with that of the wild-type virus (Table 1). The fourfold reduction in yield was not due to a residual packaging defect. Rather, it resulted from complete or partial deletion of a *cis*-acting enhancer element (element II; 6) in the three mutants. Element II functions at the left end but not when transposed to the right end of the viral chromosome (6). We conclude that the packaging defects resulting from the A5 and B1 mutations lead to approximately sixfold reductions in yield (25-fold reduction in a *dl309* background and 4-fold reduction in an *in340* background).

Late polypeptide synthesis was normal in *dl309-A5*-infected 293 cells and was reduced only twofold in *dl309-B1*-infected cells compared with wild-type virus (*dl309*)-infected cells (Fig. 1). Thus, the defect responsible for the reduced yields of these viruses likely occurs very late during the infection cycle, consistent with a defect in virion assembly.

The defect was further narrowed to the encapsidation of viral DNA into virion particles by comparing the relative amounts of mutant and wild-type DNAs present in coinfecting 293 cells to that from purified virions produced during coinfection. A mutant defective in packaging should display normal levels of intracellular DNA and reduced quantities of virion DNA. Coinfections used virus pairs in which one partner was derived from *dl309* and the other was derived from *in340*, allowing their DNAs to be distinguished easily by restriction endonuclease cleavage due to the presence of the packaging element insertion at the right end of the *in340* chromosome. Three pairs of coinfecting viruses were examined (Fig. 2). Coinfection with the two wild-type

FIG. 1. Analysis of polypeptides synthesized in 293 cells late after infection with mutant and wild-type viruses. Cultures were labeled with [³⁵S]methionine from 21 to 23 h after infection at a multiplicity of 10 PFU per cell. Cell extracts were prepared, and equal amounts of total protein were subjected to electrophoresis in a 12.5% polyacrylamide gel containing sodium dodecyl sulfate. UNINF, Uninfected.

FIG. 2. DNA accumulation and encapsidation in 293 cells late after coinfection with mutant and wild-type viruses. Total nuclear DNA was prepared at either 20 or 40 h after coinfection, and encapsidated DNA was isolated from purified virions harvested at 60 h after coinfection. Coinfections were at a multiplicity of 10 PFU per cell for each partner. DNAs were cleaved with *Hind*III and analyzed by DNA blot analysis by using an Ad5 right-end *Hind*III fragment I cloned in pBR322 as probe DNA. INTRACELL., Intracellular.

FIG. 3. Infectious yields of mutant viruses which contain deletions in the domain including the *cis*-acting packaging sequence. 293 cells were infected at a multiplicity of 5 PFU per cell, infected cultures were harvested 72 h later and subjected to three freeze-thaw cycles, and virus titers were determined on 293 cells. Yields (relative to that of the wild-type virus) are plotted against locations of deletions (segments deleted are represented by solid lines). The indicated deletions are carried in a *dl309* background, and the viruses are designated *dl309*-50, *dl309*-RH, etc. The shaded rectangles represent the location of repeats (A and A') which likely serve as core elements of the packaging signal. The sequences of the A and A' repeats are shown at the top right.

viruses (*dl309* and *in340*) generated equivalent amounts of intracellular and packaged DNA. The same result was obtained in a *dl309-in340-A5* coinfection. This makes sense because *in340-A5* carries a right-end packaging element that compensates for its left-end deletion. The results of this coinfection also support our contention (discussed above) that the fourfold reduction in *in340-A5* yield was not due to a packaging problem. The third coinfection (*dl309-A5* and *in340*) provided the key result. Although the two viral DNAs were equally represented in the intracellular pool, the proportion of *dl309-A5* DNA in the virion pool was dramatically reduced. *trans*-acting factors were equally available to the two viral DNAs within the infected cells. Therefore, we conclude that the mutation in *dl309-A5* marks a sequence required in *cis* for efficient packaging of viral DNA.

Identification of a key repeated sequence within the packaging domain. A series of mutant viruses carrying small deletions within the *cis*-acting packaging domain were produced to identify its essential sequence components. The yields of these mutant viruses in 293 cells were plotted against the locations of their deletions (Fig. 3). Two viruses carrying small deletions (*dl309-5* and *dl309-18*) exhibited growth defects nearly as severe as that observed for *dl309-A5*. Both viruses displayed a reduction in the amount of viral DNA packaged in a coinfection experiment like that for which results are shown in Fig. 2 (data not shown). Both mutations deleted a repeated element, designated A/A', whose sequence is AGTAA(G)ATTTGGG^cC. The parenthesized G is absent in the A repeat but present in the A' copy.

Several viruses carrying small deletions displayed less severe growth defects (*dl309-21*, *dl309-55*, *dl309-64*, *dl309-109*, and *dl309-141*). Most of these mutations impact on either the A or A' element, but two observations indicate that this was not the basis for their impaired growth. First, *dl309-97* lacks most of the A sequence and generated a wild-type yield. Second, we have previously shown that *dl309* viruses carrying deletion 21, 55, 64, 109, or 141, but not

97, lack key portions of enhancer element II (6). Thus, the reduced yields correlate with impairment of enhancer activity and not packaging ability. Either the A or A' sequence alone can provide the *cis*-acting packaging function defined by the deletions contained in *dl309-A5*, *dl309-5*, and *dl309-18*.

Curiously, although *dl309-B1* displayed a 25-fold reduction in yield, none of the mutants with the small deletions which subdivide the B1 domain displayed a growth defect except *dl309-21*, whose twofold impairment resulted from the impact on enhancer element II. Perhaps the signal sequence underlying the *dl309-B1* growth defect is reiterated and none of the small deletions tested removed a sufficient number of the copies to affect its function.

Packaging signal functions only near an end of the viral chromosome. We were able to propagate a virus, *in340-2*, carrying a single packaging sequence near its right end (Table 1) but were unable to generate a virus carrying a single packaging element in the 3' noncoding region of the E1A transcription unit (sequence position 1574). Thus, it seemed critical that the packaging element be located near an end of the viral chromosome.

To more carefully evaluate its positional requirement, we attempted to construct a series of viruses in which the packaging signal was moved inward from the left end in a stepwise fashion. All of these viruses lacked the packaging domain at its normal location (sequence positions 194 to 358) and contained inserts of the same segment at new locations. It was possible to move the packaging domain either 100 or 271 bp inward from the left terminus as compared with its normal position. These viruses, *dl309-2/PAC+100* and *dl309-2/PAC+271*, grew as well as the wild-type virus (Table 2). Attempts to construct viruses in which the element was displaced 655 or 863 bp inward repeatedly failed. It was also possible to move the sequence 89 bp toward the left terminus with no ill effect (*dl309-6*; Table 2), and the sequence was successfully inverted at either its normal location or with a 100-bp displacement toward the right (*dl309/PAC-INV* and *dl309-2/PAC+100-INV*, respectively; Table 2). Thus, it appears that the packaging signal can be moved and inverted, but only within a narrow range near an end of the chromosome.

TABLE 2. Effect of moving the packaging sequences on virus yield in 293 cells^a

Virus	Packaging domain alteration ^b	Yield (PFU [10 ⁹]/ml)
<i>dl309</i>	None	2.0
<i>dl309-2/PAC+100</i>	Move 100 bp right	2.2
<i>dl309-2/PAC+271</i>	Move 271 bp right	1.3
<i>dl309-2/PAC+655</i>	Move 655 bp right	Not viable ^c
<i>dl309-2/PAC+863</i>	Move 863 bp right	Not viable
<i>dl309-6</i>	Move 89 bp left	3.1
<i>dl309-INV</i>	Invert	1.6
<i>dl309-2/PAC+100-INV</i>	Move 100 bp right and invert	1.0
<i>dl309-2/PAC+271-INV</i>	Move 271 bp right and invert	Not viable

^a 293 cells were infected at a multiplicity of 5 PFU per cell, infected cultures were harvested 72 h later and subjected to three freeze-thaw cycles, and virus titers were determined on 293 cells.

^b Positions, e.g., 100 bp right, indicate the present location of the packaging segment compared with its normal location relative to the left end in wild-type virus.

^c Not viable indicates that the virus was never successfully generated from a plasmid carrying the indicated mutation.

DISCUSSION

The sequence from positions 194 to 358 on the Ad5 chromosome contains a *cis*-acting packaging element. *d*/309-A5 DNA was packaged much less efficiently than wild-type DNA in coinfecting 293 cells (Fig. 2), suggesting that the DNA segment it lacks (sequence positions 194 to 273) must include at least part of the signal. *d*/309-5 and *d*/309-18 displayed growth defects similar to that of *d*/309-A5 (Fig. 3). Their small deletions suggest that a 13- to 14-bp repeated element (designated A/A') constitutes the core signal. Results with *d*/309-97, which lacks one copy of the repeat but grew normally (Fig. 2), indicate that a single copy of the core is sufficient to signal packaging.

The A/A' repeat probably doesn't constitute the entire Ad5 packaging signal. These repeats are located entirely within the region deleted in *d*/309-A5. Several lines of evidence suggest that *d*/309-B1, which lacks a segment contiguous with the A5 deletion (sequence positions 270 to 358; Fig. 3), very likely contains a second *cis*-acting component. First, *d*/309-B1 displayed a growth defect similar to that of *d*/309-A5 (Table 1) due in part to an event which followed synthesis of late proteins (Fig. 1). Second, we have not been able to generate a virus that lacks both A5 and B1 segments (sequence positions 194 to 358) unless the domain is simultaneously inserted elsewhere on the viral chromosome (*in*340-2; Table 1). This negative result suggests that the combined B1 and A5 mutations are much more deleterious than either alone, a likely result if the two domains contain sequences which cooperate to signal efficient packaging. Experiments are in progress to better define the role of the B1 domain in packaging.

The left-end domains of Ad3, Ad7, and Ad12 contain a single copy of a sequence homologous to the Ad5 A segment. Ad3 and Ad7 contain the sequence gGTAgTTTG-GGC (10 to 13 match) beginning at position 250, and Ad12 contains the sequence tGTAAgTTTG-GGC (11 of 13 match) 101 bp from its left end. Robinson and Tibbetts (10) have characterized naturally occurring variants of Ad3 that contain deletion and substitution mutations near the left ends of their chromosomes. Surprisingly, their analysis suggests that the domain between Ad3 sequence positions 136 and 318 is not required for DNA encapsidation. Thus, at least for Ad3, the A sequence homology does not appear to play a role in packaging. Since the protein that recognizes the packaging sequence is likely coded by the virus, it might be argued that the factor and binding site have coevolved rapidly and the sequences which signal packaging in different adenovirus serotypes have drifted considerably.

The requirement that the packaging sequence be near an end of the viral chromosome (Table 2) seems to fit well with the known polarity of adenovirus packaging (3, 14). Presumably, viral DNA molecules first associate with a precapsid structure at or very near an end through the mediacy of the packaging signal. Conceivably, sequence information within the 103-bp terminal repeats also plays a role in this event. It is also possible that the terminal protein which is covalently linked to each 5' end of the viral DNA (9) is involved.

Finally, it may be no accident that the *cis*-acting A' packaging core we identified completely overlaps an adenovirus enhancer element. This enhancer element, des-

ignated element II by Hearing and Shenk (6), functions in *cis* to stimulate transcription across the entire viral chromosome. The enhancer and packaging elements could be positioned to allow a regulatory interplay between the packaging event and transcriptional enhancement. Active utilization of the enhancer element could delay the onset of packaging; the onset of packaging could in turn decrease transcription across the entire viral chromosome.

ACKNOWLEDGMENTS

This work was supported by grant MV-45 from the American Cancer Society to T.S. and Public Health Service grant AI-20412 from the National Institutes of Health and American Cancer Society grant JFRA-91 to P.H. R.J. was the recipient of a National Institutes of Health Postdoctoral Fellowship (CA 07295), W.L. is a Postdoctoral Fellow of the Damon Runyon-Walter Winchell Fund (DRG-822), and T.S. is an American Cancer Society Research Professor.

LITERATURE CITED

1. Blin, N., and D. W. Stafford. 1976. Isolation of high molecular weight DNA. *Nucleic Acids Res.* 3:2303-2308.
2. Graham, F. L., J. Smiley, W. C. Russell, and R. Nairu. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36:59-72.
3. Hammariskjold, M.-L., and G. Winberg. 1980. Encapsidation of adenovirus 16 DNA is directed by a small DNA sequence at the left end of the genome. *Cell* 20:787-795.
4. Hearing, P., and T. Shenk. 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell* 33:695-703.
5. Hearing, P., and T. Shenk. 1983. Functional analysis of the nucleotide sequence surrounding the cap site for adenovirus type 5 region E1A mRNAs. *J. Mol. Biol.* 167:809-822.
6. Hearing, P., and T. Shenk. 1986. The adenovirus type 5 E1A enhancer contains two functionally distinct domains: one is specific for E1A and the other modulates all early units in *cis*. *Cell* 45:229-236.
7. Jones, N., and T. Shenk. 1979. Isolation of Ad5 host range deletion mutants defective for transformation of rat embryo cells. *Cell* 17:683-689.
8. Pilder, S., J. Logan, and T. Shenk. 1984. Deletion of the gene encoding the adenovirus 5 early region 1B 21,000-molecular-weight polypeptide leads to degradation of viral and host cell DNA. *J. Virol.* 52:664-671.
9. Rekosh, D. M. K., W. C. Russell, A. J. D. Bellet, and A. J. Robinson. 1977. Identification of a protein linked to the ends of adenovirus DNA. *Cell* 11:283-295.
10. Robinson, C. C., and C. Tibbetts. 1984. Polar encapsidation of adenovirus DNA: evolutionary variants reveal dispensable sequences near the left ends of Ad3 genomes. *Virology* 137:276-286.
11. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.
12. Sarnow, P., Y. S. Ho, J. Williams, and A. J. Levine. 1982. Adenovirus E1b-58kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54kd cellular protein in transformed cells. *Cell* 28:387-394.
13. Stow, N. D. 1981. Cloning of a DNA fragment from the left-hand terminus of the adenovirus type 2 genome and its use in site-directed mutagenesis. *J. Virol.* 37:171-180.
14. Tibbetts, C. 1977. Viral DNA sequences from incomplete particles of human adenovirus type 7. *Cell* 17:243-249.