Cell, Vol. 27, 121–131, November 1981 (Part 2), Copyright © 1981 by MIT

The 2.2 kb E1b mRNA of Human Ad12 and Ad5 Codes for Two Tumor Antigens Starting at Different AUG Triplets

Johannes L. Bos, Laurien J. Polder, René Bernards*, Peter I. Schrier, Peter J. van den Elsen, Alex J. van der Eb and Hans van Ormondt
Department of Medical Biochemistry
Sylvius Laboratories
University of Leiden
Wassenaarseweg 72
2333 AL Leiden
The Netherlands
*and Laboratory of Gene Structure and Expression
National Institute for Medical Research
Mill Hill, London NW17 1AA
England

Summary

By nucleotide sequence analysis and S1 nuclease mapping we have determined the structural organization of early region E1b of Ad12. We have also revised the nucleotide sequence of the E1b region of Ad5. Both regions have an identical structural organization and show considerable homology at the nucleotide level. The major tumor antigens (Ad12, 19 and 54 kilodaltons [kd]); Ad5, 21 and 55 kd) are encoded in two overlapping reading frames. A single mRNA of 2.2 kilobases codes for both these proteins, depending on which AUG triplet serves as the start codon: the 19–21 kd protein initiates at the 5′-proximal AUG; the 54–55 kd protein initiates at the second AUG in another reading frame. Peptide mapping shows that the small and large tumor antigens do not share common tryptic peptides, in accordance with the nucleic acid sequence data. In addition, the 19–21 kd protein can also by synthesized from a one kilobase mRNA. Finally, the gene for the Ad12 analog of protein IX is characterized.

Introduction

Human adenoviruses can transform cultured cells of certain rodents, as can their naked DNAs. Cells thus transformed can induce tumors both in syngenic animals and in nude mice with varying degrees of efficiency. Cells transformed by adenovirus 2 (Ad2) or the closely related Ad5 induce tumors rarely, and only after a long latency time, whereas Ad12-transformed cells induce tumors at high frequency within a relatively short period (for review, see Flint, 1980).

The genes that cause transformation are located at the left end of the linear genome in all human adenoviruses investigated thus far (Graham et al., 1974; Shiroki et al., 1977, 1979; Sekikawa et al., 1978; Dijkema et al., 1979; Houweling et al., 1980; van der Eb et al., 1980). They coincide with a block of genes (region E1) that, in lytic infection, is mainly expressed in the early stages of that process. RNA mapping studies show that this region is divided into two transcription units, E1a and E1b (Flint, 1980). Most of the research on this transforming region has been directed at Ad2 and Ad5: the nucleotide sequence has been determined (van Ormondt et al., 1980b), the transcripts have been mapped (Berk and Sharp, 1978; Chow et al., 1979, 1980; Kitchingman and Westphal, 1980) and the main gene products have been determined (Harter and Lewis, 1978; Green et al., 1979; Halbert et al., 1979; Schrier et al., 1979; Ross et al., 1980; Jochemsen et al., 1981; Lupker et al., 1981). The combined evidence shows that the E1a region encodes at least several proteins that are synthesized from overlapping RNAs differing by alternative splicing. A similar situation has been suggested for the E1b region, where two overlapping mRNAs with identical 5′ and 3′ termini but different splicing patterns encode the major tumor antigens (Perricaudet et al., 1980; van Ormondt et al., 1980b).

The structural organization of the transforming region of Ad12 closely resembles that of Ad2 and Ad5. Sawada and Fujinaga (1980) established that the Ad12 transcription map is similar to that of Ad2 with only minor differences. Jochemsen et al. (1980; Jochemsen, 1979; van der Eb et al., 1980) have investigated the tumor antigens encoded in this region both in vivo and in vitro and found striking similarities between the Ad5 and Ad12 tumor antigens. The similarities were confirmed by the nucleotide sequence of the E1a region of Ad12 as determined by Sugisaki et al. (1980), which shows extensive homology with the corresponding region of Ad5 (van Ormondt et al., 1980a). It was, therefore, surprising that the nucleotide sequence of the left part of the E1b region of Ad12 as determined by Sugisaki et al. (1980) revealed a different structural organization with respect to that of Ad5. In spite of the sequence homology, the nonsense codon distribution reported for Ad12 does not permit the synthesis of E1b polypeptides comparable to those of Ad5. This is at variance with the results of Jochemsen et al. (1980; Jochemsen, 1979; van der Eb et al., 1980), who have demonstrated the synthesis of Ad12 E1b tumor antigens of 19 and 60 kilodaltons (kd) both in vivo and in vitro.

To elucidate this discrepancy we decided to reinvestigate the nucleotide sequence of this part of the E1b region of Ad12 and determine the nucleotide sequence of the remaining part of this region. In contrast to Sugisaki et al. (1980), we found the Ad12 E1b region to contain two different, overlapping reading frames that could encode the two detected E1b tumor antigens. Nevertheless, this structural organization was still different from the E1b region of Ad5, which has been reported to have a single open reading frame (van Ormondt et al., 1980b). In the present

paper we report that a single bp has to be deleted in the published Ad5 sequence, which reconciles the structural organizations of the E1b regions of the two serotypes. In our laboratory we have recently established the same structural organization for Ad7 (R. Dijkema, B. Dekker and H. van Ormondt, in preparation).

Results

Molecular Cloning of the Transforming Region of Ad12

After removal of the terminal protein by mild exonuclease III and nuclease S1 treatment and ligation of Eco RI linkers onto the termini of the viral DNA, the Eco RI fragments of Ad12 were cloned in the Eco RI site of the E. coli plasmid pAT153. Four clones containing the left-hand 5.5 kilobase (kb) Eco RI C fragment (map position 0 to 16%) were isolated and characterized. The transforming capacity of these cloned fragments was verified by transforming baby rat kidney cells with the hybrid plasmids. All four were able to transform cells with an efficiency comparable to that of Ad12 viral DNA. Furthermore, cells transformed by these clones were able to induce tumors after injection into nude mice (R. Bernards, A. Houweling, J. Hertoghs, A. Bos and A. van der Eb, in preparation).

Sequence Analysis

Figure 1 shows a restriction map of the E1b region of Ad12 and the strategy for nucleotide sequence analysis. The resulting sequence is presented in Figure 2. As shown in Figure 3A, the E1b region contains three tracts for protein synthesis. The first one runs from ATG 1541 to TAA 2030, allowing the synthesis of a polypeptide of 19 kd. From ATG 1846, the first ATG downstream from ATG 1541, a second reading frame—partially overlapping the first one—is open, coding for a polypeptide of 54 kd. Finally, a third reading frame maps between ATG 3373 and TAA 3805 and codes for the Ad12 analog of protein IX (see Discussion).

The sequence of the E1b region of Ad12 determined by us varies from that determined by Sugisaki et al. (1980), who did not find coding space for the 19 and 54 kd proteins. This difference lies in the absence of two residues in their sequence (residues 1574 and 2073). Although we cannot exclude strain differences, we feel this is an unlikely explanation of our results, the more so since one of the discrepancies with the previously published sequence can be explained easily as an artefact of the method used (see Figure 4). The difference between the organization of the Ad12 E1b region and that previously reported for the Ad5 E1b (van Ormondt et al., 1980b) led us to reinspect the sequence data from which the Ad5 sequence was deduced. We found a probable error of interpretation, and a new sequence gel (not shown) of the criticized region confirmed our suspicions. As a result, we had to delete the CG pair at position 2108 of the previously published sequence. The consequence of this deletion is that the E1b region contains, as in Ad12, two partially overlapping open frames, one stretching from ATG 1714 to TGA 2242, coding for a polypeptide of 21 kd and the other from ATG 2019 to TGA 3507, coding for a polypeptide of 55 kd (Figures 2 and 3B). As is evident from Figure 2, the Ad12 and Ad5 sequences show extensive homology, except for a region of approximately 300 bp.

RNA Mapping

The E1b region of Ad5 encodes two major "early" mRNAs and the quasi-late mRNA for protein IX, which have been mapped at the nucleotide level (Figure 3B). The two "early" mRNAs have identical 5′ and 3′ termini but differ by alternative splicing. In Ad12, E1b mRNAs of 0.45, 1.0 and 2.2 kb have been mapped by S1 nuclease analysis with relatively large DNA fragments (Sawada and Fujinaga, 1980); they show an organization similar to those in Ad5 (Figure 3). The extensive sequence homology between the Ad12 and Ad5 sequences (Figure 2), together with S1 mapping data obtained with short DNA probes (shown for 5′ terminus in Figure 5), enabled us to predict the posi-

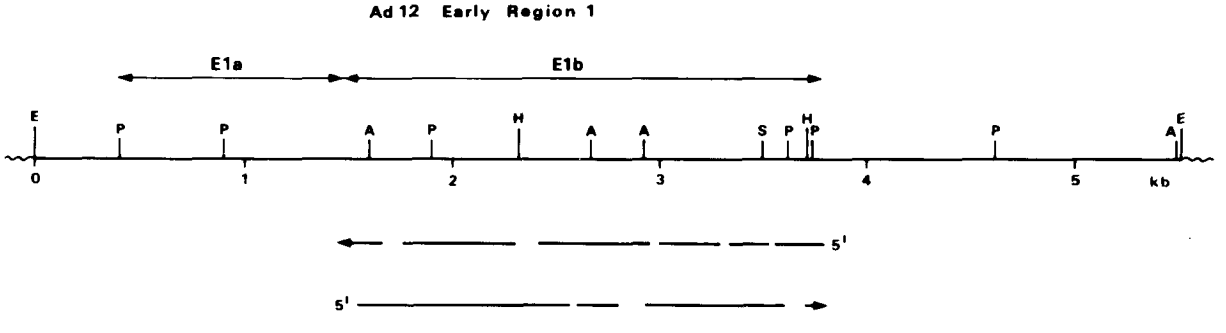

Figure 1. Restriction Map of the Cloned Eco RI C Fragment of Ad12 DNA Including the Region That Has Been Sequenced

A: Acc I. E: Eco RI. H: Hind III. P: Pvu II. S: Sal I. The bottom lines indicate which portions of both DNA strands have been sequenced.

tions of the termini and splice sites of the Ad12 E1b mRNAs at the nucleotide level. Thus the 5′ splice site of the 2.2 kb mRNA was localized at position 3295, and the common 3′ splice site of the 2.2 and 1.0 kb mRNAs at position 3368 (not shown). Only the 5′ splice site of the 1.0 kb mRNA could not be assigned by homology, since this site is located in a region where the Ad12 and Ad5 sequences diverge. S1 nuclease analysis (not shown), however, enabled us to localize this splice site around position 2050, presumably at position 2046 (surrounding sequence AG-GT), in accordance with the GU-AG rule of Breathnach et al. (1978). The common 5′ ends of the 1.0 and 2.2 kb mRNAs of Ad5 have been mapped at position 1702 or 1704 (Maat and van Ormondt, 1979; Baker and Ziff, 1980, 1981) and for Ad12 around position 1530 (Sawada and Fujinaga, 1980; see also Figure 5). This implies that the 2.2 kb mRNAs contain the coding information for both the 19-21 kd and the 54-55 kd proteins (see Figure 3), but that the 54-55 kd proteins can only be synthesized by the 2.2 kb mRNAs when the second AUG triplet is used.

To investigate whether hitherto undetected mRNAs with their 5′ ends between the two 5′-proximal AUG triplets serve as messengers for the 54-55 kd proteins, we have performed detailed S1 nuclease analyses of these termini. For Ad12 we have hybridized RNA from HT3 cells (hamster tumor cells induced by the Ad12 Eco RI C fragment [0–16%]) to a 5′-end-labeled Hpa I–Aos I fragment (positions 996 to 2089). S1 nuclease treatment yields two protected segments with lengths of 560 and 950 nucleotides (Figure 5A). The 560 nucleotide band represents the 5′-terminal part of the 2.2 kb mRNA (the 1.0 kb mRNA cannot protect the 5′-end label of the Aos I site) and the 950 nucleotide band is probably generated by protection with an E1a-E1b cotranscript (Sawada and Fujinaga, 1980), since its 5′ end maps around the 3′ splice site of the E1a mRNAs. No bands are visible, even after prolonged autoradiography, which could represent mRNAs with their 5′ end encoded between the first (1541) and second (1846) AUG triplets.

To exclude the possibility that such an mRNA is only slightly shorter than the 2.2 kb mRNA, and thus indistinguishable from the 560 nucleotide band in the above experiment, we have hybridized HT3 RNA to a 5′-end-labeled Hph I-Hinf I fragment (positions 1425 to 1666), which was strand-separated. After S1 nuclease treatment, the protected segments were run alongside a partial sequence ladder (Figure 5B) for the same fragment. A cluster of bands was detected around position 1530. Such a microheterogeneity in length of a protected segment is frequently found after S1 nuclease treatment, especially when a capped 5′ end is involved (Weaver and Weissmann, 1979). None of the bands, however, could represent an mRNA with its 5′ end encoded downstream from ATG 1546. These results further show that, in agreement with the mapping data of Sawada and Fujinaga (1980), the 2.2 kb mRNA starts around position 1530, presumably at the A residue at position 1527. For Ad5 we have hybridized either RNA from cells lytically infected with Ad5 (isolated 6 hr after infection) or RNA isolated from Ad5 Xho I C (0 to 15.5%) transformed cells to a 5′-end-labeled Hpa I–Mbo I fragment (positions 1572 to 2426). S1 nuclease treatment yields a 700 nucleotide segment generated by protection by the 2.2 kb mRNA (Figure 5C). No bands are apparent with an intensity of more than 1% of the main band, which could represent possible mRNAs with 5′ ends encoded between the two 5′-proximal ATG triplets. Furthermore, we were not able to detect, either for Ad12 or Ad5, messengers with 5′ termini encoded to the right of position 1846 (Ad12) or 2019 (Ad5) that could direct the synthesis of polypeptides initiating at an AUG downstream from those postulated above to initiate the 54-55 kd polypeptides (not shown). Since the large E1b tumor antigen is a major gene product (Jochemsen, 1979; Schrier et al., 1979; Jochemsen et al., 1980), we conclude from these results that the 2.2 kb mRNA directs the synthesis of the 54-55 kd protein.

Tryptic Peptide Maps of the Ad5 E1b Tumor Antigens

To determine whether the small and large E1b T antigens are related, they were subjected to peptide analysis. Figure 6 shows the fingerprints of the major Ad5 E1b T antigens (19 and 65 kd). None of the methionine-containing peptides are present in both proteins. A similar result was obtained with the Ad12 E1b T antigens (19 and 60 kd; H. Jochemsen, G. Daniels, J. Hertoghs, P. Schrier, P. van den Elsen and A. van der Eb, submitted). These results are in perfect agreement with the data on the primary structures of the 19-21 kd and 54-55 kd proteins predicted from the nucleic acid sequences (Figure 7). Since these proteins are encoded in different reading frames, no homology is expected.

Discussion

Coding Capacity

By nucleotide sequence analysis and RNA mapping we have investigated the structural organization of the E1b region of Ad12. Furthermore, we have revised the structural organization of the E1b region of Ad5. Both regions are very homologous and consist of two overlapping reading frames allowing the synthesis of a 19-21 kd and of a 54-55 kd protein (see Figure 3). Based on analysis of the viral gene products synthesized both in lytically infected and in transformed cells, it had previously been concluded that the E1b region of both Ad5 and Ad12 encodes two major tumor antigens of 19 and 60-65 kd (Harter and Lewis, 1978; Green et al., 1979; Halbert et al., 1979; Jo-

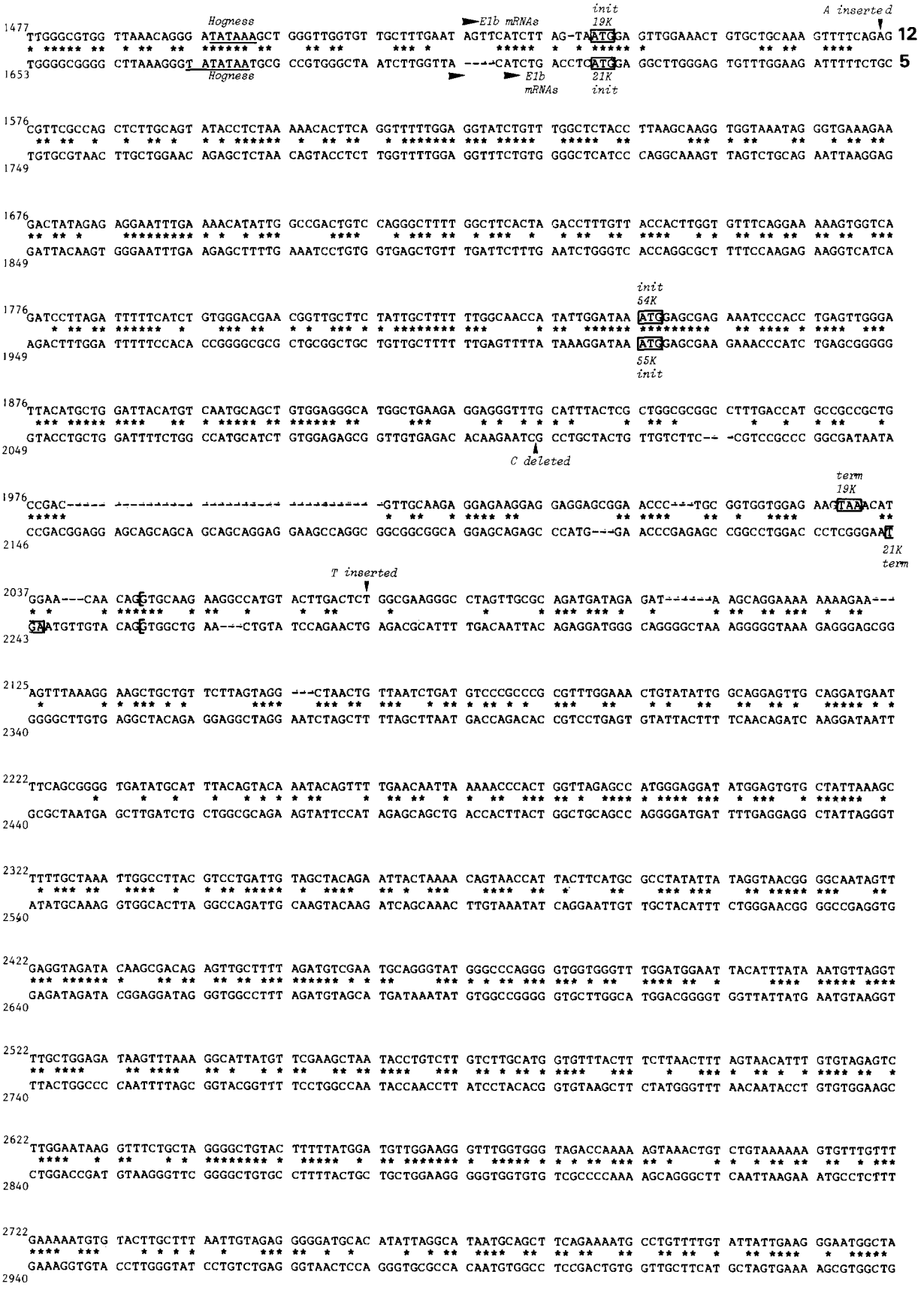
Hogness
E1b mRNAs
init
19K
A inserted
12
5
Hogness
E1b
mRNAs
21K
init
init
54K
55K
init
C deleted
term
19K
21K
term
T inserted

Figure 2. Nucleotide Sequence of the L-Strand of the E1b Regions of Ad12 and Ad5 DNA

The strands are aligned so as to give maximum homology. Asterisks: positions of identical nucleotides. In the Ad5 sequence, the 5′ ends of the major E1b mRNAs (Baker and Ziff, 1980, 1981), the 5′ end and the polyadenylation site of the mRNA for protein IX (Alestrom et al., 1980) and the splice sites of the intervening sequences (brackets; Perricaudet et al., 1980) are indicated. Initiation and termination codons are boxed, and postulated Goldberg-Hogness boxes and polyadenylation signals are underlined. The corresponding positions in Ad12 DNA are indicated (this paper).

chemsen, 1979; Schrier et al., 1979; Esche et al., 1980; Jochemsen et al., 1980, Ross et al., 1980; Spector et al., 1978; H. Jochemsen et al., submitted). Moreover, Anderson and Lewis (1980) have partially determined the N-terminal amino acids of the small E1b T antigen of Ad2 (Ad2 is closely related to Ad5) and found these to agree with a polypeptide initiating at AUG-1714. Unfortunately, the amino terminus of the 65 kd protein so far has been refractory to sequencing, but the lack of common tryptic peptides

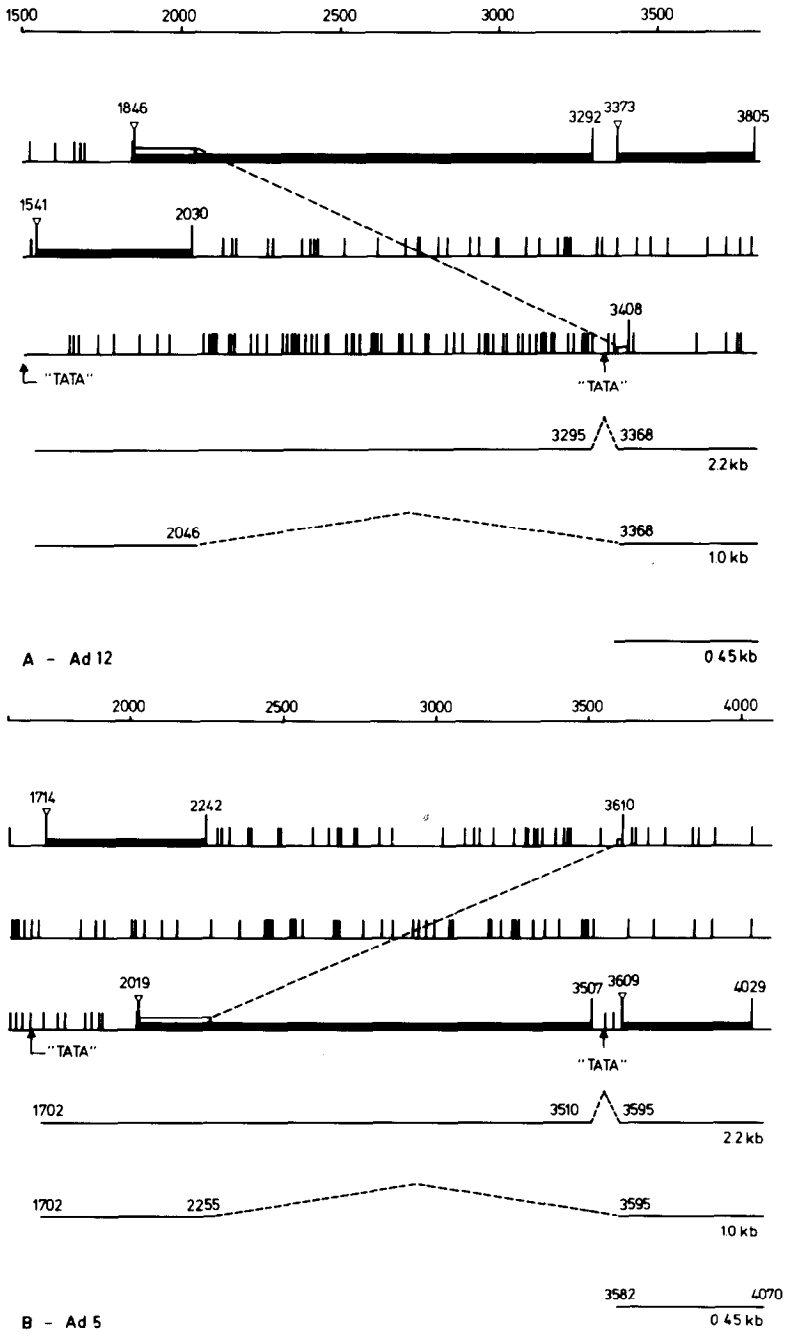

Figure 3. Schematic Representation of the Structural Organization of the E1b Regions of Ad12 and Ad5

In both A (Ad12) and B (Ad5) the three top lines represent the three possible reading frames. Tick marks: nonsense codons. Heavy lines: regions open for protein synthesis, which may be connected by RNA splicing (broken lines). Inverted triangle: the first AUG of a long open reading frame. Tall tick mark: stop codon terminating such a reading frame. TATA: a Goldberg-Hogness box. The three bottom lines represent the map positions of the major E1b mRNAs. Continuous lines represent the segments of RNA, which are linked by splicing (broken lines).

with the 19 kd protein (Figure 6) shows that these two proteins do not use the same reading frame as proposed by van Ormondt et al., 1980b). Most likely, the 60-65 kd tumor antigens are initiated at the first AUG triplet of the long open frame (AUG-1846 of Ad12 and AUG-2019 of Ad5). From these initiation codons, proteins can be synthesized with molecular weights that are in reasonable agreement with those observed experimentally. We cannot exclude, however, that an AUG triplet further downstream serves as the actual initiation codon.

Finally, in Ad12, a reading frame for a 14 kd protein between ATG-3373 and TAA 3805 is present. By analogy with the Ad5 sequence (Figure 7C) we conclude that this protein is the Ad12 analog for protein IX.

The 2.2 kb E1b mRNA Directs the Synthesis of Two Proteins

At first sight, both in Ad12 and Ad5 a corresponding mRNA can be indicated for each reading frame. Thus the 0.45 kb mRNA should encode protein IX (14 kd), the 1.0 kb mRNA the 19-21 kd protein and the 2.2 kb mRNA the 54-55 kd protein. However, when protein

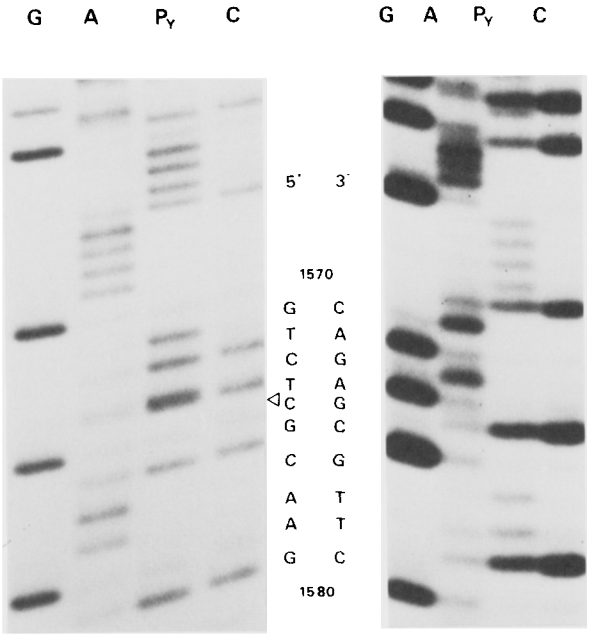

Figure 4. Autoradiograms of DNA Sequencing Gels

The gels reproduce a part of the nucleotide sequence leftwards from the Acc I site at position 1595. The sequence was obtained with use of a fragment with either a 5′-end-labeled Acc I site (left panel) or a 3′-end-labeled Acc I site (right panel). Arrowhead: a CT compression at position 1574, which could easily be misinterpreted as a single C band.

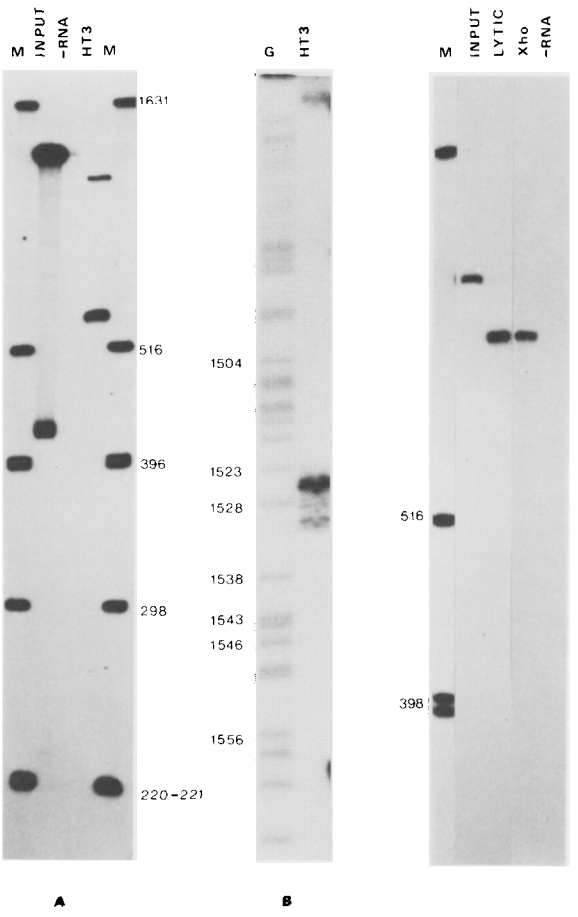

Figure 5. S1 Nuclease Mapping

5′-end-labeled DNA fragments were hybridized to RNA, treated with S1 nuclease and separated on either a 3.5% (A and C) or an 8% (B) polyacrylamide–7 M urea gel.

(A) Ad12 Hpa II–Aos I fragment (996 to 2089) hybridized to RNA isolated from HT3 cells (lane HT3) and in the absence of RNA (lane −RNA). In the other lanes the untreated DNA probe (lane INPUT) and a Hinf digest of pAT153 (lanes M; size markers) were applied.

(B) The r-strand of the Ad12 Hph I–Hinf I fragment (1425 to 1666) hybridized to RNA isolated from HT3 cells (lane HT3) and the G sequence lane for the r-strand of the same Hph I–Hinf I fragment (lane G).

(C) Ad5 Hpa I–Mbo I fragment (1572 to 2426) labeled at the Mbo I site and hybridized to RNA isolated from cells lytically infected with Ad5 (6 hr after infection; lane LYTIC), or to RNA from Ad5 Xho I C transformed cells (lane Xho), or in the absence of RNA (lane −RNA). The other lanes contain the untreated DNA probe (lane INPUT) and a Hinf digest of pAT153 as size markers (lane M).

synthesis starts at the first AUG of the 2.2 kb mRNA, the 19-21 kd protein is synthesized, just as on the 1.0 kb mRNA. In order for the 54-55 kd protein to be synthesized, protein synthesis must initiate at the second AUG of the 2.2 kb mRNA. Several arguments favor the hypothesis that the 2.2 kb mRNA can direct the synthesis of both the 19-21 kd and the 54-55 kd proteins. First, the Ad12 54 kd and the Ad5 55 kd proteins are major E1b gene products (Jochemsen, 1979; Schrier et al., 1979; Jochemsen et al., 1980; H. Jochemsen et al., submitted), and S1 nuclease analysis has shown that there is no other messenger available in appreciable amounts that could direct the synthesis of the 54-55 kd protein from the first AUG of the long open frame or, potentially, of a shorter peptide from an AUG triplet further downstream (Berk and Sharp, 1978; Chow et al., 1979, 1980; Kitchingman and Westphal, 1980; Sawada and Fujinaga, 1980; this paper; P. J. van den Elsen and A. van der Eb, in preparation). Second, in cells transformed by the Ad12 Eco RI C fragment (0–16%) and in cells transformed by the Ad5 XhoI C fragment (0–16%) the 2.2 kb mRNA is the only detectable E1b messenger (Sawada and Fujinaga, 1980; our unpublished observations; P. J. van den Elsen and A. van der Eb, in preparation) whereas both the 19-21 kd protein and the 54-55 kd protein are present in large amounts (Schrier et al., 1979; Jochemsen et al., 1980).

Finally, Esche et al. (1980) fractionated the E1b mRNAs by electrophoresis through agarose gels containing methylmercuric hydroxide and translated the separated material in vitro. In our laboratory, Lupker et al. (1981) performed similar experiments with E1b mRNAs separated by sucrose gradient centrifugation. The conclusion drawn from both investigations was that the 15-19 kd protein is specified by the 1.0 kb (13S) mRNA, as well as by longer species (larger than 2 kb [Esche] or 23S [Lupker]). Lupker et al. (1981) also showed that the proteins synthesized on both size classes of mRNA produced highly related if not

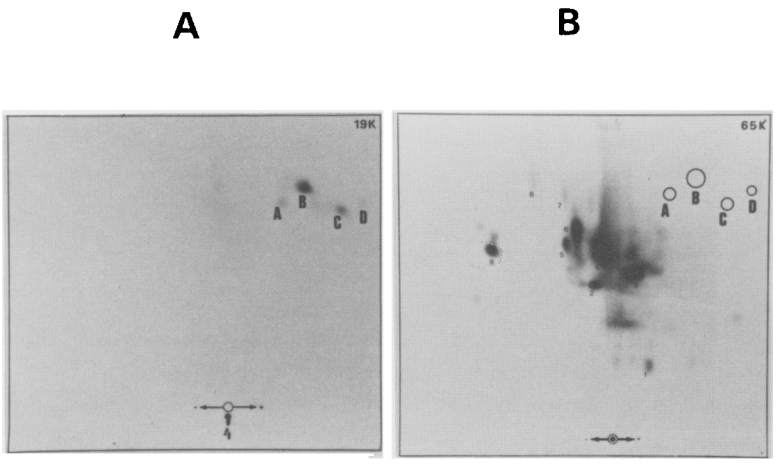

Figure 6. Autoradiograms of Tryptic Fingerprints

Fingerprints of ^{35}S-methionine-labeled Ad5 small (A) and large (B) E1b T antigens. The peptide spots marked A, B, C and D present in (A) are lacking in (B).

identical tryptic fingerprints. Hence we must conclude that the 2.2 kb mRNA is the only E1b messenger with sufficient coding space for a 54-55 kd protein. In order to synthesize the 54-55 kd product, ribosomes must be able to bypass the 5′-proximal AUG triplet. Possibly this is facilitated by the short distance between the 5′ cap and the first AUG (10–16 nucleotides). It is interesting that in vitro the 54-55 kd protein is synthesized in low amounts relative to the 19-21 kd protein, whereas in infected and transformed cells it is present in large amounts. Apparently, the cell-free systems for protein synthesis are deficient for a factor that is required for ribosomes to select the second AUG initiation site. It is tempting to speculate that this factor is of viral origin. Although the 2.2 kb mRNA differs from the vast majority of the mRNAs enumerated by Kozak (1981) in that it provides the ribosomes with two alternative initiation points, it is not the only such messenger in existence. The late 16S mRNA of SV40 also has (non-overlapping) coding space for two polypeptides that both are synthesized in the infected cell (see Jay et al., 1981). To explain the translation of this SV40 mRNA, Kozak (1981) envisages a relaxed scanning mechanism that would allow the 40S ribosomal subunit to slide by the 5′-proximal AUG and to start protein synthesis at a following initiation codon. The possibility exists that the same mechanism also holds for the 1.0 kb mRNA encoded by the adenoviral E1b region. In that case, a 8.6 kd protein would be synthesized (see Figure 3). Whether this protein is equivalent to the 10 kd E1b protein detected by Shiroki et al. (1980) is unknown.

Proteins Encoded by the E1b Region

The elucidation of the primary structure and the transcription map of the E1b regions of Ad12 and Ad5 has enabled us to predict the amino acid sequences of the proteins encoded in this region, with the assumption that the N terminus of the 54-55 kd protein is encoded by the second AUG of the 2.2 kb mRNA (see above). Comparison of these amino acid sequences in the two serotypes shows striking homologies (Figure 7). Only the carboxy-termini of the 19 and 21 kd proteins, the N-terminal parts of the 54 and 55 kd proteins and the carboxy-terminal portion of protein IX diverge. The strong homologies of the tumor antigens are consistent with transformation experiments, where deletion of the E1b region in one of the serotypes can be complemented by the E1b region of the other serotype (R. Bernards et al., in preparation). The only striking functional difference is the difference in tumorigenic potential of the Ad5- and Ad12-transformed cells in nude mice. Ad5-transformed BRK cells induce tumors rarely and only after a long latency time, whereas Ad12-transformed cells induced tumors at high frequency and within a relatively short period. The gene product that most likely is involved in tumorigenesis is the 54 kd protein, since cells transformed by the left-terminal 7% of the Ad12, which consequently lack the 54 kd tumor antigen, are incapable of inducing tumors in nude mice (Jochemsen, 1979). Whether other gene products are involved in tumorigenicity is still unknown. We are currently testing this by analyzing the tumorigenic potential of cells transformed with Ad5-Ad12 hybrid E1 regions.

Regulatory Aspects of the E1b Organization

The expression of the adenovirus genome is subject to temporal regulation that also pertains to such a small subregion as E1b. Its 2.2 kb mRNA is already apparent at the early stage of infection and remains more or less constant during the infection process (Spector et al., 1978; our unpublished observations). However, the amount of the 1.0 kb mRNA that is below our detection limit early during infection shows a sharp rise at late times so as to exceed that of the 2.2 kb species. With our present understanding of the organization of region E1b, the early 2.2 kb mRNA serves as the template for both the 54-55 and 19-21 kd proteins, whereas at late times, when apparently the infection process needs larger amounts of the 19-21 kd product, an additional mRNA (1.0 kb) is synthesized to fulfill that requirement. The details of this regulation await further investigation.

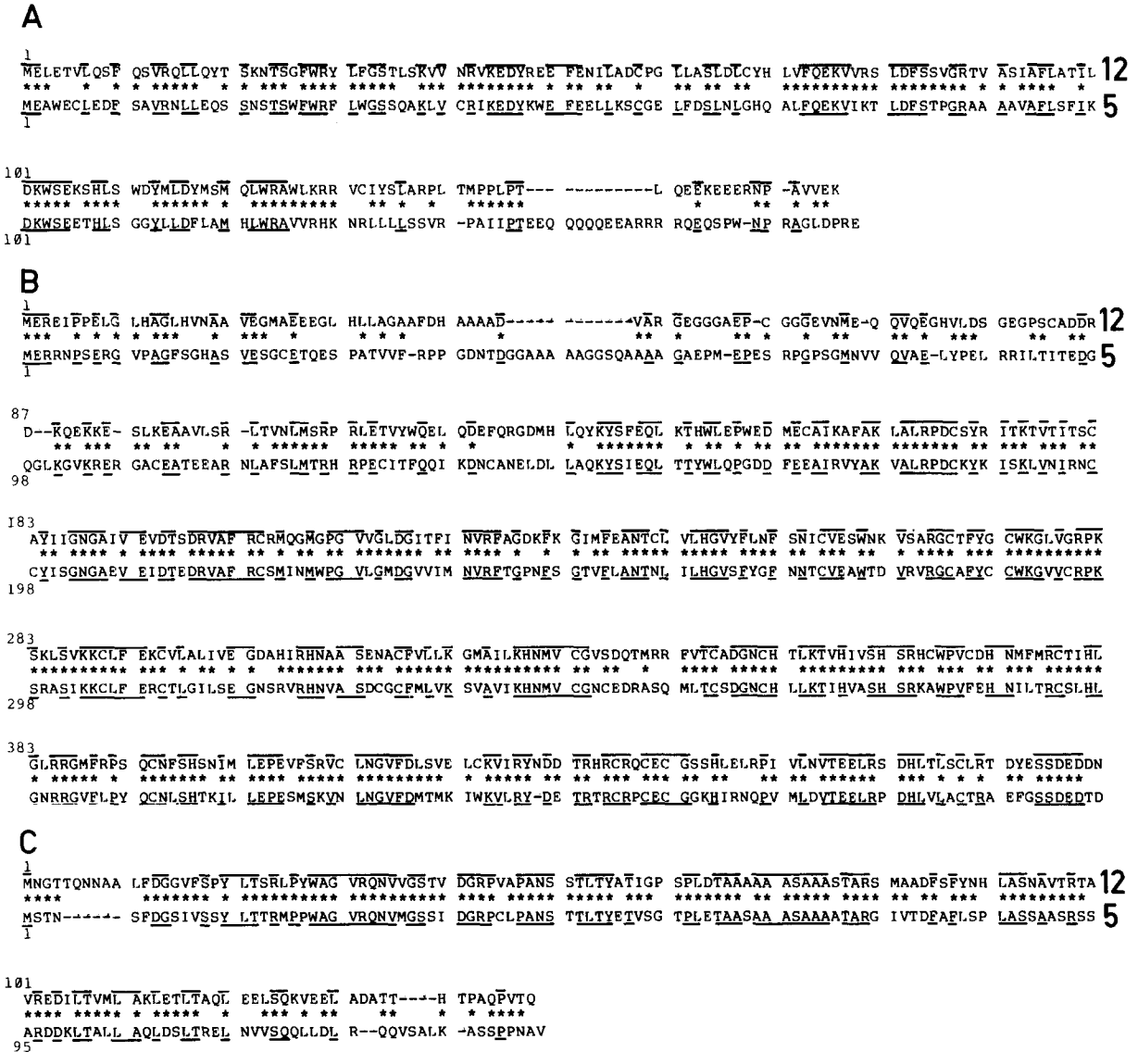

Figure 7. The Predicted Primary Structures of the E1b Proteins of Ad12 and Ad5

Upper sequence: Ad12. Lower sequence: Ad5. The sequences are aligned so as to give maximum homology between the two serotypes. Lines above and below the sequences indicate the positions where identical amino acids are found, and asterisks show the presence of amino acids of similar character (acidic, basic, hydrophobic or polar [uncharged]).
(A) Ad12 19 kd and Ad5 21 kd proteins. (B) Ad12 54 kd and Ad5 55 kd proteins. (C) Protein IX of Ad12 and Ad5.

Experimental Procedures

Virus, Viral DNA and Viral RNA

Adenovirus type 12 (Huie strain) was propagated on monolayers of KB cells. Virus purification and isolation of viral DNA has been described by van der Eb et al. (1969). RNA was extracted either from Ad12 RI C-HT3 (hamster tumor) cells (Jochemsen, 1979), from cells transformed with the Ad5 Xho I C fragment (P. J. van den Elsen and A. van der Eb, in preparation) or from cells lytically infected with Ad5 (6 hr after infection) according to the procedure of Flint et al. (1976). Transformation of baby rat kidney cells with Ad12 DNA fragments was performed as described by van der Eb and Graham (1980).

Cloning of Adenovirus 12 Left-Terminal DNA Fragment

Step 1. Preparation of Blunt-Ended Viral DNA

To remove the residual terminal protein from the viral DNA, we incubated 50 μg of viral DNA with 300 units/ml of exonuclease III in 7 mM Tris-HCl (pH 7.8), 7 mM $MgCl_2$, 10 mM 2-mercaptoethanol and 70 mM NaCl for 6 min at 15°C in a vol of 100 μl. After heat inactivation for 10 min at 70°C, the protruding 5′ ends were trimmed with 200 units/ml of S1 nuclease for 90 min at 15°C. Subsequent to phenol extraction and ethanol precipitations, the viral DNA (500 μg/ml) was repaired by incubation for 30 min at 37°C with 50 units/ml of DNA polymerase (Klenow fragment) in the presence of 40 μM of the four dNTPs. After the reaction the DNA was extracted with phenol and precipitated with ethanol.

Step 2. Ligation of Eco RI Linkers to Viral DNA

Sixty pmoles of Eco RI linkers were 5′-end-labeled by incubation with 100 μCi of γ-^{32}P-ATP and polynucleotide kinase at 37°C in a vol of 30 μl. After 10 min, ATP was added to 1 mM and the reaction was allowed to continue for another 30 min. After heat inactivation of the kinase (5 min at 70°C) the linkers were added to 50 μg of blunt-ended Ad12 DNA in 7.5 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP and six units T4 ligase (final vol 150 μl). Ligation was performed for 16 hr at

37°C. After phenol extraction and ethanol precipitation the viral DNA was cleaved with Eco RI, and the restriction fragments were isolated after agarose gel electrophoresis (Vogelstein and Gillespie, 1979).

Step 3. Cloning of the Ad12 Terminal Eco RI Fragments

The vector pAT153 (a deletion mutant of pBR322; Twigg and Sherratt, 1980) was cleaved with Eco RI and treated with calf intestine phosphatase (Weaver and Weissmann, 1979). The gel-purified Eco RI-linked terminal Eco RI fragments were ligated into the plasmid vector with T4 ligase for 16 hr at 15°C. This DNA was used to transform E. coli HB101. Ampicillin-resistant colonies were screened by in situ colony hybridization (Grunstein and Hogness, 1975). Hybridization-positive colonies were grown in 10 ml cultures and the plasmid DNA was isolated by the minilysate procedure (Birnboim and Doly, 1979). The plasmid DNA was analyzed by cleavage with various restriction endonucleases and agarose gel electrophoresis. In this way four clones were isolated containing the left-terminal Eco RI fragment. None of these clones lacked more than eight nucleotides from the left terminus, as DNA sequence analysis established.

DNA Sequence Analysis

5′- and 3′-terminal labeling of restriction fragments and nucleotide sequence determination was performed according to the method of Maxam and Gilbert (1980).

S1 Nuclease Analysis

Total cytoplasmic RNA was hybridized at 37°C to an end-labeled single-stranded DNA fragment for 3 hr, or at 53°C to an end-labeled double-stranded DNA fragment in 10–20 μl 80% formamide, 40 mM PIPES (pH 6.4), 1 mM EDTA and 0.4 M NaCl in an overnight incubation. Hybrids were incubated for 30 min at 37°C with 300 units S1 nuclease in 300 μl of 0.25 M NaCl, 30 mM Na acetate, 1 mM $ZnSO_4$ and 150 μg sonicated calf thymus DNA and separated on a 3.5% or 8% sequence gel (Maxam and Gilbert, 1980).

Peptide Maps

Tumor antigens were immunoprecipitated with hamster anti-T serum from extracts of ^{35}S-methionine-labeled transformed cells (279.C43) as described by Schrier et al. (1979). After SDS polyacrylamide gel electrophoresis and autoradiography of the wet gel, the 19 and 65 kd proteins were cut out of the gel and eluted with 1% SDS and 0.3% triethylamine (pH 10.5); 200 μg of bovine serum albumin was added as a carrier protein. After removal of the SDS and salts by gel filtration (R. Amons and P. I. Schrier, 1981), the proteins were oxidized with performic acid, lyophilized and treated two times for 2 hr at 37°C with 20 μg of TPCK-trypsin in 0.5 ml 0.1 M $(NH_4)_2$ CO_3. The digest was lyophilized and spotted on a cellulose thin-layer sheet. Chromatography was in pyridine:butanol:acetic acid:H_2O (10:15:3:12; v/v) and electrophoresis in 1.25% pyridine/acetic acid (pH 4.7).

Materials

Restriction enzymes were purchased from New England Biolabs (Beverly) except for Hind III, Acc I (gifts of H. M. C. Lupker-Wille) and Aos I (gift of M. Duyvesteyn). T4-DNA ligase was a gift of A. de Waard. T4-polynucleotide kinase was prepared as described by van Ormondt et al. (1978). Calf intestine phosphatase and Klenow enzyme were obtained from Boehringer-Mannheim and S1 nuclease from Sigma. α-^{32}P dNTPs were obtained from New England Nuclear and γ-^{32}P ATP from The Radiochemical Centre. The Eco RI-linker was a gift of H. Pannekoek. TPCK-trypsin was obtained from Worthington Biochemicals.

Acknowledgments

We thank Drs. R. A. Flavell, G. G. and F. G. Grosveld (National Institute for Medical Research) for advice on DNA cloning, Dr. K. Fujinaga for communicating his unpublished sequence data, Ms. M. Duyvesteyn for her able assistance, Ms. A. Houweling for performing the transformation experiment, Dr. A. de Waard for his support and interest and Ms. M. A. Veeren-Vink for typing the manuscript. This work was supported in part by The Netherlands Organization for the Advancement of Pure Research (ZWO) through the Netherlands Foundation for Chemical Research (SON) and Fundamental Medical Research (FUNGO).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 14, 1981

References

Aleström, P., Akusjärvi, G. Perricaudet, M., Mathews, M. B., Klessig, D. F. and Pettersson, U. (1980). The gene for polypeptide IX from adenovirus type 2 and its unspliced messenger RNA. Cell *19*, 671–681.

Amons, R. and Schrier, P. I. (1981). Removal of sodium dodecyl sulfate from proteins and peptides by gel filtration. Anal. Biochem., in press.

Anderson, C. W. and Lewis, J. B. (1980). Amino-terminal sequence of Ad2 proteins: hexon, fiber, component IX, and early protein 1B-15K. Virology *104*, 27–41.

Baker, C. C. and Ziff, E. B. (1980). Biogenesis, structures and sites of encoding of the 5′ termini of adenovirus-2 mRNAs. Cold Spring Harbor Symp. Quant. Biol. *44*, 415–428.

Baker, C. C. and Ziff, E. B. (1981). Promoters and heterogeneous 5′ termini of the messenger RNAs of adenovirus-2. J. Mol. Biol. *149*, 189–221.

Berk, A. J. and Sharp, P. A. (1978). Structure of the adenovirus 2 early mRNAs. Cell *14*, 695–711.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. *7*, 1513–1523.

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978). Evidence for a leader sequence in mRNA and DNA at the exon-intron bounderies. Proc. Nat. Acad. Sci. USA *75*, 4853–4857.

Chow, L. T., Broker, T. R. and Lewis, J. B. (1979). The complex splicing patterns of RNA from the regions of Ad2. J. Mol. Biol. *34*, 265–303.

Chow, L. T., Lewis, J. B. and Broker, T. R. (1980). RNA transcription and splicing at early and intermediate times after adenovirus-2 infection. Cold Spring Harbor Symp. Quant. Biol. *44*, 401–414.

Dijkema, R., Dekker, B. M. M., Van der Feltz, M. J. M. and van der Eb, A. J. (1979). Transformation of primary rat kidney cells by DNA fragments of weakly oncogenic adenoviruses. J. Virol. *32*, 943–950.

Esche, H., Mathews, M. B. and Lewis, J. B. (1980). Proteins and messenger RNAs of the transforming region of wild type and mutant adenoviruses. J. Mol. Biol. *142*, 399–417.

Flint, S. J. (1980). Transformation by adenoviruses. In Molecular Biology of Tumor Viruses *II:* DNA Tumor Viruses, J. Tooze, ed. Second edition. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratories) pp. 547–576.

Flint, S. J., Sambrook, J., Williams, J. F. and Sharp, P. A. (1976). Viral nucleic acid sequences in transformed cells. Virology *72*, 456–470.

Graham, F. L., Abrahams, P. J., Mulder, C., Heijneker, H. L., Warnaar, S. O., De Vries, F. A. J. and van der Eb, A. J. (1974). Studies on in vitro transformation by DNA and DNA fragments of human adenoviruses and simian virus 40. Cold Spring Harbor Symp. Quant. Biol. *39*, 637–650.

Green, M., Wold, W. S. M., Brackmann, K. H. and Cartas, M. A. (1979). Identification of families of overlapping polypeptides coded by early "transforming" gene region 1 of human Ad2. Virology *97*, 275–286.

Grunstein, M. and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Nat. Acad. Sci. USA *72*, 3961–3965.

Halbert, D. N., Spector, D. J. and Raskas, H. J. (1979). In vitro translation products specified by the transforming region of Ad2. J. Virol. *31*, 621–629.

Harter, M. L. and Lewis, J. B. (1978). Ad2 early proteins synthesized in vitro and in vivo: identification in infected cells of the 38,000–50,000 molecular weight protein encoded by the left end of the Ad2 genome. J. Virol. *26*, 736–749.

Houweling, A., van den Elsen, P. J. and van der Eb, A. J. (1980). Partial transformation of primary rat cells by the leftmost 4.5% fragment of adenovirus 5 DNA. Virology *105*, 537–550.

Jay, G., Nomura, S., Anderson, C. W. and Khoury, G. (1981). Identification of the SV40 agnogene product: a DNA binding protein. Nature *291*, 346–349.

Jochemsen, H. (1979). Studies on the transforming genes and their products of human adenovirus types 12 and 5. Ph.D. Thesis, State University Leiden.

Jochemsen, H., Daniels, G. S. G., Lupker, J. H. and van der Eb, A. J. (1980). Identification and mapping of the early gene products of adenovirus type 12. Virology *105*, 551–563.

Jochemson, H., Hertoghs, J. J. L., Lupker, J. H., Davis, A. and van der Eb, A. J. (1981). In vitro synthesis of adenovirus type 5 T antigens. II. Translation of virus-specific RNA from cells transformed by fragments of adenovirus type 5 DNA. J. Virol. *37*, 530–534.

Kitchingman, G. R. and Westphal H. (1980). The structure of Ad2 early nuclear and cytoplasmic RNAs. J. Mol. Biol. *137*, 23–48.

Kozak, M. (1981). Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. Curr. Topics Microbiol. Immunol. *93*, 81–123.

Lupker, J. H., Davis, A., Jochemsen, H. and van der Eb, A. J. (1981). In vitro synthesis of adenovirus type 5 antigens. I. Translation of early region-1 specific RNA from lytically infected cells. J. Virol. *37*, 524–529.

Maat, J. and van Ormondt, H. (1979). The nucleotide sequence of the transforming *Hin*dIII-G fragment of adenovirus type 5 DNA. The region between positions 4.5 (*Hpa*I site) and 8.0 (*Hin*dIII site). Gene *6*, 75–90.

Maxam, A. M. and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. *65*, 499–559.

Perricaudet, M., Le Moullec, J. M. and Pettersson, U. (1980). The predicted structure of two adenovirus T antigens. Proc. Nat. Acad. Sci. USA *77*, 3778–3782.

Ross, S. R., Flint, S. J. and Levine, A. J. (1980). Identification of the adenovirus early proteins and their genomic map positions. Virology *100*, 419–432.

Sawada, Y. and Fujinaga, K. (1980). Mapping of adenovirus 12 mRNAs transcribed from the transforming region. J. Virol. *36*, 639–651.

Schrier, P. I., van den Elsen, P. J., Hertoghs, J. J. L. and van der Eb, A. J. (1979). Characterization of tumor antigens in cells transformed by fragments of adenovirus type 5 DNA. Virology *99*, 372–385.

Sekikawa, K., Shiroki, K., Shimojo, H., Ojima, S. and Fujinaga, K. (1978). Transformation of rat cell line by an adenovirus 7 DNA fragment. Virology *88*, 1–7.

Shiroki, K., Handa, H., Shimojo, H., Yano, S., Ojima, S. and Fujinaga, K. (1977). Establishment and characterization of rat cell lines transformed by restriction endonuclease fragments of adenovirus 12 DNA. Virology *82*, 462–471.

Shiroki, K., Shimojo, H., Sawada, Y., Uemizu, Y. and Fujinaga, K. (1979). Incomplete transformation of rat cells by a small fragment of adenovirus 12 DNA fragment. Virology *95*, 127–138.

Shiroki, K., Segawa, K. and Shimojo, H. (1980). Two tumor antigens and their polypeptides in adenovirus type 12-infected and transformed cells. Proc. Nat. Acad. Sci. USA *77*, 2274–2278.

Spector, D. J., McGrogan, M. and Raskas, H. J. (1978). Regulation of the apparance of cytoplasmic RNA from region 1 of the adenovirus 2 genome. J. Mol. Biol. *126*, 395–414.

Sugisaki, H., Sugimoto, K., Takanami, M., Shiroki, K., Saito, I., Shimojo, H., Sawada, Y., Uemizu, Y., Uesugi, S.-I. and Fujinaga, K. 1980. Structure and gene organization in the transforming Hind III-G fragment of Ad12. Cell *20*, 777–786.

Twigg, A. J. and Sherratt, D. (1980). Trans-complementable copy-number mutants of plasmid Col E1. Nature *283*, 216–218.

van der Eb, A. J., van Kesteren, L. W. and van Bruggen, E. F. J. (1969). Structural properties of adenovirus DNAs. Biochem. Biophys. Acta *182*, 530–541.

van der Eb, A. J. and Graham, F. L. (1980). Assay of transforming activity of tumor virus DNA. Meth. Enzymol. *65*, 826–839.

van der Eb, A. J., van Ormondt, H., Schrier, P. I., Lupker, J. H., Jochemsen, H., van den Elsen, P. J., DeLeys, R. J., Maat, J., van Beveren, C. P., Dijkema, R. and de Waard, A. (1980). Structure and function of the transforming genes of human adenoviruses and SV40. Cold Spring Harbor Symp. Quant. Biol. *44*, 383–399.

van Ormondt, H., Maat, J., de Waard, A. and van der Eb, A. J. (1978). The nucleotide sequence of the transforming *Hpa*I E fragment of Ad5 DNA. Gene *4*, 309–328.

van Ormondt, H., Maat, J. and Dijkema, R. (1980a). Comparison of nucleotide sequences of the early E1a regions for subgroup A, B and C of human adenoviruses. Gene *12*, 63–76.

van Ormondt, H., Maat, J. and van Beveren, C. P. (1980b). The nucleotide sequence of the transforming early region E1 of adenovirus type 5 DNA. Gene *11*, 299–309.

Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. Proc. Nat. Acad. Sci. USA *76*, 615–619.

Weaver, R. F. and Weissmann, C. (1979). Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15 S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucl. Acids Res. *7*, 1175–1193.