Isolation of a cDNA encoding the adenovirus E1A enhancer binding protein: a new human member of the *ets* oncogene family

Fumihiro Higashino, Koichi Yoshida, Yukako Fujinaga, Koichi Kamio and Kei Fujinaga* Department of Molecular Biology, Cancer Research Institute, Sapporo Medical College, S-1, W-17, Chuo-ku, Sapporo 060, Japan

Received October 19, 1992; Revised and Accepted January 11, 1993

DDBJ accession no. D12765

ABSTRACT

The cDNA encoding adenovirus E1A enhancer-binding protein E1A-F was isolated by screening a HeLa cell λ gt11 expression library for E1A-F site-specific DNA binding. One cDNA clone produced recombinant E1A-F protein with the same DNA binding specificity as that endogenous to HeLa cells. Sequence analysis of the cDNA showed homology with the ETS-domain, a region required for sequence-specific DNA binding and common to all ets oncogene members. Analysis of the longest cDNA revealed about a 94% identity in amino acids between human E1A-F and mouse PEA3 (polyomavirus enhancer activator 3), a recently characterized ets oncogene member. E1A-F was encoded by a 2.5kb mRNA in HeLa cells, which was found to increase during the early period of adenovirus infection. In contrast, ets-2 mRNA was significantly reduced in infected HeLa cells. The results indicate that E1A enhancer binding protein E1A-F is a member of the ets oncogene family and is probably a human homologue of mouse PEA3.

INTRODUCTION

The adenovirus(Ad) E1A gene is the first transcription unit to be activated in productive infection. The E1A gene encodes a potent activator for transcription of viral early and late genes(1). Defects of the E1A gene cause delayed onset of transcription from other early genes, E1B through E4, and viral mutants grow at extremely reduced levels(2,3). At least four separate enhancer elements are required for active transcription of the adenovirus type 5(Ad5) E1A gene(4-7). Transient expression assay with a series of deletion plasmids has showed a far upstream enhancer element, located around position -340 relative to the E1A cap site(4). Hearing and Shenk found element I of E1A enhancer at two separate sites, positions -200 and -300(5,6); Ad5 mutants with deletions of either copy of the element I produce low levels of E1A-specific mRNA. Recent analysis of viral mutants has showed transcriptional significance of an another copy of element I, located at position -270(7).

We have previously identified at least twenty-one protein-

Here, we have isolated the cDNA encoding Ad5 E1A enhancer-binding protein (called E1A factor: E1A-F) from a HeLa cell λ gt11 expression library by utilizing its ability to bind E1A-F sites. The cloned cDNA contained the ETS-domain, a region conserved in all *ets* oncogene members and required for specific DNA binding. Deduced amino acid sequences showed significant homology with recently characterized mouse PEA3, a member of the *ets* oncogene family.

MATERIALS AND METHODS

Cloning of the E1A-F cDNA

A HeLa cell cDNA expression library in λ gt11 vector (Clontech Laboratories,Inc.) was screened using *in situ* sequence-specific DNA binding activity as described(18). Screening was carried out with a 5'end-labeled DNA probe containing six tandem E1A enhancer elements (F-wt:5'-AC<u>AGGAAGT</u>GACA<u>CGGATGT</u>GGC-3', core sequences are underlined) and four tandem mutant

binding sites on the upstream region of the Ad5 E1A gene(8). Several of those contain sequences identical or similar to the sequence recognized by transcription factors such as ATF(9), NFIII(Oct-1)(10) and E2F(11). Interestingly, the factor bound to copies of the E1A enhancer element I and far upstream enhancer was recovered in a single fraction of heparin-agarose chromatography(8). Nucleotide sequences involved in protein-DNA contacts were similar with each other(8,12). These suggest E1A enhancer elements may interact with common factor(s). E1A enhancer elements contain core binding sequences 5'-A/CGGAA/TGT-3'(3,4,12), which overlap the sequence recognized by ets oncogene family members. Several ets-related genes encode putative transcription factors for viral and cellular genes(13). The ets-1 is a cellular proto-oncogene of the v-ets of avian leukosis virus E26(14). Exogenous expression of the ets-1 activates the polyomavirus enhancer through site-specific DNA binding(15). The ets-1 as well as AP-1 binding sequences mediates serum and oncogene-induced activation of some cellular genes(16). The ets oncogene family may play a critical role in cell growth regulation(17).

^{*} To whom correspondence should be addressed

sequences (F-mut:5'-AC<u>ACCAAGT</u>GACA<u>CCCATGT</u>GGC-3'). After four rounds of plaque-purification, the cDNA insert was cleaved out with EcoRI digestion, recloned into pUC119 and sequenced by the dideoxy method(19). The isolated cDNA was used as a probe for further screening of the HeLa cell λ gt11 library by plaque hybridization and the KB cell cDNA library in pCD vector by colony hybridization to isolate longer cDNA. These cDNA were also recloned into pUC119 or pUC118 and sequenced.

For DNA binding specificity, Y1089 lysogens were isolated(20). Induced synthesis of β -galactosidase fusion proteins with Isopropyl β -D(-)-Thiogalactopyranoside (IPTG) was carried out as described(18). Cells from 4 ml aliquots of the induced lysogen cultures were rapidly pelleted and resuspended in 50 μ l of lysis buffer (50 mM Tris pH 7.5, 1mM EDTA, 1mM DTT, 0.1mM α PMSF). Cell suspensions were subjected to three cycles of freezing and thawing. After a 30 min spin in a microfuge at 4°C, the supernatants were directly used for gel retardation and methylation interference assays.

Gel retardation assay

Standard binding reactions containing the β -galactosidase fusion protein and 5'-end labeled probes were carried out in 15 μ l mixtures in binding buffer (20mM Hepes-NaOH pH 7.8, 50mM NaCl, 0.5mM EDTA, 5% glycerol, 2 μ g of poly dI-dC/dI-dC and 6 μ l of the lysogen extract). For competitions, 5 –200 fold molar excess amounts of DNA were added into the reaction mixtures. The mixtures were kept on ice for 10 min and then ³²P-labeled DNA probe (~0.5ng) was added. After incubation for 15 min at room temperature, DNA-protein complex was resolved in 4% non-denaturing polyacrylamide gel.

Methylation interference assay

DNA labeled at the 5' end was partially methylated by treating for 3 min with 1μ l Dimethyl sulfate(DMS) as described(21). After ethanol-precipitation twice, DNA was incubated with the lysogen extract and electrophoresed in 4% non-denaturing polTyacrylamide gel. After short term autoradiography, DNA was eluted from the region of gels containing free DNA and protein-DNA complexes as described(22). DNA was purified from co-eluted acrylamide by a NACS column chromatography(BRL) according to the supplier's manual. DNA was cleaved with 100 μ l of 1M piperidine and equal amounts of DNA (in cerenkov count) were electrophoresed in 12% polyacrylamide sequencing gels containing 8M urea with A + G Marker.

DNA sequencing

DNA sequencing was performed by the dideoxy sequencing method(19), according to the suppliers manual (U.S.B.corp.). The longest cDNA clone was sequenced with Applied Biosystems Model 373A auto-sequencer.

Construction of effector and reporter plasmids

The expression vector pSTC contains the human CMV promoter, HSV tk gene 5' untranslated leader and initiation codon, rabbit β -globin gene splicing and polyadenylation signals, and the replication origin of SV40(23). To produce the pCMVE1A-F plasmid, the E1A-F cDNA coding sequences were cut out from the longest cDNA clone with Eco47III and XbaI, filled in with Klenow enzyme, and inserted into the SmaI site of the pSTC plasmid. The coding sequences begin at the tk gene AUG codon and encode an extra 7 N-terminal amino acid and 453 amino acids long E1A-F protein. The chloramphenicol acetyltransferase (CAT) expression plasmids were constructed by inserting three or six copies of the wild type E1A enhancer core element (F-wt) or two copies of the mutated version (F-mut) into the unique BgIII site of the SV40 early core promoter-CAT (pA10CAT2, ref.24).

DNA transfection and CAT assay

Human 293 cells(5×10^{6} per 10cm plate), were co-transfected with 5.0µg of a reporter CAT plasmid, 2.5µg of the pCMVE1A-F plasmid and 5.0µg of the pRSV-βgal plasmid by using the calcium phosphate-DNA coprecipitation method(25). Cells were harvested 48hr post-transfection and lysed by three successive cycles of freezing and thawing. Cell extracts were assayed for CAT activity by incubation with acetyl coenzymeA and [¹⁴C] chloramphenicol(26), followed by thin-layer chromatography and quantitated by cutting the radioactive spots out of the plates after autoradiography. β -Galactosidase activity was determined by using O-nitrophenyl- β -D-galactopyranoside as a substrate. CAT activity was corrected for differences in transfection efficiency by normalizing to β -galactosidase activity.

Northern blot hybridization

HeLa cells were infected with human Ad5 (strain 300) at 20 plaque-forming units (PFU)/cell or with Ad5 dl312 (E1A genedeleted mutants) at 50 PFU/cell. Cells were suspended in hypotonic buffer and lysed with 0.5% NP-40(27). Cytoplasmic RNA was extracted twice with phenol/chloroform mixture and precipitated with 2 volumes of ethanol. RNA ($20\mu g$ /lane) was applied to 1.4% agarose gels containing 2.2M formaldehyde in MOPS-running buffer(28), transferred onto nitrocellulose filter (S&S, BA85) and probed in high stringency with ³²P-DNA (specific activity, $1-3 \times 10^8 \text{cpm}/\mu g$) labeled by the random priming method(28). Filters were washed twice with $2 \times \text{SSC}/0.2\%$ SDS at 55°C and exposed to Fuji RX X-ray films with intensifying screens at -70° C.

RESULTS

Isolation of the cDNA encoding E1A enhancer binding protein, E1A-F

HeLa cell λ gt11 expression library was screened by using catenated E1A enhancer core elements that contained either wild type or mutant sequences (Fig.1). About 1.2×10^6 plaques were screened in non-denaturing binding buffer without guanidine-HCl according to Singh et al.(18). Several phage plaques were positive in the first screening. After four rounds of plaque purifications, only one clone, designated λ 3-16, bound to the wild type probe (F-wt) though not to the mutant probe (F-mut)(Fig.1B). Thus the λ 3-16 was a candidate for the cDNA clone to encode E1A-F.

Binding specificity of the λ 3-16 β -galactosidase fusion protein to E1A enhancer core sequences

In order to analyse the DNA binding specificity of the cDNAencoding protein, lysogens were isolated with the λ 3-16 and λ gt11 vector phages. Extracts of the IPTG-induced lysogens, which were expected to contain β -galactosidase fusion protein, were subjected to gel retardation assays. As shown in Figure 2, extract of the λ 3-16 lysogen contained the binding activity to ³²P-labeled F-wt but not to the F-mut, as well as DNA-affinity chromatography fraction of HeLa cell E1A-F (Fig.2, lanes 2 and 8, lanes 9 and 10). The binding activity was not detected in the



Figure 1. Specific binding of λ 3-16 phage plaques with E1A enhancer core sequences. (A) Locations of Ad5 E1A enhancer core sequences. (B) The λ 3-16 phage cDNA clone possibly encode E1A-F was isolated from HeLa cell λ gt11 cDNA library by the method of Singh et al.(18). The λ 3-16 plaques were lifted onto a nitrocellulose filter and probed with F-wt DNA containing three tandem copies of two E1A enhancer core elements (5'-AGGAAGT-3' and 5'-CGGATGT-3'). A half of the filter was also probed with ³²P-labeled F-mut DNA containing two tandem copies of mutant core elements (5'-ACCAAGT-3' and 5'-CCCATGT-3'). Filters were exposed overnight to X ray films at -70° C.



Figure 2. Competition gel mobility-shift assay. Lysates of λ gt11 (lane 1) and λ 3-16 (lanes 2–8) lysogens and partially purified HeLa cell E1A-F (lanes 9–10) were incubated with ³²P-labeled F-wt (lanes 1–7 and 9) and F-mut probes (lanes 8 and 10). DNA-protein complex(C) was separated from free probe(F) by 4% PAGE. For competition, excess molar amounts of unlabeled F-wt (lanes 3–6) or F-mut (lane 7) were added as indicated.



Figure 3. Methylation interference assay with HeLa cell E1A-F protein and extracts of the λ 3-16 lysogen. DNA-protein complexes (lanes 3 and 4) were formed by incubation with partially methylated ³²P-labeled Ad5 RsaI/SspI fragments (-303 to -158 relative to the E1A cap site) and separated from the free probe (lane 2) by 4% PAGE. A+G (lane 1) refers to chemical cleavage products that were used as markers. The regions corresponding to the E1A enhancer core elements were indicated by their sequences. G residues whose methylation affected bindings were marked with triangles and nt numbers.

 λ gt11 lysogen extract (Fig.2, lane 1). For specific binding, molarexcess amounts of competitor oligonucleotides were added in binding reactions. A 20-fold molar excess of F-wt competitor almost completely eliminated the binding activity (Fig.2, lane 4), while F-mut DNA had little effect even in the 200-fold molar excess amount (Fig.2, lane 7). Almost similar profiles of competition were obtained when probed with the Ad5 RsaI/SspI fragment containing E1A enhancer element (-303 to -158 relative to the E1A cap site)(data not shown).

The binding site of λ 3-16 β -galactosidase fusion protein was determined by the methylation interference assay and compared with that of E1A-F protein endogenous to HeLa cells. The Ad5 RsaI/SspI fragment, partially methylated with DMS, was incubated with extracts of the λ 3-16 lysogen and endogenous HeLa cell E1A-F. Free DNA and protein–DNA complexes were separated by gel electrophoresis. DNA was extracted, cleaved with piperidine and resolved in sequencing gels. Methylation of G residues at –298, –299, –268 and –269 affected DNA-binding with extracts of the λ 3-16 lysogen (Fig.3, lane 4). These G residues resided in the E1A enhancer core elements (Fig.1A). The methylation effect was comparable with that obtained by HeLa cell E1A-F protein (compare Fig.3, lanes 3 and 4). Similar methylation interferences were observed when probed with

1	CCG P	GAA E	ATG M	GGA G	GAC D	TTG L	CGC R	GAA E	GCG A	CTG L	ATC 1	GGC G	CCG P	CTO L	GGG G	AAG K	СТС L	ATG M	GAC D	CCG P	GGC G	TCC S	CTG L	CCG P	CCC P	75
26	։ Ե	GAC D	TCT S	GAA E	GAT D	СТС L	TTC F	CAG Q	GAT D	CTA L	AGT S	CAC H	TTC F	CAG Q	GAG E	ACG T	TGG W	CTC L	GCT A	GAA E	GCT A	CAG Q	GTA V	CCA P	GAC D	150
51	AGT S	GAT D	GAG E	CAG Q	TTT F	GTT V	сст Р	GAT D	TTC F	CAT H	TCA S	GAA E	AAC N	CTA L	GCT A	TTC F	CAC H	AGC S	CCC P	ACC T	ACC T	AGG R	ATC I	AAG K	AAG K	225
76	GAG E	CCC P	CAG Q	AGT S	CCC P	CGC R	ACA T	GAC D	CCG P	GCC A	CTG L	TCC S	TGC C	AGC S	AGG R	AAG K	CCG P	CCA P	CTC L	CCC P	TAC Y	CAC H	CAT N	00C 0	GAG B	300
101	CAG Q	TGC C	CTT L	TAC Y	TCC S	AGT S	GCC A	TAT Y	GAC D	CCC P	CCC P	AGA R	CAA Q	ATC 1	GCC A	ATC I	AAG K	TCC B	CCT P	acc A	CCT P	00T G	GCC A	стт L	GGA G	376
126	CAG Q	TCG S	CCC P	CTA L	CAG Q	CCC P	TTT F	CCC P	CGG R	GCA A	GAG E	CAA Q	COO R	AAT N	TTC F	CTG L	AGA R	TCC S	TCT 8	00C G	ACC T	TCC S	CAG `Q	CCC P	CAC H	450
151	P	GGC G	CAT H	GGG G	TAC Y	CTC	000 G	GAA E	CAT H	AGC S	TCC 8	GTC V	TTC F	CAG Q	CAG Q	CCC P	CTG L	GAC D	ATT	TOC C	CAC N	TCC 8	TTC F	ACA T	TCT S	525
176	CAG Q	GGA G	600 6	00C G	CGG R	GAA B	CCC P	CTC L	CCA P	OCC A	CCC P	TAC Y	CAA Q	CAC	CAG Q	CTG L	TCG 8	GÁG E	CCC P	TOC C	CCA P	CCC P	TAT Y	CCC P	CAG Q	600
201	CAG Q	AGC S	TTT F	AAG K	CAA Q	GAA E	TAC Y	CAT H	ĞAT D	CCC P	CTG L	TAT Y	GAA E	CAG Q	0CG A -	00C 0	CAG Q	CCA P	GCC A	oto V	GAC D	CAG Q	G G G	000 G	otc v	675
226	AAT N	GGG G	CAC H	AGG R	TAC Y	CCA P	000 0	GCG A	GGC G	ata V	GTO V	ATC I	AAA K	CAG Q	GAA E	CAG Q	ACG T	GAC D	TTC F	GCC A	TAC Y	GAC D	TCA B	GAT D	otc V	750
251	ACC T	660 G	TGC C	GCA A	TCA S	ATG H	TAC Y	CTC L	CAC H	ACA T	GAG E	00C 0	TTC F	TCT S	000 G	CCC P	тст 8	CCA P		GAC D	000 0	GCC A	ATG H	00C 0	TAT Y	825
276	GGC G	TAT Y	GAG E	AAA K	CCT P	CTG L	CGA R	CCA P	TTC F	CCA P	GAT D	GAT D	GTC V	TOC C	GTT V	GTC V	CCT P	GAG E	AAA K	TTT F	GAA E	GGA G	GAC D	ATC I	AAG K	900
301	CAG Q	GAÀ E	000 G	GTC V	GCT G	GCA A	TTT F	CGA R	GAG E	000 G	CCG P	CCC P	TAC Y	CAG Q	COC R	COG R	GGT G	OCC A	CTO L	CAG Q	CTG L	100 W	CAA Q	TTT	CTO L	975
326	GTG V	GCC A	TTC L	CTG L	GAT D	GAC D	CCA P	ACA T	AAT	GCC	CAT H	TTC F	ATT	GCC A	тос ¥	ACG T	66C 6	CGG R	GOA G	ATG H	GAG E	TTC F	AAG K	CTC L	ATT I	1050
351	GAG E	CCT P	GAG E	GAG E	GTC V	GCC A	AGG R	CTC L	TGG W	GGC G	ATC I	CAG Q	AAG K	AAC N	CGG R	CCA P	GCC A	ATG M	AAT N	TAC Y	GAC D	AAG K	CTG L	AGC 8	COC R	1125
376	TCG S	CTC L	CGA R	TAC Y	TAT Y	TAT Y	GAG E	AAA K	00C 0	ATC 1	ATG M	CAG Q	AAG K	ata V	GCT A	001 0	GAG E	COT R	TAC Y	GTO V	TAC Y	AAG K	TTT P	GTG V	TOT C	1200
401	GAG E	CCC P	GAG E	GCC A	CTC L	TTC F	TCT 8	TTO L	GCC A	TTC F	CCO P	GAC D	AAT N	CAG Q	CGT R	CCA P	GCT A	CTC L	AAG K	GCT A	GAG B	TTT	GAC D	COG R	CCT	1275
126	стс v	AGT S	GAG E	GAG E	GAC D	ACA T	GTC V	CCT P	TTG L	TCC S	CAC H	TTG L	GAT D	GAG E	AGC S	CCC P	acc A	TAC Y	CTC L	CCA P	GAG E	ста L	OCT A	00C G	CCC P	1350
151	GCC A	CAG Q	CCA P	TTT F	66C G	CCC P	AAG K	GOT G	00C G	TAC Y	TCT S	TAC Y	TAG *	ccco	CAG	COOC1	IGTTO	cccc	TOCO	GCAG	1010 0	00100	TOC	CTG	GTAC	: 1437
	ATATAAATGAATCT0GT0TT0G00AAACCTTCATCT0AAACCCACAGAT0TCTCT0G000CA0ATCCCCACT0TCCTACCA0TT0CCCTA0CCCAGACTCT 1												1537													
	GAGO	TGC1	CACO	GGAC	ITCA1	TGGC	GAAGO	JAAA/	AGTGO	AGA/	ATG	CAAC	тст/	GAG	rctc/	AGAA/	ACTCO	CCT	36666	TTT	CACCI	10060	CCT	GAGO	38811	1637
	CAGO	TCAC	CTTO	TTCC	TAGO	tcc/	AGCO	ccco	CACAG	CTT	rtcco	CAAC	CAC	GAG	ACA	GAGI	rttgi	TCTO	ITTCI	1000	MAC/	GAG/	1000	COCTI	rccc/	1737
	ACTTCATACTGGCAGGAGGGTGAGGAGGTTCACTGAGCTCCCCAGATCTCCCCACTGCGGGGGAGACAGAAGCCTGGACTCTGCCCCCACGCTGTGGCCCTGG													1837												
	AGGG	тссс	GGTT	TGTO	AGTI	CTTG	атас	тста	TOTI	·ccc/	GAGO	CAGO	caa	OOT	GAAD	34440	GAAC	CTG	MATO	AGGO	00TO	CT G G G	TAT/	AGC/	GAGA	1937
	NGRATGGRTTCCTGCTCCAAGGGACCCTTTGCCTTTCTTGCCCTTTCCTAGGCCCAGGCCTGGGTTTGTACTTCCACCTCCACCACATCTGCCAGACC 20													2037												
	T T <u>AA</u>	TAAA	leecc	CCCA	CTTC	tccc	****								•											2093

Figure 4. Nucleotide and deduced amino acid sequences of E1A-F cDNA. Nucleotide numbers are shown on the right side; amino acid numbers are on the left. The boxed region denotes the ETS-domain common to all of the *ets* oncogene members(13). Acidic and glutamine-rich regions are underlined and putative ATP-binding domain was indicated by an arrow. A putative poly(A) signal (5'-AATAAA-3') is double underlined.

consensus		*S*****LW*FLL*LL*****	***I*W*******	**************************************	R*WG**K**K**M*Y****R*LR*Y**K****K*				
		Y		н	R	R	ĸ		
3	Identity	£							
ELA-E	100%	RRGA LQLWQFLVALLDDPT	AHFIAWT GRGMBFK	LIEPEEVA RLWGIQK	NRPAMNYDKLS	RSLRYYYEKGIMQKV	GERYVYKFV		
ets-J	62%	GS-P ILET-KS	QSSD-W	-SD-DRKR-	-K-KB	-GD-N-IH-T-	-KR		
ets-2	2 61%	GS-P ILBS-KS	QSSD-W	-AD-DRKR-	-K-KB	-GD-N-IH-TS	3-KR		
erg	59%	GS-Q ILES-SS	SSC-T-E -TNG	MTD-DRBR-	SK-N	-AD-NTH	I-KAD		
elk	608	PS VTLQREQG	G-I-SSRD-G	-VDA LR-	- KTN	-AD-N-IR8	3-QKF		
PU.1	328	GSKKKIR-YLDRSGD	KDS-W-VDKDKGT-Q	FSSKHKE-LAHR	GKK-T-Q-M2	A-AN-G -TGEV-K	/KKKLT-Q-S		
GABPO	2 53%	NN-Q ILET-KDA	RDC-S-V -DBG	-NQL QKQR-	-K-TB	-ADGDMIC()-K-F		
PEA3	100%								

Figure 5. Comparison of amino acid sequences in the ETS-domain between E1A-F and other *ets* oncogene members. The *ets* oncogenes examined are indicated at the left with% identity of amino acids. Amino acid sequences identical with E1A-F are hyphenated. Consensus amino acids according to Karim et al.(13) are shown at the top of the figure.

catenated E1A enhancer core elements (data not shown). These results showed that the cloned cDNA encoded a E1A enhancer binding protein almost identical in binding specificity with E1A-F protein endogenous to HeLa cells.

Nucleotide sequences of E1A-F cDNA

The λ 3-16 clone contained about a 0.46kb-long cDNA, which was considerably shorter than the full length cDNA as estimated from the size of mRNA (Fig.6, approximately 2.5kb). Further

screening was carried out by plaque-hybridization with the 0.46kb cDNA to isolate longer cDNA. A Human KB cell cDNA library in the pCD vector(29,30) was also searched by colony-hybridization. The longest cDNA was obtained by screening $\sim 4.5 \times 10^5$ colonies of KB cell cDNA library. Nucleotide sequences are shown in Figure 4. The cDNA comprised 2093bp with a large open reading frame (ORF) and the carboxyl-terminal region contained nucleotide sequences identical to those of the 0.46kb cDNA originally isolated. The ORF comprised 462 amino



Figure 6. Activation of the E1A enhancer core element-containing SV40 core promoter by cloned E1A-F protein. CAT reporter constructs with wild-type or mutated E1A enhancer core elements were transfected into 293 cells with or without the pCMVE1A-F plasmid expressing E1A-F protein, using the calcium phosphate precipitation technique described in Materials and Methods. CAT activity was measured 48hr after transfection and expressed in relation to the basal activity of the SV40 core promoter with wild-type E1A enhancer core element (6xF-wt). Experiments were repeated three times with similar quantitative results. One representative autoradiograph is shown. Abbreviations: 6xF-wt, six tandem copies of wild-type E1A enhancer core element; 2xF-mut, two tandem copies of mutant E1A enhancer core element.

acids. Judging from the molecular mass of the HeLa cell E1A-F protein (60kD, unpublished data), we estimated that the E1A-F cDNA clone may lack a coding sequence for about 90 amino acids on the amino-terminal of the ORF. Following the stop codon at nucleotide (nt) 1387(TAG), the 3' non-coding region of the cDNA comprised 704bp including a 29nt poly(A) tail and a putative poly(A) signal at nt 2040. A computer search using the Genbank and NBRF databases revealed that the carboxyl-terminal region had high homology with the ETS-domain common to the members of ets oncogene family (Fig.4, boxed region). The ETSdomain consists of about 85 amino acids and is involved in sequence-specific DNA binding(13). Figure 5 shows the homology in the ETS-domain between the E1A-F and other ets oncogene members. E1A-F has an approximately 60% identity with ets-1(31), ets-2(32,33), erg(34), elk(35) and GABPa(36), though only a 32% identity with PU.1(37). Twenty-five of twenty-seven ETS consensus amino acids were conserved in E1A-F. The E1A-F ETS-domain was identical to that of the polyomavirus(Py) enhancer activator 3 (PEA3, ref.38)(Fig.5). Acidic and glutanime-rich regions were also conserved in E1A-F and PEA3 (Fig.4, amino acid numbers 27 to 53 and 126 to 222, underlined). Overall, there was an about 94% identity in amino acids between them (E1A-F: 462 residues determined so far; PEA3: 555 residues). E1A-F may be a human homologue of mouse PEA3.

Transcriptional activation by E1A-F

To elucidate whether E1A-F could activate the promoter, we performed transient expression assay in human 293 cells by cotransfection of effector and reporter plasmids. Human 293 cells do not express E1A-F mRNA (Fig.7B, lane 6). The effector expression plasmid (pCMVE1A-F) contains the cytomegalovirus promoter, the herpes thymidine kinase untranslated leader sequence and initiation codon, the coding region of the E1A-F cDNA, which is 453 amino acids long, and the β -globin splicing and poly(A) signals. The reporter CAT plasmids contain the SV40 early core promoter carrying six tandem copies of wild-



Figure 7. Northern blot analysis of E1A-F mRNA in HeLa cells infected with adenoviruses. Cytoplasmic RNA was isolated from HeLa cells infected with Ad5 or Ad5 dl312 at the indicated infection times. RNA ($20\mu g$ /lane) was probed in high stringent conditions with (A) Ad5 E1A DNA (nucleotides 1–1338), (B) E1A-F cDNA (nucleotides 922–1392), or (C) human ets-2 cDNA (EcoRI fragment of cDNA 14 clone, ref.39). A blot was also probed with human β -actin to ensure equal loading of the sample.

type E1A enhancer core elements ($6 \times F$ -wt) or two tandem copies of mutant E1A-F enhancer core elements ($2 \times F$ -mut): these sites are positioned approximately 140bp upstream of the SV40 early RNA start site. As shown in Figure 6, the SV40 core promoter with wild-type E1A enhancer core elements was activated an average of 3.4-fold by co-transfection with pCMVE1A-F. However, E1A-F has little effect on the activity of the SV40 core promoter with mutant E1A enhancer core elements. These results indicate that E1A-F is capable of activating the SV40 core promoter through E1A-F binding sites.

Effects of adenovirus infection on the expression of E1A-F gene

Assuming that the cloned E1A-F gene encodes an activator for E1A gene transcription, the expression of E1A-F can be increased by Ad infection in parallel with an activation of the E1A gene. HeLa cells were infected with Ad5 and Ad5 dl312 mutant and cytoplasmic RNA was extracted at different times after infections. RNA was analyzed by Northern blot hybridization with ³²Plabeled Ad5 E1A DNA, E1A-F cDNA and human ets-2 cDNA (Fig.7). Ets-2 is a member of the ets oncogene family and is expressed in many kinds of cells with different origins. A blot was also probed with human β -actin to ensure equal loading of RNA. An about 2.5kb mRNA was clearly detected with E1A-F cDNA probe in HeLa cells (Fig.7B, lane 1), but was not found in 293 cells (Fig.7B, lane 6). No detectable hybridization was obtained in the 293 cells under stringent conditions even when 10µg of polyA-containing RNA was used (data not shown). Ets-2 specific mRNA (4.7, 3.2 and 2.7kb, ref.39) was detected in both HeLa and 293 cells (Fig.7C, lanes 1 and 6). Ad5 E1A mRNA was detectable at 3hr post-infection by longer exposure to a Xray film and after that increased gradually (Fig.7A). E1A-F mRNA increased in some extent (about 2 fold) during early times of Ad5 infections (Fig.7B, lanes 1-5), while ets-2 mRNA

reduced significantly with delay at the first 3hr post-infection (Fig.7C, lanes 1-5). Levels of E1A-F and *ets*-2 mRNA did not change by infection with E1A gene-deleted Ad5 mutant, dl312 (Fig.7B and C, lanes 7-11), suggesting differential regulation of *ets* oncogene members by Ad5 E1A. The results support the idea that E1A-F gene encodes an activator for transcription of Ad E1A gene.

DISCUSSION

We have isolated the cDNA which encodes the Ad E1A enhancerbinding protein, E1A-F. The deduced amino acid sequences showed high homology with the ETS-domain, the DNA binding region in the *ets* oncogenes. The ETS-domain is about 85 amino acids long(13); it contains three tryptophan repeats spaced 18 or 19 residues apart in the amino-terminal half and is rich in basic amino acids in the carboxyl-terminal half(13). Furthermore, E1A-F conserves virtually all of the structural features of the ETSdomain (Fig.5). Therefore, we conclude that E1A-F is a member of the *ets* oncogene family.

Other domains to be found in such genes have been enumerated by Seth et al., who predicted the presence of an amphipathic helixloop-helix, a cell division motif, a nuclear localization signal, an ATP-binding domain and a Cx type metal-binding finger(40). ATP-binding domains exist in many kinds of kinases, containing glycine-rich sequences with the consensus motif of GXGXXG(41). At 10 to 15 amino acids apart from the carboxylterminal of this motif, a conserved lysine residue is present. In E1A-F, ATP-binding like domain was observed upstream of the ETS-domain (GDGAMG, see Fig.4) and a lysine residue was found 20 amino acids downstream of the consensus motif. No other putative domains have yet been established in E1A-F.

Amino acid sequences of E1A-F showed striking homology (about 94% identity) with that of mouse PEA3(38). Furthermore, the ETS-domain of E1A-F and PEA3 proteins was identical. Since the PEA3 motif overlapped E1A enhancer core sequence (5, 6, 8, 12), it seems reasonable to conclude that E1A-F is a human homologue of mouse PEA3. Acidic and glutaminerich regions resided in the amino-terminal side of mouse PEA3(38). These amino acids generally constitute activation domains of transcription factors(42,43). Both acidic and glutamine-rich regions were also conserved in E1A-F (Fig.4). The acidic region was the same as that of PEA3, but the glutamine-rich region was slightly different (86% identity). Indeed, exogenous expression of E1A-F was found to activate the CAT reporter gene through multimers of E1A-F sites (see Fig.6). These domains may play an important role in transcriptional activation. However, further experiment is required to elucidate the role of E1A-F in the regulation of E1A gene expression.

It is interesting that two members of the *ets* oncogene family, E1A-F and *ets*-2, were oppositely influenced by Ad infection, and probably by the E1A gene. The expression of E1A-F was activated by E1A, suggesting that E1A may regulate itself positively through enhanced E1A-F binding to enhancer elements. On the other hand, *ets*-2 mRNA was significantly decreased by E1A. E1A-F and *ets*-2 genes may have a reciprocal relationship in the regulation of Ad E1A gene. Xin et al.(38) found two mRNA species (2.4 and 4.1kb) in HeLa and 293 cells by hybridization with a mouse PEA3 probe. In this study, the 2.4kb mRNA was detectable in HeLa cells though not in 293 cells even when polyA-containing RNA was used. We also noted that the hybridization signal with the PEA3 probe was faint. The reason for this discrepancy is not clear. However, the possibility is that there is some variety of 293 cells dependent on laboratory growth conditions and thus selection pressures. A 4.1kb mRNA was observed in HeLa and 293 cells only under low stringent conditions adopted in the study of Xin et al.(38)(data not shown).

Bruder and Hearing(12) identified cellular protein(s) bound to Ad5 E1A enhancer element I, termed EF-1A. Watanabe et al.(44) purified HeLa cell protein bound to an Ad5 E4 transcription control element (5'-ACGGAAGTG-3'), termed E4TF1. EF-1A, E4TF1 and E1A-F in this study recognize common or related sequences. It remains to be resolved what role, if any, they play in the activation of the E1A gene. *Ets*-1 and -2 oncogenes activated the Py enhancer through the PEA3-binding motif(15). The cloned mouse PEA3 gene also activated artificial promoter through the PEA3-binding motif(38). It is possible that Ad E1A enhancer harboring 5'-^A/_CGGA^A/_TGT-3' sites is recognized and regulated by multiple members of the *ets* oncogene family.

ACKNOWLEDGEMENTS

We thank Shigeki Koizumi for helpful discussions and Takeshi Ohno for technical supports. We also thank Takis S.Papas for c-ets-1 and c-ets-2 plasmids and Walter Schaffner for pSTC plasmid. This work was supported by Grants in Aid for Cancer Research and for Scientific Research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

- 1. Berk, A.J. (1986) Annu. Rev. Genet., 20, 45-79.
- 2. Harrison, T.J., Graham, F.L. and Williams, J.F. (1977) Virology, 77, 319-329.
- 3. Jones, N. and Shenk, T. (1979) Cell, 17, 683-689.
- Hen, R., Borrelli, E., Sassone-Corsi, P. and Chambon, P. (1983) Nucleic Acids Res., 11, 8747-8760.
- 5. Hearing, P. and Shenk, T. (1983) Cell, 33, 695-703.
- 6. Hearing, P. and Shenk, T. (1986) Cell, 45, 229-236.
- 7. Bruder, J.T. and Hearing, P. (1991) J. Virol., 65, 5084-5087.
- Yoshida, K., Narita, M. and Fujinaga, K. (1989) Nucleic Acids Res., 17, 10015-10034.
- Lee, K.A.W., Hai, T.Y., SivaRaman, L., Thimmappaya, B., Hurst, H.C., Jones, N.C. and Green, M.R. (1987) Proc. Natl. Acad. Sci. USA, 84, 8355-8359.
- Harvey, R.P., Robins, A.J. and Wells, J.R.E. (1982) Nucleic Acids Res., 10, 7851-7863.
- 11. Kovesdi, I., Reichel, R. and Nevins, J.R. (1986) Cell, 45, 219-228.
- 12. Bruder, J.T. and Hearing, P. (1989) Mol. Cell. Biol., 9, 5143-5153.
- Karim, F.D., Urness, L.D., Thummel, C.S., Klemsz, M.J., Mckercher, S.R., Celada, A., Van Beveren, C., Maki, R.A., Gunther, C.V., Nye, J.A. and Graves, B.J. (1990) Genes Dev., 4, 1451-1453.
- Leprince, D., Gegonne, A., Colli, J., de Taisne, C., Schneeberger, A., Lagrou, C. and Stehelin, D. (1983) Nature, 306, 395-397.
- Wasylyk, B., Wasylyk, C., Flores, P., Begue, A., Leprince, D. and Stehelin, D. (1990) Nature, 346, 191-193.
- Wasylyk, C., Flores, P., Gutman, A. and Wasylyk, B. (1989) EMBO J., 8, 3371-3378.
- 17. Gutman, A. and Wasylyk, B. (1991) Trends Genet., 7, 49-54.
- Singh, H., LeBowitz, J.H., Baldwin, A.S., Jr. and Sharp, P.A. (1988) Cell, 52, 415-423.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467.
- Huynh, T.V., Young, R.A. and Davis, R.W. (1985) In Glover, D.M. (ed.), DNA cloning—A Practical Approach. IRL Press, Oxford, Vol. I, pp.49–78.
- 21. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol., 65, 499-560.
- 22. Treisman, R. (1988) Cell, 46, 567-574.
- Severne, Y., Wieland, S., Schaffner, W. and Rusconi, S. (1988) EMBO J., 7, 2503-2508.

- Laimins, L.A., Grass, G., Pozzatti, R. and Khoury, G. (1984) J. Virol., 49, 183-189.
- 25. Graham, F.L. and Van der Eb, A.J. (1973) Virology, 52, 456-457.
- Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Mol. Cell. Biol., 2, 1044-1051.
- Schreiber, E., Matthias, P., Muller, M.M. and Schaffner, W. (1989) Nucleic Acids Res., 17, 6419.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor.
- Okayama, H., Kawaichi, M., Brownstein, M., Lee, F., Yokota, T. and Arai, K. (1987) Methods Enzymol., 154, 3-28.
- Nishikawa, T., Yamashita, T., Yamada, T., Kobayashi, H., Ohkawara, A. and Fujinaga, K. (1991) Jpn. J. Cancer Res., 82, 1340-1343.
- Gunther, C.V., Nye, J.A., Bryner, R.S. and Graves, B.J. (1990) Genes Dev., 4, 667-679.
- Boulukos, K.E., Pognonec, P., Begue, A., Galibert, F., Gesquiere, J.C., Stehelin, D. and Ghysdael, J. (1988) EMBO J., 7, 697-705.
- Watson, D.K., McWilliams, M.J., Lapis, P., Lautenberger, J.A., Schweinfest, C.W. and Papas, T.S. (1988) Proc. Natl. Acad. Sci. USA, 85, 7862-7866.
- Reddy, E.S.P., Rao, V.N. and Papas, T.S. (1987) Proc. Natl. Acad. Sci. USA, 84, 6131-6135.
- Rao, V.N., Huebner, K., Isobe, M., Ar-Rushdi, A., Croce, C.M. and Reddy, E.S.P. (1989) Science, 244, 66–70.
- LaMarco, K., Thompson, C.C., Byers, B.P., Walton, E.M. and McKnight, S.L. (1991) Science, 253, 789-792.
- Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C. and Maki, R.A. (1990) Cell, 61, 113-124.
- Xin, J-H., Cowie, A., Lachance, P. and Hassell, J.A. (1992) Genes Dev., 6, 481-496.
- Watson, D.K., McWilliams-Smith, M.J., Nunn, M.F., Duesberg, P.H., O'Brien, S.J. and Papas, T.S. (1985) Proc. Natl. Acad. Sci. USA, 82, 7294-7298.
- Seth, A., Ascione, R., Konopka, A. and Papas, T.S. (1990) Oncogenesis. Gulf Publishing Co., Houston, Tx, pp.297-311.
- 41. Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science, 241, 42-52.
- 42. Johnson, P.F. and McKnight, S.L. (1989) Annu. Rev. Biochem., 58, 799-839.
- 43. Mitchell, P.J. and Tjian, R. (1989) Science, 245, 371-378.
- Watanabe, H., Imai, T., Sharp, P.A. and Handa, H. (1988) Mol. Cell. Biol., 8, 1290-1300.