BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **243,** 307–316 (1998)
ARTICLE NO. RC978035

Identification of a Novel AP-2 Consensus DNA Binding Site

Michael S. Gee, Christopher J. Sarkisian, and Wafik S. El-Deiry[1]

Laboratory of Molecular Oncology and Cell Cycle Regulation, Howard Hughes Medical Institute, Department of Medicine and Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

Received December 22, 1997

Activator Protein (AP)-2 is a transcription factor that is required for mouse development. AP-2 activates expression of positive and negative growth regulators including erbB-2 and p21$^{WAF1/CIP1}$. Induction of p21 has been correlated with cell cycle and growth inhibition of human cancer cells. Because several endogenous AP-2 binding sites do not fit the known consensus sequences well, we sought to define AP-2's interaction with DNA more precisely. Using *C*yclic *A*mplification and *S*election of *T*argets (CAST'ing) of random oligonucleotide sequences and recombinant human AP-2 protein, we identified 17 novel AP-2 binding sites. Mobility shift assays showed significant AP-2 binding of the novel sites as compared to p21, erbB-2 and hMtIIa sites. Several sites that bound with high specificity and affinity did not fit known AP-2 consensus sequences. A sequence comparison based on several of the novel sequences yielded a putative consensus binding sequence of 5′-TAGAAAGNYCYNG-3′. These DNA binding sites may help identify novel targets of AP-2 and aid in further understanding AP-2 function. © 1998 Academic Press

The transcription factor AP-2 was originally identified as a nuclear factor regulating the expression of human metallothionein IIa (1,2). AP-2 has subsequently been found to be involved in the transcriptional regulation of many cellular genes (3-14). Analysis of the pattern of AP-2 expression in mouse embryos showed that AP-2 is involved in the development of the hindbrain, spinal cord, face and limbs (3). The spatial and temporal appearance of AP-2 in neural folds and head mesenchyme corresponds closely to that of the earliest neural crest cells. AP-2 null mutant mice die around the time of birth, exhibiting a range of congenital abnormalities including cranio-abdominoschisis, midline facial clefting, anencephaly, axial skeleton contortion and shrinkage (4,5).

The effects of AP-2 are diverse and involve a wide range of tissues. The c-erbB-2 proto-oncogene has been shown to be regulated by AP-2 in human mammary carcinoma cells (6-8). c-erbB-2 is a tyrosine kinase that is overexpressed in a significant proportion of carcinomas of several tissues, and in breast cancer is correlated with poor prognosis (6). AP-2 levels correlate well with c-erbB-2 levels in mammary carcinoma cells (7). Transfection of AP-2 into mammary cell lines activates the c-erbB-2 promoter, while transfection of a truncated form of AP-2 lacking the DNA-binding domain into c-erbB-2 overexpressing cells abolishes promoter activity (8). AP-2 also plays an important role in the differential expression of keratin genes in epidermal keratinocytes (9-11). The human epidermal keratin gene K14 is expressed during the mitotically active stage of keratinocyte differentiation, and progression through subsequent stages is accompanied by downregulation of K14 expression. AP-2 is able to activate expression of a reporter gene containing a K14 upstream sequence, and high levels of AP-2 correspond to increased expression of K14. A transient increase in AP-2 expression has been observed during retinoic acid-induced differentiation of NT2 human teratocarcinoma cells (12,13), suggesting that AP-2 may be a potential downstream mediator of retinoic acid-induced alteration in cell morphology.

Recently, AP-2 has been shown to have a negative effect on cell cycle progression by regulating the expression of the universal cell cycle inhibitor p21$^{WAF1/CIP1}$ (14). p21, via its associations with cyclin/CDK complexes, is a potent mediator of growth arrest and inhibition of DNA synthesis in mammalian cells. A 27 bp AP-2-binding sequence upstream of the p21 transcription initiation site was shown to be required for TPA-induced p21 expression. AP-2 transfection into human

[1] To whom correspondence should be addressed. Fax: (215) 573-9139.

The first two authors contributed equally to this work.

0006-291X/98 $25.00

hepatoblastoma cells induced transcription of a p21 promoter-luciferase reporter. This activation was inhibited by mutagenesis of an AP-2 binding site within the 27 bp region. Transfection of AP-2 into human colon carcinoma cells led to upregulation of endogenous p21 expression, along with a decrease in cell division and colony formation.

Several attempts have been made to define a consensus DNA-binding site for AP-2 (15,16). DNAse footprint assays using the AP-2 binding sites present in SV40, hGH, H-2Kb and hc-myc derived a preliminary consensus sequence: 5′-TCCCCANGCG-3′(15). Since then, methylation interference and missing contact probe assays performed on the AP-2 sites present in the SV40 enhancer and hMtIIa BLE yielded another consensus sequence: 5′-GCCNNNGGC-3′(16).

We chose to use a biochemical method to identify DNA sequences capable of binding AP-2. Cyclic Amplification and Selection of Targets (CAST'ing) has been used successfully to identify consensus DNA binding sequences for p53 and myogenin (17,18). This technique provides an unbiased method to detect DNA binding sequences, since it begins with a degenerate pool containing over 10^{13} oligonucleotides. Potential binding sites within the pool are isolated and enriched by repeated cycles of protein-DNA binding and PCR amplification. Previous studies have tried to extrapolate a consensus site from the limited number of AP-2 binding sites that have been identified from functional studies. The degeneracy of the sequence derived from those studies makes it difficult to identify other genes that may be under the transcriptional control of AP-2. A biochemical approach to identify AP-2 binding sequences may also serve to help elucidate the relationship between the DNA-binding ability of AP-2 and its numerous roles in gene regulation. It appears that AP-2 can serve as both a positive and a negative regulator of gene expression and cell proliferation. This duality in function may involve the ability of AP-2 to inhibit its own ability to trans-activate target gene expression. This phenomenon, known as "self-interference," has been implicated in the N-ras oncogene-mediated transformation of human teratocarcinoma cells (19). A better understanding of the ability of AP-2 to bind to DNA may shed light upon the nature of this complexity.

MATERIALS AND METHODS

Preparation of GST-AP-2 fusion protein. AP-2 was expressed as a fusion protein with glutathione-S-transferase from an expression plasmid provided by Richard Gaynor (20, 21). Briefly, 500 mL of Luria-Bertani broth containing 50 μg/mL ampicillin was inoculated with bacteria transformed with the expression vector and grown at 37°C until an OD_{600} level of 0.3. Overexpression of the fusion protein was induced by the addition of 0.5 mM IPTG into the growth media for 3 hours. The cells were then pelleted and resuspended in 10 mL of purification buffer (phosphate-buffered saline containing 1% Triton X-100). The cells were then sonicated, and the cell lysate was centrifuged for 30 minutes at 18000 rpm. The supernatant was then incubated with 500 μL of a 50% solution of glutathione-agarose beads (Sigma) in PBS for 30 minutes at 4°C. The fusion protein was then dissociated from the beads by incubating the beads in 5 mM reduced glutathione for 30 minutes at 4°C. An appropriate-sized band was visualized upon resolution of the fusion protein on a polyacrylamide gel, and protein concentration was determined using the Bio-Rad Coomassie protein assay (Bio-Rad).

CAST'ing for the AP-2 consensus DNA-binding site. CAST'ing was performed essentially as described previously, with some modifications (17,18). The target oligonucleotide consisted of 14 degenerate bases with flanking restriction sites (figure 1). A total of 60 pmoles of the degenerate oligonucleotide mix were converted to double-stranded DNA in a single PCR cycle (94°C, 1 min; 55°C, 1 min; 72°C, 60 min) using a 10-fold molar excess of primer complementary to the 3′-flanking region (5′GCGTCGACAAGCTTTCTAGA-3′). The DNA was separated from the *Taq* polymerase by incubating the PCR product in 1% SDS and 0.5 mg/mL Proteinase K for 1 hour at 48°C. The double-stranded DNA was then extracted with buffered phenol/chloroform (PC8, 3 parts phenol:2 parts chloroform:2 parts 500 mM Tris-HCl, pH 8.0, 20 mM EDTA, 10 mM NaCl) and precipitated with 3.5 volumes of EtOH, 1M ammonium acetate and 60 μg glycogen (Boehringer Mannheim).

The precipitated DNA was mixed with 20 pmoles of purified AP-2 protein fused to glutathione-S-transferase (20,21) in 50 μL of binding buffer [10 mM Tris-HCl (pH 7.9), 4.5% Ficoll 400, 60 mM KCl, 4 mM $MgCl_2$, 0.1 mM EDTA, 50 μg/mL bovine serum albumin, 0.2% Nonidet P-40] (4). Following a 30 minute incubation at 4°C, 20 μL of glutathione-agarose beads (Sigma) were added (diluted 1:1 in binding buffer), and the entire mixture was agitated for one hour at 4′C. The beads were then washed four times with 750 μL of wash buffer (100 mM NaCl, 0.1% Nonidet P-40, 0.1 mg/mL bovine serum albumin). The AP-2 binding DNA was recovered from the protein/beads complex by an SDS/Proteinase K incubation as described above, followed by a PC8 extraction and EtOH precipitation. The recovered DNA was then subjected to a standard PCR reaction (94°C, 1′; 55°C, 1′; 72°C, 2′; 15 cycles followed by a final 72°C extension for 10 minutes) in a total volume of 50 μL containing 1 μM each of primer corresponding to the sense (5′-CGCTCGAGGGATCCGAATTC-3′) and antisense (5′-GCGTCGACA AGCTTTCTAGA-3′) 5′-flanking regions. Following another SDS/Proteinase K incubation, PC8 extraction and EtOH precipitation, the PCR product was resuspended in 20 μL of binding buffer. A total of 10 μL was used as the template for the second CAST'ing cycle. For the first cycle, 40 ng/μL of poly (dI-dC) was added to the binding reaction as a non-specific competitor (Pharmacia). The poly (dI-dC) was omitted in subsequent cycles. After six cycles of CAST'ing, the final purified PCR product was digested with BamHI and HindIII and then cloned into pBluescript II-SK (Stratagene). The individual inserts were then sequenced using the Sanger dideoxy method (US Biochemicals).

EMSA. Electrophoretic mobility shift assays (EMSA) were performed as described (22). The sequences isolated by CAST'ing were ^{32}P end-labeled using polynucleotide kinase (New England Biolabs). An EMSA was then performed on each labeled oligonucleotide using 400 ng of GST-AP-2 fusion protein, 1 mg/mL of bovine serum albumin, 50 ng/μL of poly (dI-dC) and reaction buffer (20 mM HEPES, 25 mM KCl, 2 mM $MgCl_2$, 2 mM DTT, 10% glycerol, 0.2% Nonidet P-40) up to a total volume of 20 μL. The reactions were incubated on ice for 20 minutes. Rabbit anti-AP-2 polyclonal antibody (0.5 μg, Santa Cruz) was then added to some of the reactions for an additional twenty minutes to super-shift the AP-2-DNA complexes. The reaction products were then resolved on a 15% nondenaturing polyacrylamide gel, fixed, dried and autoradiographed.

RESULTS

Amplification of AP-2-specific DNA-binding sequences. We used the CAST'ing technique in an at-

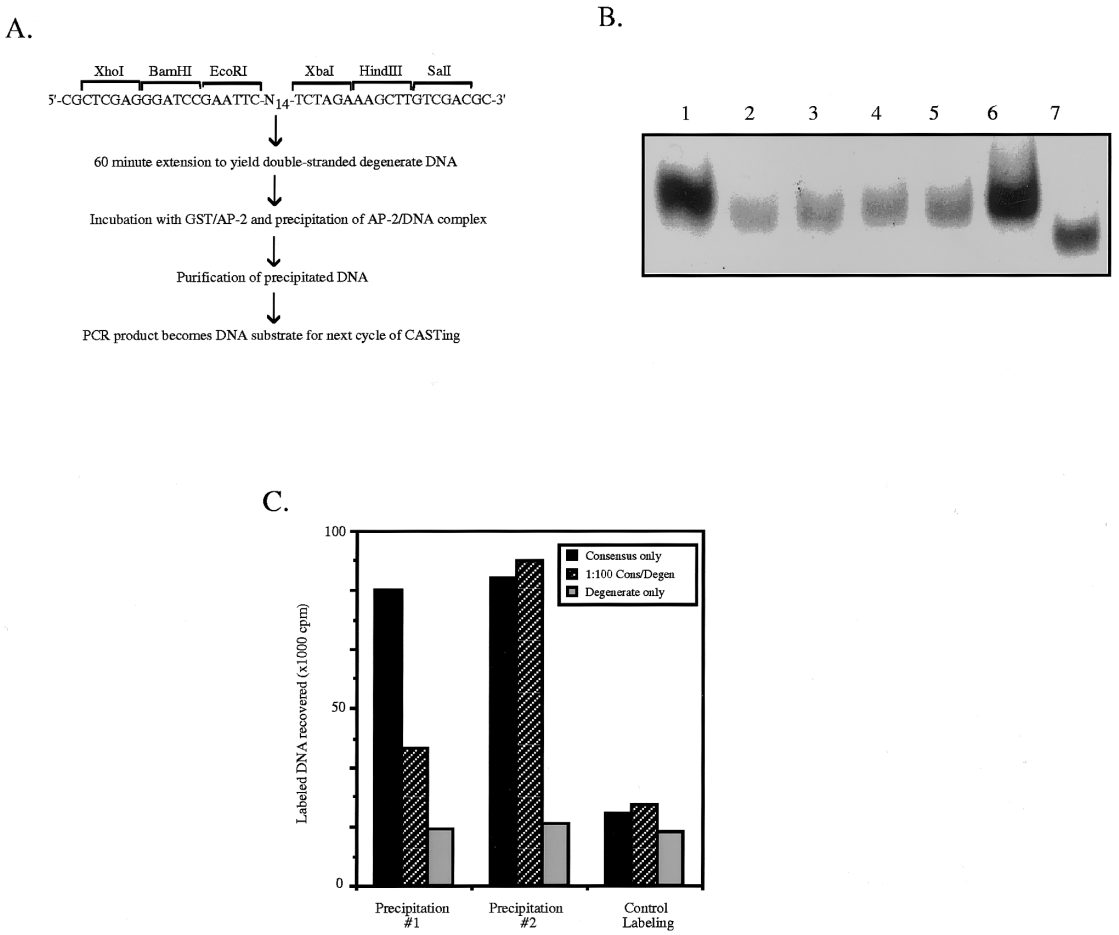

FIG. 1. Optimization of CAST'ing. CAST'ing was performed essentially as described by Wright et al. (1991) with some modifications. (A) Protocol for a single CAST'ing cycle. Six cycles of CAST'ing were used to generate AP-2 binding sequences. (B) Optimization of extension time to produce a double-stranded oligonucleotide. Single stranded degenerate oligonucleotide was annealed to primer complementary to its 3′-terminus and incubated at 72°C under standard PCR conditions. Lanes 1-5: Aliquots removed after 60′, 0′, 15′, 30′, and 45′ extension times respectively. Lanes 6-7: Double-stranded and single-stranded AP-2 consensus oligonucleotide as size comparison. All samples were ^{32}P end-labeled and resolved on a 15% non-denaturing polyacrylamide gel. (C) Selective enrichment of AP-2 binding DNA using the CAST'ing technique. Three pools of DNA, each with varying ratios of degenerate to consensus AP-2 binding DNA, were individually subjected to two cycles of CAST'ing with AP-2/GST. DNA recovered from the two cycles was ^{32}P end-labeled, and radioactivity was measured on a scintillation counter. Equal amounts of DNA from the three pools were control labeled and counted to demonstrate uniformity of labeling.

tempt to isolate AP-2-specific DNA-binding sequences. An overview of the CAST'ing protocol for a single cycle is shown in Figure 1A. We determined the optimal extension time necessary to generate a double stranded degenerate oligonucleotide (Figure 1B). There appeared to be a progressive increase in the size of the bands with longer extension times. For example, at 60 minutes, the band migrated at the same size as a known double stranded oligo of the same length (compare lane 1 with lane 6), indicating that the extension reaction was complete. Based on these results, we chose an extension time of 60 minutes.

In order to demonstrate that the CAST'ing technique could amplify AP-2 binding sequences specifically, we synthesized an "AP-2 consensus" single stranded oligonucleotide 5′CGCTCGAGGGATCCGAATTCCCG-CCCGCGGCCCGTTCTAGA AAGCTTGTCGACGC-3′) with the same length and flanking restriction sites as the degenerate oligonucleotide. The middle 14 bases corresponded to the AP-2 binding site present in the hMtIIa promoter (16). This physiologic binding site fits one of the published consensus binding sequences perfectly (16). Three pools of single stranded DNA were subjected to two cycles of CAST'ing, beginning with the initial generation of double stranded target oligonucleotide and ending with the second precipitation with AP-2. Each pool contained 60 pmoles of total DNA; the first pool consisting of entirely degenerate oligomers, the second of entirely positive control (AP-2 consensus) oligomers and the third of a 1:100 mixture of consensus

Sequence:		AP-2 Binding:
ABS.1	AAGCTTTCTAGATCCCGAAAGTGGAAGAATTCGGATCC TTCGAAAGATCTAGGGCTTTCAGGTTCTTAAGCCTAGG	+
ABS.2	GGATCCGAATTCTTAAAGCCCTGGAATTCCGCTCGAGGGATCC CCTAGGCTTAAGAATTTCGGGACCTTAAGGCGAGCTCCCTAGG	++
ABS.3	AAGCTTTCTAGAGAAGCCCTGCAGCGGGGCCTGAATTCGGATCC TTCGAAAGATCTCTTCGGGACGTCGCCCCGGACTTAAGCCTAGG	++
ABS.4 (x3)	AAGCTTTCTAGAACCTCGATCCCTACCAAACCGAATTCGGATCC TTCGAAAGATCTTGGAGCTAGGGATGGTTTGGCTTAAGCCTAGG	N.D.
ABS.5 (x2)	GGATCCGAATTCTCGTGCTGCGTGCTCTTGATCTAGAAAGCTT CCTAGGCTTAAGAGCACGACGCACGAGAACTAGATCTTTCGAA	N.D.
ABS.6	GGATCCGAATCTAAAGCCCTGGATTCCGCTCGAG CCTAGGCTTAGATTTCGGGACCTAAGGCGAGCTC	N.D.
ABS.7 (x2)	AAGCTTTCTAGACAGCGAACCCTAGCAACATGGAATTCGGATCC TTCGAAAGATCTGTCGCTTGGGATCGTTGTACCTTAAGCCTAGG	+
ABS.8	GGATCCGAATTCCTCCGGTGCGCCCTCTAGAAAGCTT CCTAGGCTTAAGGAGGCCACGCGGGAGATCTTTCGAA	++
ABS.9	GGATCCGAATTCGAGCCTGGGGACTTTCTAGAAAGCTT CCTAGGCTTAAGCTCGGACCCCTGAAAGATCTTTCGAA	++
ABS.10	GCATCCGAATTCGACCCTGGGGACTTTCTAGAAAGCTT CGTAGGCTTAAGCTGGGACCCCTGAAAGATCTTTCGAA	++
ABS.11	GGATCCGAATTCGCCCTCTAGAAACATCCTGTCTAGAAAGCTT CCTAGGCTTAAGCGGGAGATCTTTGTAGGACAGATCTTTCGAA	++
ABS.12 (x2)	AAGCTTTCTAGATAGCAAAAAACCAACATGTTGAATTCGGATCC TTCGAAAGATCTATCGTTTTTTGGTTGTACAACTTAAGCCTAGG	N.D.
ABS.13 (x2)	AAGCTTTCTAGAAAGTCCCAGGCTCGAATTCGGATCC TTCGAAAGATCTTTCAGGGTCCGAGCTTAAGCCTAGG	N.D.
ABS.14 (x2)	GGATCCGAATTCGATCGCATATTAATGACCCTTGCTCTCTAGAAAGCTT CCTAGGCTTAAGCTAGCGTATAATTACTGGGAACGAGAGATCTTTCGAA	-
ABS.15	GGATCCGAATTCCTCCGGTGCGCCCTCTAGAAAGCTT CCTAGGCTTAAGGAGGCCACGCGGGAGATCTTTCGAA	-
ABS.16	AAGCTTTCTAGATGCTTGCTGACCCTGATCCGAATTCGGATCC TTCGAAAGATCTACGAACGACTGGGACTAGGCTTAAGCCTAGG	-
ABS.17	AAGCTTTCTAGAGTTTCAACCCTACCAGCAGGGAATTCGGATCC TTCGAAAGATCTCAAAGTTGGGATGGTCGTCCCTTAAGCCTAGG	-

FIG. 2. DNA sequences obtained using the CAST'ing procedure.The DNA recovered after the sixth cycle of CAST'ing was digested with BamHI and HindIII, and cloned into pBluescriptII-KS. Suquencing was performed using the dideoxy method. Some sequences were isolated from more than one clone, and the frequency with which each sequence was encountered is listed underneath. Ability to bind AP-2 was determined by EMSA. (+) indicates binding observed, (++) indicates strong binding observed and (N.D.) indicates binding affinity was not determined.

to degenerate oligomers. Aliquots were taken from the DNA recovered after each precipitation, ^{32}P end-labeled and quantitated by liquid scintillation spectrometry. Equal aliquots were removed from each DNA pool and end-labeled prior to CAST'ing in order to demonstrate that all of the oligos were labeled uniformly (Figure 1C, column 4). The results in Figure 1C demonstrate that the CAST'ing procedure leads to the selective enrichment of the AP-2 binding sequence.

In the mixed oligonucleotide sample, AP-2 was able to bind twice as much DNA during the second precipitation compared with the first. Presumably, this was due to PCR amplification of the DNA that bound to AP-2 during the first precipitation. The question remained as to whether the precipitation was specific for the AP-2 binding sequence. This was addressed by comparing the results with those from the other two pools. The sample containing consensus DNA alone exhibited

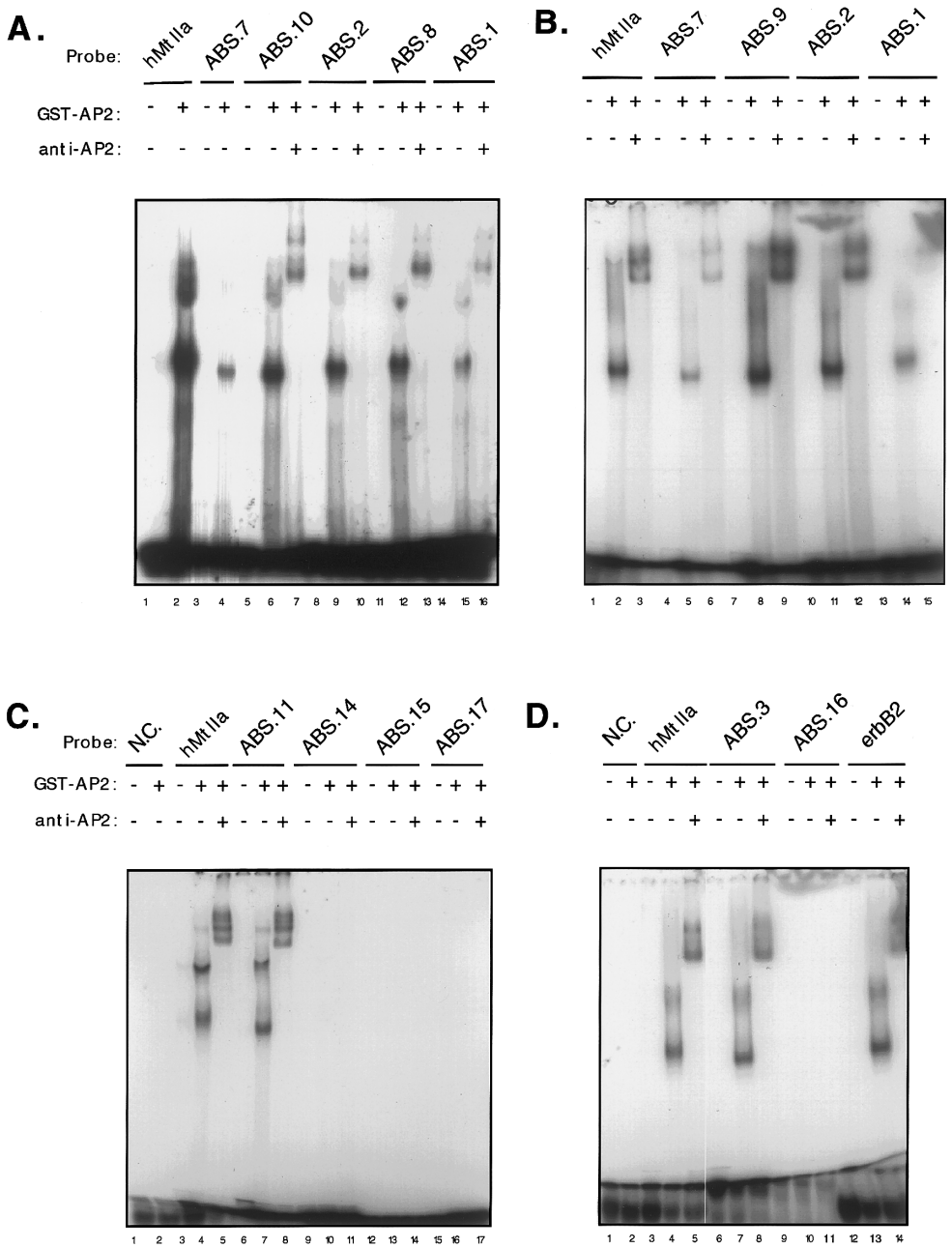

FIG. 3. Binding and specificity of AP-2 for novel oligonucleotide sequences. Electrophoretic mobility shift assays were performed as described in the Methods section, using 30,000 cpm of the specified ^{32}P labelled oligonucleotides and 400 ng GST-AP2. Where shown, the AP2-DNA complexes were supershifted with 0.5 μg rabbit polyclonal anti-AP2. These assays were performed both in the absence (A) and the presence (B-D) of 100 ng poly (dI-dC) as a nonspecific competitor for AP2 binding. The negative control (N.C., lanes 1-2 of C and D) is the p53 binding region of the p21 promoter.

strong binding to AP-2 during both cycles of precipitation, while the degenerate DNA alone sample was precipitated poorly by AP-2 during both cycles. These results taken together indicate that the CAST'ing technique led to the specific precipitation and amplification of the AP-2 binding sequence. The DNA in the mixed sample in effect simulated the oligonucleotide pool precipitated from a later CAST'ing cycle that is enriched in sequences that bind AP-2. It is interesting to note that in the consensus pool, there was no change in the amount of DNA precipitated between the two cycles. Presumably, 60 pmoles is at or above the binding capacity of the amount of AP-2 used for the precipitation. This experiment demonstrated that several cycles of

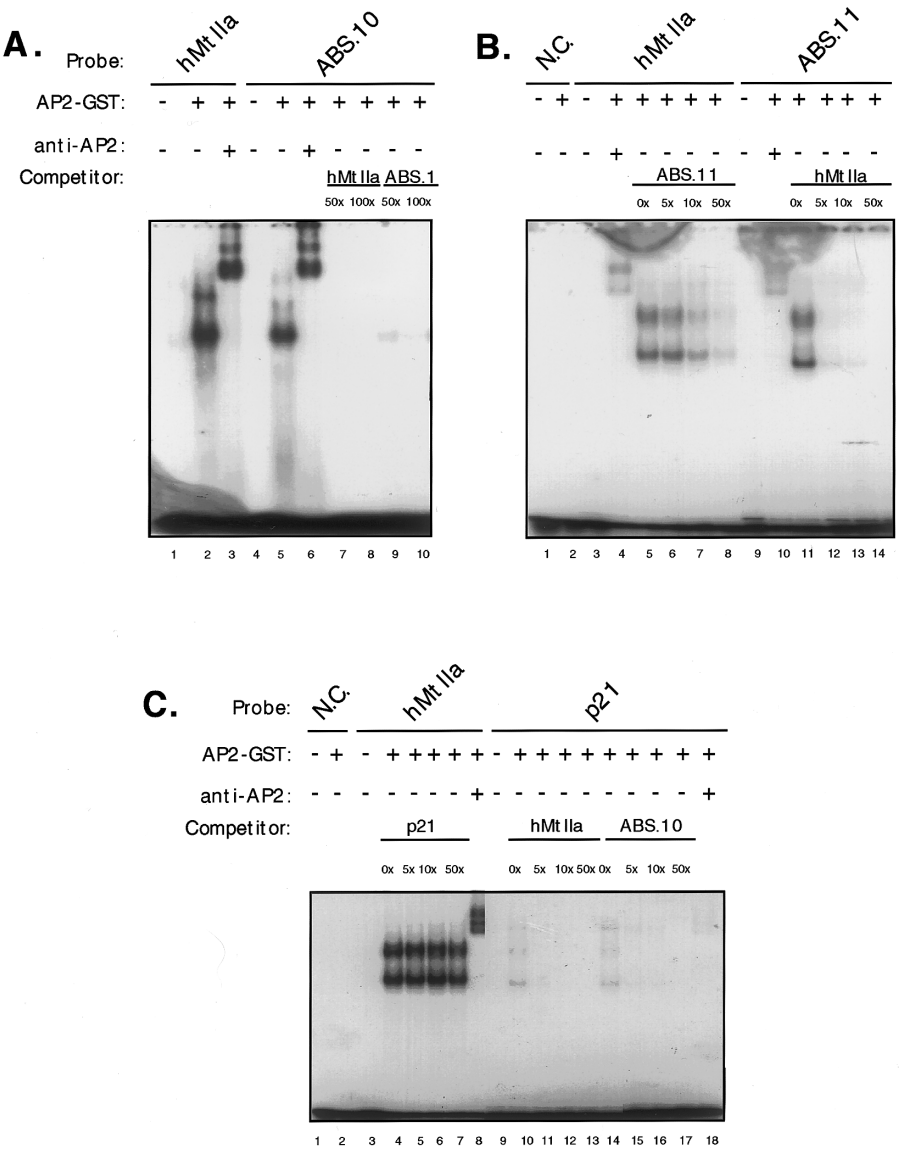

FIG. 4. Determining relative affinity of AP-2 binding to novel sequences through specific competition. Competition experiments were performed as in Figure 3, with the exception that in the denoted lanes unlabelled competitor oligonucleotides were pre-incubated with GST-AP2 prior to labelled probe addition. The designated competitor concentrations correspond as 5X = 15 ng, 10X = 30 ng, and 50X = 150 ng of unlabelled specific competitor. Supershifts with anti-AP2 were performed in the absence of specific competition. All incubations were conducted in the presence of 100 ng poly dI-dC, in addition to any specific competition.

CAST'ing would be necessary to generate a detectable enrichment of specifically bound DNA.

Isolation of novel AP-2 binding sequences. Having established the efficacy of the CAST'ing technique for AP-2, we then performed six cycles of CAST'ing on a 60 pmole pool of degenerate oligonucleotides. The DNA recovered was subcloned and sequenced by the dideoxy method, revealing that seventeen unique sequences had been recovered (Figure 2). Several sequences were present in more than one clone, indicating that they may have been amplified during the CAST'ing process. The sequences were quite diverse, and no single consensus was evident. It was interesting that no sequence contained either of the two published AP-2 consensus sequences, 5′-TCCCCANGCG-3′ (15) and 5′-GCCNNNGGC-3′ (16), in its entirety. This observation led us to characterize the ability of the novel sequences to bind AP-2.

Novel sequences bind AP-2. Several of the sequences identified via the CAST'ing protocol were examined using the mobility shift assays to determine their relative affinity for AP-2. To demonstrate that the binding of AP-2 protein to the putative AP-2 bind-

ing site (ABS) oligonucleotides was specific, EMSA's was carried out in the absence (Figure 3A) or the presence (Figure 3B-D) of 40 ng/μL poly dI-dC, a nonspecific competitor for DNA binding to AP-2. We chose to study the novel sites illustrated in Figure 3 based on an early motif observed among many of the sites surrounding a CCC sequence, although this motif has been subsequently discounted.

An oligonucleotide containing the hMtIIa sequence previously demonstrated to be bound by AP-2 was used as a positive control for sequences that strongly bound AP-2. The results indicate that the hMtIIa sequence was most strongly bound by AP-2 in the absence or presence of poly (dI-dC) (Figure 3A lanes 1-2, 3B lanes 3-5, 3C lanes 3-5, and 3D lanes 3-5). However, ABS.10 (Figure 3A lanes 5-7), ABS.9 (Figure 3B lanes 7-9), ABS.11 (Figure 3C lanes 6-8) and ABS.3 (Figure 3D lanes 6-8) were bound by AP-2 at similar levels. In contrast, other novel sequences identified through CAST'ing were not bound by AP-2 in the presence of poly (dI-dC) (ABS.14, 15, and 17, Figure 3C lanes 9-17; ABS.16 Figure 3D lanes 9-11). Finally, a promoter element of the c-erbB2 gene that has been reported to be activated by AP-2 was also analyzed for its relative affinity for AP-2. We observed that this c-erbB2 promoter region could be shifted to levels comparable to that of the hMtIIa promoter element. These results indicated that AP-2 protein could bind some novel DNA sequences that contain little homology to the defined consensus, at levels similar to physiologically relevant targets.

Relative affinities of novel binding sequences for AP-2. Having identified several sequences that were bound strongly by AP-2, we proceeded to compete these sequences with one another in order to identify the preferred DNA binding sequence of AP-2 more precisely. EMSA's of AP-2-DNA complexes were performed such that the GST-AP-2 was pre-incubated with unlabeled specific competitor oligonucleotides prior to incubation with the radiolabeled probe. Figure 4A (lanes 7-10) shows that ABS.10 binding to AP-2 was effectively competed by large amounts of the high affinity sequence hMtIIa, or poorly competed by large amounts of a weakly binding sequence, ABS.1.

Since the hMtIIa and ABS.11 sequences appeared to be two of the strongest potential AP-2 binding sequences, either of the two labeled oligonucleotides was reciprocally gel shifted with GST-AP-2 following competitive incubation with the other unlabeled oligonucleotide (Figure 4B). The results indicate that although both oligonucleotides appeared to bind AP-2 in the absence of specific competition, the hMtIIa sequence competed for AP-2 approximately ten times more efficiently than ABS.11 (compare Figure 4B lane 8 vs. 12).

We next tested the ability of the novel sites to compete with other physiological targets of AP-2. Zeng et al. identified a region in the p21 promoter necessary for

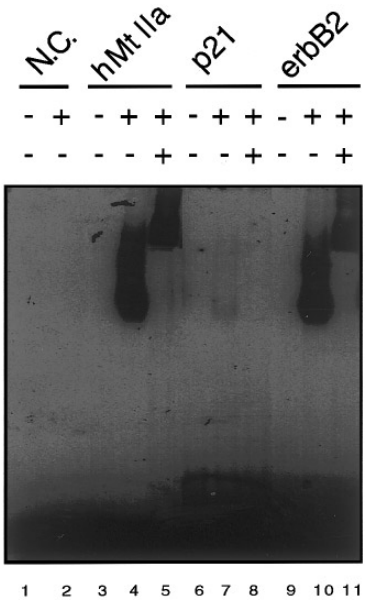

FIG. 5. Differential AP-2 binding to relevant *in vivo* targets. The EMSA's shown were performed as described in Figure 3, using 100 ng poly (dI-dC) in all incubations.

binding to and activation by AP-2 (14). By performing EMSA's we found that AP-2 protein bound to this region of the p21 promoter less efficiently than the hMtIIa sequence (Figure 4C, lane 4 vs. lane 10). It was also strongly competed by hMtIIa or by ABS.10, another novel high affinity AP-2 binding sequence. Collectively these results indicated that the novel sites were not as strongly bound by AP-2 as hMtIIa but were stronger than the p21 promoter binding region and thus fall in the range of potential physiologically relevant activation sites for AP-2.

Comparison of Physiologic Targets of AP-2. Finally, we compared directly the ability of the hMtIIa, p21, and c-erbB2 elements to bind AP-2. Gel shifts were performed using each of these targets (Figure 5). The p21 promoter element bound GST-AP-2 weakly, whereas the c-erbB2 and hMtIIa promoter regions bound GST-AP-2 at comparable levels. Because the region of the p21 promoter necessary for AP-2 activation was bound several fold less efficiently than the hMtIIa or c-erbB2 probes, it appears that the affinity of binding may not necessarily predict the activation potential of these targets.

Prediction of a novel AP-2 consensus binding site. Based on the results of EMSA assays using the novel DNA sequences isolated by CAST'ing, we performed a sequence alignment in search of similar motifs. Comparison of the two novel oligomers exhibiting the highest AP-2 binding affinity (ABS.10 and ABS.11) yielded a common sequence of 5′-TAGAAANNY-

A.

ABS.10

5′-AAGCTTTC**TAGAAAGTCCCCA**GGGTCGAATTCGGATGC-3′
3′-TTCGAAAG**ATCTTTCAGGGGT**CCCAGCTTAAGCCTACG-5′

ABS.11

5′-GGATCCGAATTCGCCCTC**TAGAAACATCCTG**TCTAGAAAGCTT-3′
3′-CCTAGGCTTAAGCGGGAG**ATCTTTGTAGGAC**AGATCTTTCGAA-5′

Consensus sequence: 5′-**TAGAAAGNYCYNG**-3′

B.

BASE #	CONSENSUS SEQUENCE	AP-2 BINDING SEQUENCES (% DISTRIBUTION)				AP-2 NON-BINDING (% DISTRIBUTION)			
		A	**C**	**G**	**T**	**A**	**C**	**G**	**T**
1		25	25	25	25	0	0	100	0
2		25	0	0	75	67	0	33	0
3		25	50	0	25	25	0	25	50
4	**T**	0	25	0	**75**	0	25	25	**50**
5	**A**	**75**	0	0	25	**0**	75	25	0
6	**G**	0	0	**75**	25	0	50	**50**	0
7	**A**	**100**	0	0	0	**25**	0	50	25
8	**A**	**100**	0	0	0	**75**	0	0	25
9	**A**	**100**	0	0	0	**50**	0	25	25
10	**G**	0	25	**75**	0	0	25	**0**	75
11	**N**	25	50	0	25	0	50	50	0
12	**Y**	0	**75**	0	**25**	25	**50**	25	**0**
13	**C**	0	**100**	0	0	0	**75**	25	0
14	**Y**	0	**50**	0	**50**	**25**	50	0	**25**
15	**N**	0	25	50	25	0	25	25	50
16	**G**	0	0	**100**	0	0	75	**0**	25
17		0	25	25	50	25	25	25	25
18		0	25	25	50	0	0	75	25
19		25	50	25	0	25	0	75	0

FIG. 6. Sequence comparison of novel oligomers. (A) A sequence comparison among ABS.10 and ABS.11 yielded a shared stretch of 10 oligonucleotides (boldfaced and underlined). (B) This shared sequence was then aligned with other sequences recovered by CAST'ing. Each sequence was designated as either an AP-2 binding (ABS.2, ABS.9, ABS.10, ABS.11) or non-binding (ABS.14, ABS.15, ABS.16, ABS.17) oligonucleotide based on EMSA results described previously. The distribution of bases is shown as a percentage that reflects all of the sequences in a given group.

CCYR-3′ (Figure 6A). We then analyzed other sequences recovered by CAST'ing for the presence of this motif (Figure 6B) and arrived at a consensus sequence of 5′-TAGAAAGNYCYNG-3′. The AAA at positions 4-6 of the consensus, the C at position 10 and the G at position 13 were present in each of the four strong AP-2 binding sites. Several elements of the consensus sequence such as the AAA and YCY motifs appear multiple times within the strongly binding AP-2 sites. This multimerization of binding motifs is consistent with the observation that AP-2 binds to DNA as a dimer (16).

DISCUSSION

The most important result of our work is the demonstration that DNA sequences that do not conform to the consensus AP-2 binding sequence are still capable of binding AP-2 with high affinity *in vitro*. None of the sequences obtained by CAST'ing contained either of the two published consensus sequences in its entirety. This is particularly surprising given the degenerate nature of the consensus sequence (it only contains six fixed nucleotides with proper spacing). Several of the recovered sequences contained either the GCC or GGC motif, but none contained both. Two of the novel sequences (ABS.10 and ABS.11) were found to bind to AP-2 with an affinity comparable to the hMtIIa site as evidenced by reciprocal competition gel shift assays (Figure 4A and 4B). The erbB2 proto-oncogene also bound AP-2 protein at levels similar to hMtIIa and the stronger novel sequences. The sequences ABS.1 and ABS.7 appear to bind AP-2 at a low to intermediate level. Several

other apparently strongly binding novel sequences were ABS.2, ABS.3, ABS.8, and ABS.9, although these have not been directly competed with other novel or physiological binding sites.

A sequence comparison of the two novel oligomers exhibiting the highest AP-2 binding affinity (ABS.10 and ABS.11) yielded a common sequence of 5′-TAGAAANNYCCYR-3′ (Figure 6A). We then analyzed other sequences recovered by CAST'ing for the presence of this motif and found that four (ABS.2, ABS.9, ABS.10 and ABS.11) of the sequences that bound AP-2 contained a motif resembling this one, while all of the oligomers without AP-2 binding capability (ABS.14, ABS.15, ABS.16 and ABS.17) did not (Figure 6B). Thus, the presence of the sequence 5′-TAGAAAGNYCYNG-3′ correlated with high AP-2 binding affinity, and therefore represents a novel consensus binding sequence for AP-2. The fact that several oligomers capable of binding AP-2 with high affinity do not contain this sequence (ABS.3, ABS.8, hMtIIa,c-erb-B2), indicates to us that the consensus sequence we derived may represent a distinct binding sequence from those previously identified. A search using BLAST and TFD revealed no significant homology with other sequences. Future experiments using mutant ABS oligomers will determine which bases within this new motif are most important for AP-2 binding. It is possible that AP-2 may bind several distinct, unrelated DNA sequences strongly. The fact that several of the novel oligonucleotides (e.g. ABS.3 and ABS.8) exhibit strong AP-2 binding activity in the absence of either the published consensus or our novel binding sequence is an indication that this may be true.

We have demonstrated that not all of the sequences identified by CAST'ing are specific for AP-2 because four of them (ABS.14, ABS.15, ABS.16, and ABS.17) did not bind AP-2 in the presence of poly (dI-dC). Future experiments such as DNase footprinting, methylation interference and missing contact probe analysis may determine the individual contributions of the constituent nucleotides to AP-2 binding. Overall, our results indicate that the novel sites fall within the range of binding affinity displayed by physiologically relevant activation sites for AP-2.

Biological interpretations of these data are difficult because the affinity of a particular DNA sequence for AP-2 is likely to represent only one factor in determining its enhancer activity via AP-2-mediated transactivation. Further studies are necessary to determine whether these new sequences have biological activity. These include testing the ability of AP-2 in cells to activate expression of a reporter gene containing the novel sequences within its promoter.

These experiments represent the first nonbiased assay to screen for AP-2 binding sequences. The discovery of novel sites using the CAST'ing technique is not surprising. Previous attempts to define the DNA-binding specificity of AP-2 have only included binding sequences with known biological activity. There are presently only a few known physiologic AP-2 targets, and these likely represent only a subset of the total. The degenerate nature of the oligomers used in the CAST'ing procedure enables it to screen all possible sequences for AP-2 binding ability. In this way, the CAST'ing technique should detect both novel physiologic targets for AP-2 as well as sequences that physically bind AP-2 but have no biological significance. Several of the sequences recovered by CAST'ing were found by EMSA not to bind to AP-2. Although poly (dI-dC) was included in the first step of CAST'ing, it may be that these sequences bound nonspecifically to AP-2 and were amplified by the subsequent CAST'ing steps lacking poly (dI-dC). For this reason, the EMSA binding buffer (see Materials and Methods) contained poly (dI-dC) to reduce non-specific protein-DNA interactions and ensure that only sequence-specific binding to AP-2 would lead to a shift in mobility.

In summary, a novel AP-2 binding consensus has been identified. This site may help to identify novel AP-2-regulated genes, as well as understand the interaction between AP-2 and DNA.

ACKNOWLEDGMENTS

We thank Richard Gaynor for the GST-AP2 plasmid and Kumaravel Somasundaram for valuable advice and critical comments. M.S.G. received support from the NIH MSTP grant and C.S. was supported by NIH Training Grant 5-T32GM7229. W.S.E.-D. is an Assistant Investigator of the Howard Hughes Medical Institute.

REFERENCES

1. Lee, W., Haslinger, A., Karin, M., and Tjian, R. (1987) *Nature* **325,** 368–372.
2. Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) *Genes & Development* **2,** 1557–1569.
3. Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W. J., and Tjian, R. (1991) *Genes & Development* **5,** 105–119.
4. Schorle, H., Maier, P., Buchert, M., Jaenisch, R., and Mitchell, P. (1996) *Nature* **381,** 235–238.
5. Zhang, J., Hagopian-Donaldson, S., Serbedzija, G., Elsemore, J., Plehn-Dujowich, D., McMahon, A. P, Flavell, R. A., and Williams, T. (1996) *Nature* **381,** 238–241.
6. Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ulrich, A., and McGuire, W. L. (1987) *Science* **235,** 177–182.
7. Hollywood, D. P., and Hurst, H. C. (1993) *EMBO* **12,** 2369–2375.
8. Bosher, J. M., Williams, T., and Hurst, H. C. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 744–747.
9. Leask, A., Byrne, C., and Fuchs, E. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 7948–7952.
10. Leask, A., Rosenberg, M., Vassar, R., and Fuchs, E. (1990) *Genes & Development* **4,** 1985–1998.
11. Byrne,C., Tainsky, M., and Fuchs, E. (1994) *Development* **120,** 2369–2383.

12. Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) *Genes & Development* **2,** 1557–1569.
13. Luscher, B., Mitchell, P., Williams, T., and Tjian, R. (1989) *Genes & Development* **3,** 1507–1517.
14. Zeng, Y. X., Somasundaram, K., and El-Deiry, W. S. (1997) *Nature Genetics* **15,** 78–82.
15. Imagawa, M., Chiu, R., and Karin, M. (1987) *Cell* **51,** 251–260.
16. Williams, T., and Tjian, R. (1991) *Genes & Development* **5,** 670–682.
17. Wright, E. W., Binder, M. and Funk, W. (1991) *Molecular and Cellular Biology* **11,** 4104–4110.
18. Funk, W. D., Pak, D. T., Karas, R. H., Wright, W. E., and Shay, J. (1992) *Molecular and Cellular Biology* **12,** 2866–2871.
19. Kannan, P., Buettner, R., Chiao, P. J., Yim, S. O., Sarkiss, M., and Tainsky, M. A. (1994) *Genes & Development* **8,** 1258–1269.
20. Smith, D. B., and Johnson, S. J. (1988) *Gene* **67,** 31–40.
21. Muchardt, C., Seeler, J. S., and Gaynor, R. B. (1992) *EMBO J.* **7,** 2573–2581.
22. Somasundaram, K., and El-Deiry, W. S. (1997) *Oncogene* **14,** 1047–1057.