

Human CCAAT-Binding Proteins Have Heterologous Subunits

Lewis A. Chodosh,[†] Albert S. Baldwin,^{*}
Richard W. Carthew,^{*‡} and Phillip A. Sharp^{*}

^{*} Center for Cancer Research
and Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

[†] Harvard Medical School
Boston, Massachusetts 02115

Summary

We have characterized three distinct proteins present in HeLa cell extracts that specifically recognize different subsets of transcriptional elements containing the pentanucleotide sequence CCAAT. One of these CCAAT-binding proteins, CP1, binds with high affinity to CCAAT elements present in the human α -globin promoter and the adenovirus major late promoter (MLP). A second protein, CP2, binds with high affinity to a CCAAT element present in the rat γ -fibrinogen promoter. Finally, the third CCAAT-binding protein is nuclear factor I (NF-I), a cellular DNA-binding protein that binds to the adenovirus origin of replication and is required for the initiation of adenoviral replication. CP1, CP2, and NF-I are distinct activities in that each binds to its own recognition site with an affinity that is at least three orders of magnitude higher than that with which it binds to the recognition sites of the other two proteins. Surprisingly, CP1, CP2, and NF-I each appear to recognize their binding site with highest affinity as a multisubunit complex composed of heterologous subunits. In the case of CP1, two different types of subunits form a stable complex in the absence of a DNA-binding site. Moreover, both subunits are present in the CP1-DNA complex. We thus propose the existence of a family of related multisubunit CCAAT-binding proteins that are composed of heterologous subunits.

Introduction

The central challenge in the study of eukaryotic gene expression is understanding how individual cells are able to stably maintain distinct levels and patterns of transcription from thousands of promoters. The rate of transcription initiation from promoters is specified by the interaction of sequence-specific DNA-binding proteins with *cis*-acting DNA control elements (reviewed in Maniatis et al., 1987). Thus, the activity of a promoter is determined both by its architecture, consisting of the type of control elements contained in the promoter and the context in which they appear, and by the regulatory state of the proteins that recognize these control elements. In view of this, a considerable degree of promoter diversity can be generated by

the combinatorial use of a relatively small number of modular transcriptional elements.

Promoter diversity is further increased by the ability of related sets of control elements to be recognized by a family of proteins that possess distinct regulatory properties. A striking example of such a family is that of the nuclear hormone receptors. Nuclear hormone receptors are inducible *trans*-acting factors that stimulate transcription by binding to specific DNA control elements (reviewed in Yamamoto, 1985). Comparison of the cDNA clones encoding the glucocorticoid, estrogen, progesterone, triiodothyronine, mineralocorticoid, retinoic acid, and vitamin D receptors has revealed conserved primary amino acid sequences in the DNA-binding domains of these receptors (Hollengerg et al., 1985; Weinberger et al., 1985; Green et al., 1986; Loosfelt et al., 1986; Weinberger et al., 1986; Arriza et al., 1987; Petkovich et al., 1987; McDonnell et al., 1987). Correspondingly, the control element sequences that these receptors specifically recognize are also conserved (Von der Ahe et al., 1986; Klock et al., 1987; Strahle et al., 1987). Thus, binding of a family of related transcription factors to a related set of control elements may result in a diverse set of regulatory consequences.

Many eukaryotic promoters possess a transcriptional regulatory element that contains the pentanucleotide sequence CCAAT between 60 and 80 bp upstream of the initiation site (Benoist et al., 1980; Efstratiadis et al., 1980; McKnight and Tjian, 1986). Several proteins that specifically recognize CCAAT elements have been described, however, the relationship between these proteins has been difficult to assess (Leegwater et al., 1985; Jones et al., 1985; Graves et al., 1986; Cohen et al., 1986; Hata-mochi et al., 1986; Jones et al., 1987; Wu et al., 1987; Dorn et al., 1987; Barberis et al., 1987). Although these various CCAAT-containing elements appear grossly similar in sequence, recent evidence suggests that different DNA-binding factors are capable of distinguishing between these elements (Dorn et al., 1987; Oikarinen et al., 1987).

In this paper, we have characterized three related but distinct proteins from human cells which specifically recognize particular subsets of CCAAT elements. Surprisingly, all three of these CCAAT-binding proteins appear to be composed of at least two different subunits.

Results

HeLa Cell Proteins Specifically Bind to Three Different CCAAT Elements

The gel electrophoresis DNA-binding assay (Fried and Crothers, 1981; Garner and Revzin, 1981; Carthew et al., 1985) was used to detect proteins in HeLa cell extracts that specifically bind to CCAAT elements present in the adenovirus major late promoter (MLP), the rat γ -fibrinogen (γ -FBG) promoter, and the consensus binding site sequence for NF-I (Figure 1). Each of the three probes containing a CCAAT element formed a specific protein-DNA

[‡] Present address: Department of Biochemistry, University of California, Berkeley, California 94720.

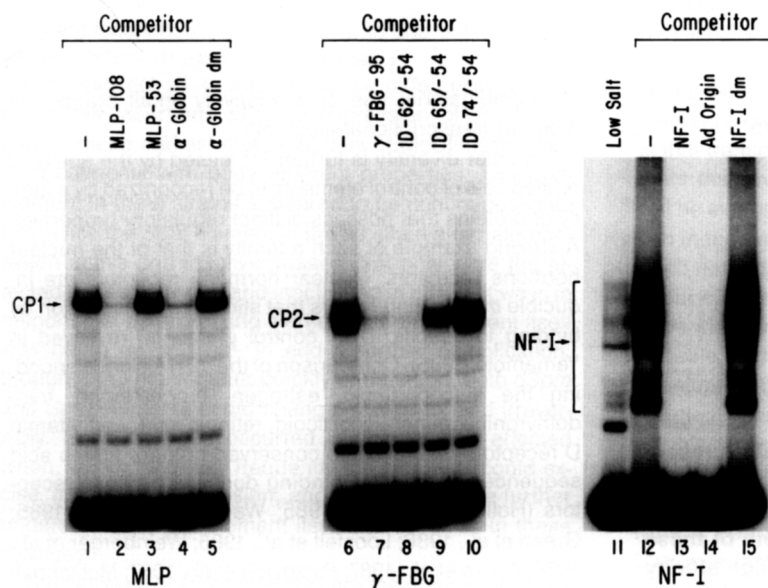


Figure 1. Specific Binding of HeLa Cell Proteins to Three CCAAT Elements
 A mixture of 12 μ g of fraction [AB] and 4 μ g of fraction [CB] was added to a standard binding reaction containing 2 μ g poly(dI-dC)-(dI-dC), and 20,000 cpm of either 0.1 ng of a radiolabeled 74 bp MLP probe (lanes 1–5) containing MLP sequences from –108 to –62, including the MLP CCAAT element located at –76, 0.2 ng of a radiolabeled 90 bp γ -FBG probe (lanes 6–10) containing sequences from –95 to –39, including the γ -FBG CCAAT element located at –65, or 0.2 ng of a 77 bp radiolabeled oligonucleotide probe containing the NF-I consensus binding site sequence (lanes 11–15), which contains sequences corresponding to +21 to +45 of the left terminus of Ad 2 with a point mutation at +28 as described in Experimental Procedures. Reactions also contained either no additional DNA (lanes 1, 6, 11, and 12), or a 50-fold molar excess of an unlabeled competitor DNA fragment (lanes 2–5, 7–10, and 13–15). Competitor fragments contained sequences from: –108 to –62 of the Ad MLP (lane 2); –53 to +33 of the Ad MLP (lane 3); –81 to –58 of the human α -globin promoter (lane 4); –81 to –58 of a mutant α -globin promoter containing a double-point mutation in the CCAAT element (lane 5); –95 to –39 of the rat γ -FBG promoter (lane 7); –95 to +35 of the rat γ -FBG promoter with internal deletion of –62 to –54 (lane 8), –65 to –54 (lane 9), or –74 to –54 (lane 10); the NF-I consensus binding site (lane 13); the adenovirus origin of replication (lane 14); or a double-point mutant of the NF-I consensus binding site, NF-I-dm (lane 15). Radiolabeled probe, unlabeled competitor fragments, and poly(dI-dC)-(dI-dC) were mixed prior to the addition of protein fractions [AB] and [CB]. Final reaction conditions were 60 mM KCl, 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 4 mM Tris-HCl (pH 7.9), 1 mM EDTA, and 1 mM dithiothreitol for MLP binding reactions (lanes 1–5), for γ -FBG binding reactions (lanes 6–10), and for the low salt NF-I binding reaction (lane 11). Final reaction conditions were 150 mM KCl, 5 mM MgCl₂, 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 4 mM Tris-HCl (pH 7.9), 1 mM EDTA, and 1 mM dithiothreitol for the remaining NF-I binding reactions (lanes 12–15). Binding reactions, extract preparation, probe preparation, chromatography, and the gel electrophoresis DNA binding assay were as described in Experimental Procedures. Arrows indicate the position of specific protein–DNA complexes formed by CP1, CP2, and NF-I.

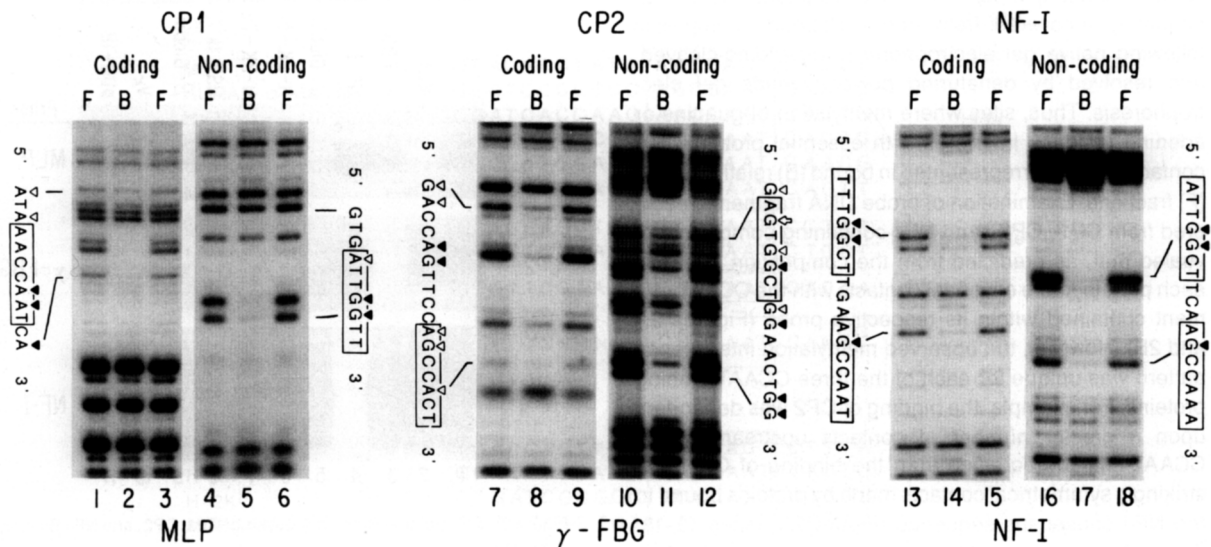
complex when mixed with partially purified chromatographic fractions from HeLa whole-cell extracts (WCE) (Figure 1, lanes 1, 6, and 12). An identical pattern of complexes was generated when unfractionated WCE was used (data not shown). A protein present in HeLa cell extracts, designated CP1, specifically recognized sequences present between –108 and –62 bp relative to the MLP initiation site (Figure 1, lanes 1–3). A CCAAT pentanucleotide sequence is present in the MLP at –76 (Figure 2B). An oligonucleotide containing the human α -globin CCAAT element competed efficiently for the binding of CP1 to the MLP probe (Figure 1, lane 4). A double-point mutant in which the α -globin CCAAT sequence was mutated to CGAGT (Figure 2B) failed to compete for CP1 binding to the MLP probe (Figure 1, lane 5). This suggests that CP1 recognizes sequences within the CCAAT pentanucleotide of the MLP.

Similarly, a protein present in HeLa cell extracts, designated CP2, specifically recognized sequences between –95 and –39 of the rat γ -FBG promoter (Figure 1, lanes 1 and 2). A CCAAT-related sequence, AGCCACT, is present in the γ -FBG promoter at –65 (Figure 2B). An internal deletion mutant (ID–62/–54) of the promoter which retains only the sequence AGCCA competed efficiently for binding of CP2 to the wild-type γ -FBG probe (Figure 1, lane 8). Another mutant (ID–65/–54) with an internal deletion that is 3 bp larger (Figure 2B), and thus missing the CCA trinucleotide of the γ -FBG CCAAT-like element, com-

peted poorly for CP2 binding (Figure 1, lane 9). This suggests that CP2 recognizes sequences within the CCAAT-like element of the γ -FBG promoter. A mutant (ID–74/–54) with a larger internal deletion encompassing the entire CCAAT-like element and additional upstream sequences was unable to compete for CP2 binding (Figure 1, lane 10). This indicates that CP2 makes additional important contacts upstream of the CCAAT-like element. A point mutant of the γ -FBG promoter that contained a canonical CCAAT pentanucleotide sequence competed for CP2 binding to the wild-type γ -FBG probe as efficiently as the wild-type promoter (data not shown). Thus, CP2 specifically recognizes CCAAT sequence elements.

In contrast to the MLP and γ -FBG promoter fragments, specific binding of HeLa cell proteins to an oligonucleotide probe containing the NF-I consensus binding site sequence, TTTGGCN₅GCCAAT (Leegwater et al., 1986; De Vries et al., 1987; Rosenfeld et al., 1987), was difficult to detect using the same reaction conditions that optimize the specific binding of CP1 and CP2 (Figure 1, lane 11). However, addition of MgCl₂ to 5 mM and KCl to 150 mM resulted in a striking stimulation of specific binding activity to the NF-I probe (Figure 1, compare lanes 11 and 12). These conditions are known to increase NF-I binding activity (Rosenfeld and Kelly, 1986; Jones et al., 1987; De Vries et al., 1987). Moreover, the unusual pattern of protein–DNA complexes generated under these conditions was similar to that previously observed for affinity-purified

A



B

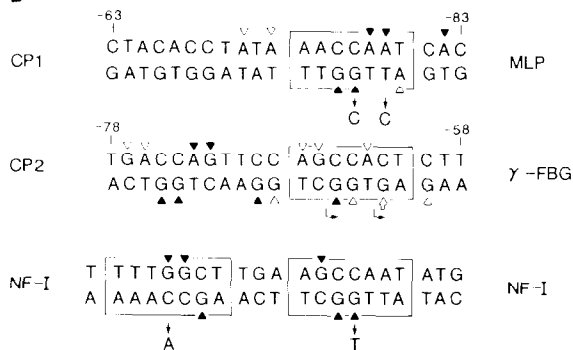


Figure 2. Methylation Interference Analysis of CP1, CP2, and NF-I Binding Activities

(A) Methylation Interference Analysis. Methylation interference was performed as described in Experimental Procedures. 200,000 cpm of a 3' end-labeled probe, containing either MLP sequences from -108 to -62, γ -FBG sequences from -95 to -39, or the NF-I consensus binding site sequence was partially methylated with dimethylsulfate, purified, and added to a scaled-up CP1, CP2, or NF-I binding reaction containing 10 μ g of poly(dI-dC)-(dI-dC), and either no protein (lanes 1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, and 18), or a mixture of 180 μ g of fraction [AB] and 58 μ g of [CB]. Bound, (B), and free, (F), probes were separated by native gel electrophoresis, visualized by autoradiography, eluted, purified, and cleaved with piperidine. An equal number of counts of free and bound DNA were electrophoresed on denaturing polyacrylamide gels. Coding and noncoding sequences are displayed along with a summary of the effects of partial methylation on the binding

of CP1, CP2, and NF-I to their recognition sites. A box surrounds the heptanucleotide sequence corresponding either to the extended CCAAT consensus sequence, RRCCAAT, or its complement. Residues at which methylation interfered with the binding of a protein to its recognition sequence are indicated by triangles. Solid triangles represent a greater degree of interference than hollow triangles. A residue at which methylation enhanced binding is indicated by a hollow arrow.

(B) Summary of Methylation Interference Analysis and Gel Electrophoresis DNA-Binding Analysis. Each of the three promoter sequences MLP, γ -FBG, and the NF-I consensus binding site sequence, recognized by CP1, CP2, and NF-I, respectively, are displayed along with a summary of the methylation interference analysis as performed in (A). Straight, solid arrows represent double-point mutations that abolish the binding of CP1 and NF-I to their respective recognition sites as performed in Figure 1. Bent arrows represent the 5' internal deletion endpoints of two γ -FBG mutants that define essential contacts within the γ -FBG CCAAT element, as analyzed in Figure 1.

NF-I (Rosenfeld et al., 1987; Jones et al., 1987). As has been noted, this pattern probably reflects heterogeneity in the group of polypeptides forming these complexes since affinity-purified NF-I is similarly heterogeneous following SDS-polyacrylamide gel electrophoresis (Rosenfeld and Kelly, 1986; Jones et al., 1987). Binding to the NF-I consensus sequence probe was specifically competed by the homologous consensus sequence as well as by a DNA fragment containing the high-affinity NF-I binding site present in the adenovirus origin of replication (Figure 1, lanes 13 and 14; Nagata et al., 1983). Mutation of a single base in each GCCAA half site of the NF-I consensus sequence (Figure 2B) produced a sequence (NF-I-dm) that did not compete for the binding of NF-I polypeptides (Fig-

ure 1, lane 15). Therefore, this entire set of proteins must make essential contacts with at least one CCAAT element in this probe. This double-point mutation has been shown to severely decrease NF-I binding affinity (Rosenfeld et al., 1987). Based upon the characteristic pattern of protein-DNA complexes formed on the NF-I probe, as well as the specific competition of these complexes by wild-type but not mutant sites, we conclude that the group of proteins in these complexes most likely corresponds to the previously identified NF-I.

In order to assess more closely the DNA contacts made by each of the CCAAT-binding proteins described above, a methylation interference experiment was performed (Siebenlist et al., 1980). DNA probes were partially methyl-

ated by reaction with dimethylsulfate, and used as substrate in binding reactions. Probe fragments were subsequently recovered from free and bound complexes following native gel electrophoresis, piperidine-cleaved, and resolved by denaturing polyacrylamide gel electrophoresis. Thus, sites where methylation of guanine or adenine residues interferes with essential protein-DNA contacts are underrepresented in bound (B) relative to free (F) fractions. Examination of probe DNA fragments recovered from CP1-, CP2-, and NF-I-containing complexes revealed that, as predicted from the competition analysis, each protein made essential contacts with the CCAAT element contained within its respective probe (Figures 2A and 2B). However, the observed methylation interference pattern was unique for each of the three CCAAT-binding proteins. For example, the binding of CP2 was dependent upon a greater number of contacts upstream of the CCAAT-related sequence than the binding of CP1. The strikingly symmetrical contacts made by proteins bound to the NF-I consensus sequence (Figure 2A, lanes 13-18; Figure 2B) are identical to those previously described for the binding of purified NF-I (De Vries et al., 1987).

In summary, the methylation interference analysis confirms that CP1, CP2, and NF-I are CCAAT-binding proteins in that each protein makes essential contacts within the CCAAT pentanucleotide. In addition, variation in the patterns of methylation interference, as well as differences in the mobility of the protein-DNA complexes, suggest that CP1, CP2, and NF-I are distinct proteins.

CP1, CP2, and NF-I Are Distinct CCAAT-Binding Proteins

In order to examine the relationship between the specificities of proteins binding to CCAAT elements in a variety of promoters, a competition analysis was performed in which a radiolabeled probe was incubated in the presence of a 100-fold molar excess of an unlabeled DNA fragment. Each of the CCAAT elements examined has been shown to contribute to the transcriptional activation of the promoter from which it is derived, except for the V_{κ} CCAAT element (L. Chodosh and R. Carthew, unpublished results; Morgan et al., 1988; Chodosh et al., 1987; Miksicek et al., 1987; Mellon et al., 1981; A. Baldwin, unpublished results; Jones et al., 1985; Graves et al., 1986; Charnay et al., 1985; Myers et al., 1986). The results of these competition experiments are shown in Figure 3, in which panels representing the bound complexes generated on three different probes are mounted beneath each other. Thus, the ability of a particular DNA fragment to compete for complexes formed on each of the three probes can be directly compared. The MLP fragment competed for the binding of CP1 to the MLP probe, and had no effect upon the binding of CP2 to the γ -FBG probe or upon the binding of NF-I to the NF-I consensus sequence probe (Figure 3, lanes 1 and 2). Similarly, the γ -FBG promoter fragment competed for the binding of CP2 to the homologous probe, and had no effect upon the binding of CP1 to the MLP probe or upon the binding of NF-I to the NF-I consensus sequence probe (Figure 3, lane 3). Finally, the NF-I consensus sequence fragment competed for the binding

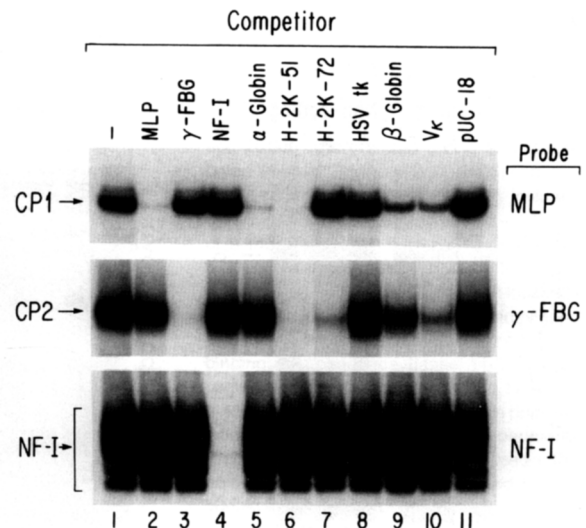


Figure 3. Competition Binding Analysis of CP1, CP2, and NF-I Binding Activities

A mixture of 12 μ g of HeLa fraction [AB] and 4 μ g of HeLa fraction [CB] (lanes 1-11) was added to a standard binding reaction for either CP1 (top panel), CP2 (middle panel), or NF-I (bottom panel) containing a mixture of the indicated probe, 2 μ g poly(dI-dC)-(dI-dC), and either no competitor DNA (lane 1) or a 100-fold molar excess of an unlabeled competitor DNA fragment (lanes 2-11). Competitor fragments contained sequences from: -108 to -62 of the Ad MLP (lane 2); -95 to -39 of the rat γ -FBG promoter (lane 3); the consensus binding site of NF-I, corresponding to +21 to +45 of the left terminus of Ad 2 with a point mutation at +28 (lane 4); -81 to -58 of the human α -globin promoter (lane 5); -65 to +3 of the mouse H-2K^b promoter (lane 6); -90 to -64 of the mouse H-2K^b promoter (lane 7); -94 to -70 of the HSV thymidine kinase promoter (lane 8); -344 to -15 of the mouse β -globin promoter (lane 9); -297 to -47 of the MOPC-41 V_{κ} promoter (lane 10); or the 50 bp polylinker of pUC18 (lane 11). Binding reactions were analyzed by native gel electrophoresis as above. Following autoradiography, panels corresponding to the bound complexes formed on each of the three probes were excised and mounted together for direct comparison. Arrows indicate the position of specific protein-DNA complexes formed by CP1, CP2, and NF-I.

of NF-I to the consensus sequence probe, and had no effect upon the binding of CP1 to the MLP probe or upon the binding of CP2 to the γ -FBG probe (Figure 3, lane 4). This suggests that CP1, CP2, and NF-I are distinct CCAAT-binding proteins with dramatically different affinities for these three binding sites.

In support of the conclusion that CP1, CP2, and NF-I possess distinct binding specificities, each CCAAT-binding protein displayed a characteristic pattern of binding of a panel of unlabeled competitor fragments containing different CCAAT elements (Figure 3; Table 1). For example, CP1 bound with high affinity to fragments containing the MLP, the human α -globin CCAAT element, and the mouse class I major histocompatibility gene H-2K^b -51 CCAAT element, and with somewhat lower affinity to fragments containing the murine β -globin promoter and the MOPC-41 V_{κ} promoter (Figure 3, lanes 2, 5, 6, 9, and 10). CP1 also bound with high affinity to the murine sarcoma virus (MSV) LTR and the human heat shock protein 70 (hsp70) promoter, and with low affinity to the herpes

Table 1. Sequence Comparison of CP1, CP2, and NF-I Binding Sites

CP1					Affinity	
MLP	(-76)	CTACACCTATA	AACCAAT	CACCT	High	
HUMAN α -GLOBIN	(-70)	GGCTCCGCGCC	AGCCAAT	GAGCG		
H-2K ^b	(-51)	CGATCACCAAG	AACCAAT	CAGTG		
HSP70 ¹	(-70)	TTTCCCTTCTG	AGCCAAT	CACCG		
UAS2UP1-CYC1	(-204)	ACCGCTTGATC	CACCAAT	CAACG		
MSV LTR	(-81)	TTATTTGAACT	AACCAAT	CAGTT	Moderate	
MOPC-41 V _K	(-89)	ACAGCAGCTGT	GACCAAT	CCTAA		
MOUSE β -GLOBIN	(-77)	ACTCTGGTAAG	GGCCAAT	CTGCT		
HSV tk	(-82)	GTGTTCGAATT	CGCCAAT	GACAA	Low	
CP1 Consensus		^C _T NNNNNN	^A _{GG} CCAAT	CAN ^C _{TT}		
CP2						
FBG	(-65)	TGACCAGTTCC	AGCCAAT	CTTTA	High	
H-2K ^b	(-51)	CGATCACCAAG	AACCAAT	CAGTG		
H-2K ^b	(-72)	GTGATCGCGCC	ACCCAAT	GGGGG		
MOPC-41 V _K	(-89)	CTTGCAGCTGT	GACCAAT	CCTAA		
MOUSE β -GLOBIN	(-77)	ACCCTGGTAAG	GGCCAAT	CTGCT	Moderate	
CP2 Consensus		^C _T AG ^C _T NNN	^A _G RCCAAT	CNNNR		
NF-I						
NF-I		TTTGGCT	TGA	AGCCAAT	ATGAG	High
Ad ORIGIN		TTTGGAT	TGA	AGCCAAT	ATGAG	
NF-I Consensus ²		NTTGGCN	NNN	NGCCAAN		

The sequences of a number of CCAAT elements referred to in the text were aligned according to the best fit to the boxed CCAAT-element motif, RRCCAAT, and grouped by relative affinity for CP1, CP2, and NF-I, as measured by the gel electrophoresis DNA-binding assay. High, moderate, and low affinities differ by a factor of at least five. The number listed immediately to the left of each sequence indicates the position of the most extreme C of the CCAAT pentanucleotide relative to the transcription initiation site. A consensus binding site sequence is presented for each of the three CCAAT-binding proteins. The consensus binding site sequences for CP1 and CP2 were derived from an affinity-weighted analysis of each of the binding sites listed. Consensus sequence positions which have one nucleotide listed above another (e.g., ^A_T) indicate a preference for the nucleotide listed on top. R indicates a similar preference for A or G.

¹ M. Pelletier and R. Kingston, personal communication and L. Chodosh, data not shown.

² The NF-I consensus binding site sequence is derived from Rosenfeld et al. (1987) and De Vries et al. (1987).

simplex virus thymidine kinase (HSV tk) CCAAT element (data not shown; Figure 3, lane 8; M. Pelletier and R. Kingston, personal communication). Like CP1, CP2 recognized a fragment containing the H-2K^b -51 CCAAT element with high affinity (Figure 3, lane 6). Unlike CP1, CP2 also recognized fragments containing the H-2K^b -72 CCAAT element with high affinity, and did not bind to a fragment containing the α -globin CCAAT element (Figure 3, lanes 5 and 7). In marked contrast to both CP1 and CP2, NF-I did not bind to any other CCAAT-containing fragment tested (Figure 3, lanes 2-11). In particular, the α -globin promoter failed to even partially compete at a 100-fold molar excess with any complex characteristic of NF-I binding (Figure 3, lane 5). Thus, the group of proteins which we have characterized as NF-I recognize the α -globin CCAAT element with an affinity that is at least three orders of magnitude lower than the affinity with which they recognize the NF-I consensus binding site. This was surprising since NF-I had previously been reported to bind to the α -globin

CCAAT element with an affinity that was only five times lower than that with which NF-I bound to the adenovirus origin of replication (Jones et al., 1987). In order to test for weak binding to the α -globin promoter by NF-I, a probe containing the α -globin CCAAT element was radiolabeled and used as substrate in a binding reaction under conditions optimal for NF-I binding. Complexes characteristic of NF-I binding, either by electrophoretic mobility or specific competition, were not detected. Complexes characteristic of NF-I binding were also not detected under the low salt conditions optimal for CP1 binding to the α -globin promoter (data not shown).

Competition analysis using extremely high molar ratios of competitor to probe demonstrated that each of the three CCAAT-binding proteins recognized binding sites for the other two proteins with a higher affinity than that with which it recognized non-CCAAT-element containing control fragments of comparable length (data not shown). This was not unexpected, since the core sequences of

each of these CCAAT-containing elements are quite similar. The distinct binding specificities of the three CCAAT-binding proteins must therefore reflect contacts outside of the CCAAT pentanucleotide sequence. Comparison of the methylation interference contacts of CP1, CP2, and NF-I, as well as the consensus binding site sequences for each of these three proteins, supports this conclusion (Figures 2A and 2B; Table 1). High-affinity CP1 binding shows a strong preference for a CA dinucleotide immediately 3' to the CCAAT pentanucleotide. In contrast, high-affinity CP2 binding shows a strong preference for sequences 5' to the CCAAT pentanucleotide. Interestingly, despite their differences, the consensus binding site sequences of CP1 and CP2 do share similarities at several positions outside of the CCAAT pentanucleotide sequence. Unlike both CP1 and CP2, NF-I clearly binds with high affinity to two symmetrically disposed CCAAT elements.

The competition binding patterns of CP1, CP2, and NF-I revealed examples of CCAAT elements that bound only one of the three proteins with high affinity (MLP, γ -FBG, NF-I, α -globin, H-2K^b -72), as well as examples of CCAAT elements that bound two CCAAT proteins with moderate to high affinity (H-2K^b -51, MOPC-41 V_K, β -globin). The HSV tk promoter bound CP1 with low affinity, and did not detectably bind to either CP2 or NF-I. Analysis of the binding of HeLa cell proteins to a labeled fragment containing the HSV tk CCAAT element confirmed the specific binding of CP1, and also revealed the presence of an additional rapidly migrating specific complex of comparable abundance to that formed by CP1. Generation of this complex may be due to the previously described CCAAT-binding protein, CBP (data not shown). The H-2K^b promoter presented an interesting example in which one CCAAT element (-51) bound both CP1 and CP2 with high affinity, while an adjacent CCAAT element (-72) selectively bound only CP2 with high affinity (Figure 3, lanes 6 and 7). Deletion of sequences containing either one of these CCAAT elements resulted in a 10-fold reduction in transcription *in vivo* (A. Baldwin, unpublished results). However, on the basis of this data alone, it is unclear whether the loss of transcription activity resulted from the loss of binding of a particular CCAAT-binding factor.

The binding specificities of CP1, CP2, and NF-I were reproducibly observed in multiple nuclear extracts prepared from HeLa cells. The same three binding specificities were observed in HeLa whole-cell extracts. Thus the binding specificities defining CP1, CP2, and NF-I are not dependent upon the method of extract preparation. It seems highly unlikely that these three binding specificities result from the *in vitro* modification of a single progenitor protein. Such a modification would have to occur rapidly and precisely in multiple extracts to alter extensively the binding specificity of the protein. Protein modifications of this kind have not been described previously.

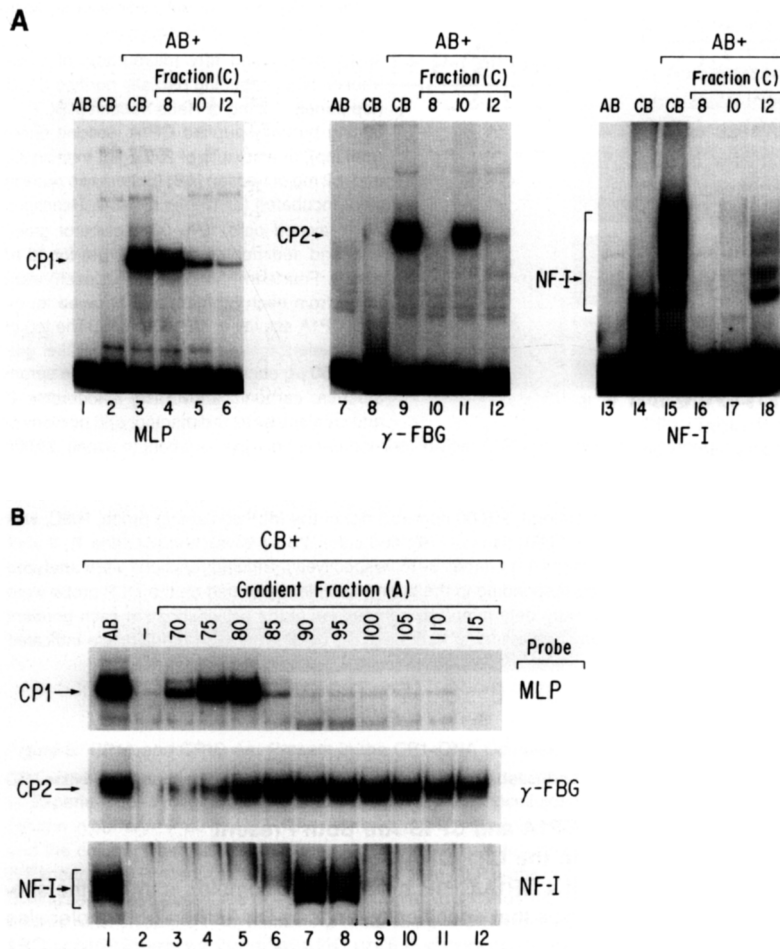
Optimal Binding of CP1, CP2, and NF-I Requires Two Chromatographic Fractions

Since HeLa cell extracts had been shown to contain at least three distinct CCAAT-binding proteins by the above analysis, we wished to determine if these activities could

be separated chromatographically. Surprisingly, resolution of HeLa whole cell extracts into four fractions by phosphocellulose chromatography resulted in an almost complete loss of binding activity for each of the three CCAAT-binding proteins (data not shown). Pair-wise combination of the four phosphocellulose fractions revealed that addition of the 0.04 M KCl flowthrough fraction A to the 0.6 M fraction C resulted in a striking restoration of binding activity for each of the three CCAAT-binding proteins (data not shown). A similar result was observed following chromatography of a murine erythroleukemia cell extract on heparin-agarose (data not shown). For each of the CCAAT-binding proteins, the activities present in fraction A and in fraction C were followed through an additional chromatographic step. The activity present in fraction A bound to a DEAE column and was eluted by a 0.15 to 0.35 M KCl step (fraction AB; see Experimental Procedures). Similarly, the activity present in fraction C bound to a single-stranded DNA cellulose column and was eluted by a 0.1 to 0.3 M KCl step (fraction CB; see Experimental Procedures). Although neither fraction AB nor fraction CB exhibited significant specific DNA-binding activity alone, combination of AB and CB resulted in a dramatic increase in the specific binding activity of CP1, CP2, and NF-I (Figure 4A, lanes 1-3, 7-9, and 13-15). Combining fractions AB and CB resulted in a 30- to 50-fold stimulation of CP1 binding activity, and in a 5- to 10-fold stimulation of CP2 and NF-I binding activities. No other combination of chromatographic fractions resulted in a significant stimulation of binding (data not shown). Trypsin treatment demonstrated that the active component in each fraction was a protein (data not shown). Binding was not stimulated by the addition of unrelated proteins such as bovine serum albumin to either fraction (data not shown). Thus, fraction AB contained binding activities CP1A, CP2A, and NF-IA. Similarly, fraction CB contained binding activities CP1B, CP2B, and NF-IB.

CP1, CP2, and NF-I Do Not Share Common Subunit Binding Activities

The surprisingly similar chromatographic behavior of the subunit activities of CP1, CP2, and NF-I raised the possibility that the three CCAAT-binding proteins might have common subunits. This possibility was examined by eluting CP1A, CP2A, and NF-IA activities from a DEAE column, and CP1B, CP2B, and NF-IB activities from a single-stranded DNA-cellulose column, with linear salt gradients. This provides better resolution than the discontinuous steps used to generate the original fractions AB and CB. Chromatography of phosphocellulose fraction C on single-stranded DNA-cellulose resulted in the clear separation of each of the three component activities (Figure 4A). CP1B activity eluted primarily in fraction 8, CP2B activity eluted almost exclusively in fraction 10, and NF-IB activity eluted primarily in fraction 12 (Figure 4A, lanes 4, 11, and 18). Therefore, CP1B, CP2B, and NF-IB each have distinct chromatographic behavior. Similarly, chromatography of phosphocellulose fraction A on DEAE-Sephacel revealed that while CP1A activity peaked in fraction 80, CP2A exhibited a broad elution profile peaking in fraction



85, and NF-IA activity peaked in fraction 90 (Figure 4B, lanes 5–7). Clearly, CP1A activity chromatographs in a distinct fashion from CP2A and NF-IA. The relationship between the chromatographic behavior of CP2A and NF-IA activities, however, is less clear. The observation that CP2A and NF-IA activities peak in separate fractions, and that the CP2A activity peak is quite broad while the NF-IA peak is relatively narrow, suggests that CP2A and NF-IA activities are distinct. This hypothesis is further supported by competition experiments performed in the presence of limiting amounts of fraction AB (data not shown). Under these conditions, a 100-fold molar excess of an unlabeled γ -FBG promoter fragment failed to compete for NF-I binding. Similarly, the NF-I consensus binding site fragment failed to compete for CP2 binding.

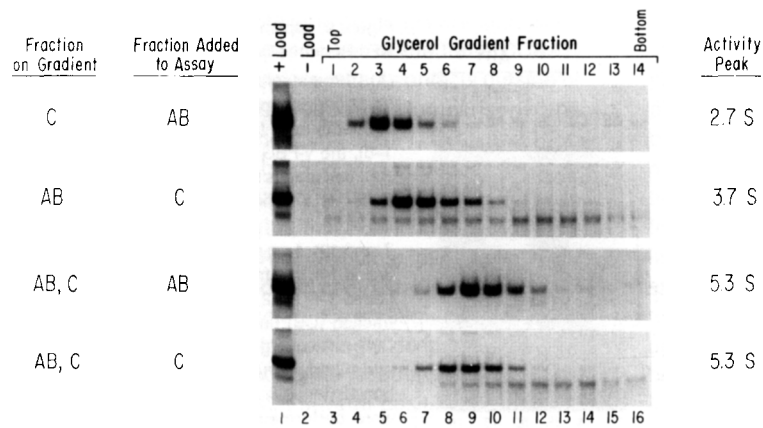
CP1A and CP1B Form a Stable Complex in the Absence of a DNA-Binding Site

The observation that maximum binding activity of each of three CCAAT factors required the addition of two chromatographic fractions suggested that at least two distinct subunits might be present in the CCAAT factor–DNA com-

plex. Thus, if these subunits form a stable complex in HeLa cell extracts, but are separated by cation-exchange chromatography, a binding activity present in unfractionated extracts would appear absent in chromatographic fractions. When the appropriate fractions containing the two subunits were combined, they could reassociate and thereby regenerate the original binding activity. In order to test this possibility, partially purified fractions containing CP1A alone, CP1B alone, or a combination of CP1A and CP1B were sedimented through glycerol gradients and the resulting fractions assayed for CP1A and CP1B binding activities (Figure 5). CP1B activity derived from phosphocellulose fraction C peaked in gradient fraction 3, corresponding to a Svedberg coefficient of approximately 2.7 S (Figure 5, top panel, lane 5). CP1A activity derived from DEAE fraction AB peaked in gradient fraction 4, corresponding to a Svedberg coefficient of 3.7 S (Figure 5, second panel from top, lane 6). When fractions AB and C were mixed and preincubated prior to loading on the glycerol gradient, both CP1A and CP1B activities sedimented considerably more rapidly and both activities peaked in fraction 7 (Figure 5, bottom two panels, lane 9).

Figure 4. Optimal Binding of CP1, CP2, and NF-I Requires Two Chromatographic Fractions (A) Chromatography of CP1B, CP2B, and NF-IB. 20,000 cpm of a radiolabeled probe containing sequences from either the Ad MLP (lanes 1–6), the γ -FBG promoter (lanes 7–12), or the NF-I consensus binding site (lanes 13–18) was added to a standard binding reaction containing either 12 μ g of fraction [AB] alone (lanes 1, 7, and 13), 4 μ g of fraction [CB] alone (lanes 2, 8, and 14), or a mixture of 12 μ g of fraction [AB] (lanes 3–6, 9–12, and 15–18) and either 4 μ g of fraction [CB] (lanes 3, 9, and 15), or 4 μ g of DNA-cellulose gradient fraction 8 (lanes 4, 10, and 16), fraction 10 (lanes 5, 11, and 17), or fraction 12 (lanes 6, 12, and 18). Binding reactions were performed and analyzed by native gel electrophoresis as above. Chromatography of HeLa fraction [C] over DNA-cellulose was performed as in Experimental Procedures. Arrows indicate specific protein–DNA complexes formed by CP1, CP2, and NF-I. (B) Chromatography of CP1A, CP2A, and NF-IA. 20,000 cpm of a radiolabeled probe containing sequences from either the Ad MLP (top panel), the γ -FBG promoter (middle panel), or the NF-I consensus sequence (bottom panel) was added to a standard binding reaction containing 4 μ g of HeLa fraction [CB] (lanes 1–12), and either 12 μ g of HeLa fraction [AB] (lane 1), 1 μ l of chromatography buffer (lane 2), or 1 μ l of the indicated chromatographic gradient fraction derived from fractionation of HeLa fraction [A] over DEAE-Sephacel as in Experimental Procedures (lanes 3–12). Binding reactions were analyzed by native gel electrophoresis as above. Following autoradiography, panels corresponding to the bound complexes formed on each of the three probes were excised and mounted together for direct comparison. Arrows indicate the position of specific protein–DNA complexes formed by CP1, CP2, and NF-I.

plex. Thus, if these subunits form a stable complex in HeLa cell extracts, but are separated by cation-exchange chromatography, a binding activity present in unfractionated extracts would appear absent in chromatographic fractions. When the appropriate fractions containing the two subunits were combined, they could reassociate and thereby regenerate the original binding activity. In order to test this possibility, partially purified fractions containing CP1A alone, CP1B alone, or a combination of CP1A and CP1B were sedimented through glycerol gradients and the resulting fractions assayed for CP1A and CP1B binding activities (Figure 5). CP1B activity derived from phosphocellulose fraction C peaked in gradient fraction 3, corresponding to a Svedberg coefficient of approximately 2.7 S (Figure 5, top panel, lane 5). CP1A activity derived from DEAE fraction AB peaked in gradient fraction 4, corresponding to a Svedberg coefficient of 3.7 S (Figure 5, second panel from top, lane 6). When fractions AB and C were mixed and preincubated prior to loading on the glycerol gradient, both CP1A and CP1B activities sedimented considerably more rapidly and both activities peaked in fraction 7 (Figure 5, bottom two panels, lane 9).



each fraction were analyzed by SDS-polyacrylamide electrophoresis. To assay for CP1A activity (second panel from top, and bottom panel), 20,000 cpm (0.3 ng) of the MLP-containing probe, RMC, was added to a standard binding reaction containing a mixture of 7.5 μ g of fraction [C] (lanes 1–16), and either 36 μ g fraction of [AB] (lane 1), 7 μ l of glycerol gradient buffer containing no protein (lane 2), or 7 μ l of gradient fractions 1–14 (lanes 3–16, respectively). To assay for CP1B activity (top panel and second panel from bottom), 20,000 cpm (0.3 ng) of the MLP-containing probe, RMC, was added to a standard binding reaction containing a mixture of 36 μ g of fraction [AB] (lanes 1–16), and either 7.5 μ g of fraction [C] (lane 1), 7 μ l of glycerol gradient buffer containing no protein (lane 2), or 7 μ l of gradient fractions 1–14 (lanes 3–16, respectively). Binding reactions were analyzed by native gel electrophoresis as above. Following autoradiography, panels corresponding to the bound complexes formed on the MLP probe were excised and mounted together for direct comparison. Svedberg coefficients were determined by comparison of the activity peak in each gradient to the sedimentation profile of protein standards. The Svedberg coefficients corresponding to the activity peak from each gradient are indicated to the right of each panel.

This sedimentation rate corresponded to a Svedberg coefficient of 5.3 S, indicating that CP1A and CP1B were present as part of a larger complex. This complex was stable at 4°C over the period of 16 hr required for centrifugation. Binding activity assayed in the absence of exogenously added fractions AB or C indicated that glycerol gradient fraction 7 contained all of the components required for specific CP1 binding (data not shown).

Interestingly, in the gradients containing both CP1A and CP1B activities, the activity peaks corresponding to the individually sedimented factors at 2.7 S and 3.7 S were not detectable (Figure 5, bottom two panels, lanes 3 and 4). Since roughly equivalent amounts of CP1A and CP1B activity were loaded onto the gradient (data not shown), it is possible that the two activities efficiently associated with one another. If this occurred, it can be inferred that CP1A and CP1B bind to each other with high affinity.

The Svedberg coefficients of CP1A and CP1B correspond to molecular weights of approximately 45 and 25 kd, respectively, assuming they are globular proteins. The sedimentation rate of the CP1A/CP1B complex corresponds to a molecular weight of approximately 140–150 kd. Thus if CP1 and its component subunits behave as globular proteins, it is probable that the complex formed in solution is composed of more than a monomer of CP1A and a monomer of CP1B. CP1 may in fact be a heterologous complex composed of two molecules each of CP1A and CP1B. However, since there is no information regarding the shape of these molecules, and since there may be a significant conformational change following their association, the stoichiometry of components contained in CP1 cannot be determined from this experiment.

Figure 5. Glycerol Gradient Sedimentation Analysis of CP1A and CP1B Activities

Two hundred and fifty micrograms of HeLa fraction [C], containing partially purified CP1B (top panel), 1.2 mg of HeLa fraction [AB], containing partially purified CP1A (second panel from top), or a mixture of 250 μ g of fraction [C] and 1.2 mg of fraction [AB] (bottom two panels) were incubated for 10 min at 30°C. Reactions were layered onto 10%–30% glycerol gradients and sedimented at 45,000 rpm for 19 hr at 4°C. Fourteen fractions (275 μ l each) were taken from each gradient and assayed for either CP1A activity or CP1B activity. The top of the gradient was fraction 1. In a parallel gradient, 50 μ g each of ovalbumin, bovine serum albumin, carbonic anhydrase, cytochrome C and catalase were sedimented, and portions of

CP1A and CP1B Are Both Present in the CP1–DNA Complex

If a CP1A/CP1B heterologous complex is in fact the species that specifically binds to DNA, then both molecules should copurify during DNA-affinity chromatography. CP1 was purified by chromatography on a specific, MLP-containing DNA-affinity column. Affinity-purified CP1 was subsequently subjected to phosphocellulose chromatography under conditions similar to those that separate CP1A from CP1B (Figure 6; see Experimental Procedures). Following chromatography, each fraction was assayed for the presence of CP1A and CP1B activities. The affinity-purified CP1 binding activity loaded onto the column could not be detected either in the 0.06 M KCl flowthrough from the column or in the 0.6 M KCl eluate (Figure 6, lanes 4, 5, and 8). Addition of partially purified CP1B to the flowthrough fraction resulted in a dramatic stimulation of specific CP1 binding activity (Figure 6, lane 7). Conversely, addition of partially purified CP1A to the flowthrough fraction had no effect upon binding activity (Figure 6, lane 6). This suggested that CP1A was present in the flowthrough fraction. Similarly, addition of partially purified CP1A to the 0.6 M KCl phosphocellulose fraction resulted in a dramatic stimulation of binding activity. Conversely, addition of partially purified CP1B to this same fraction had no effect upon binding activity (Figure 6, lanes 9 and 10). This suggested that CP1B was present in the 0.6 M phosphocellulose fraction. The eluate from a DNA-affinity column formed with a control fragment contained neither CP1A nor CP1B (data not shown). Thus, CP1 purified by DNA-affinity chromatography contained both CP1A and CP1B activities. Therefore not only are

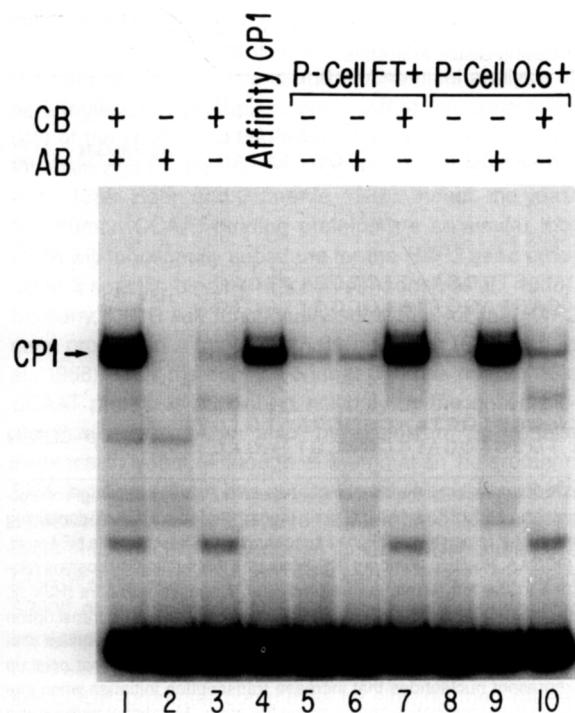


Figure 6. CP1A and CP1B Are Present in the CP1-DNA Complex

CP1 activity was purified by DNA-affinity chromatography as described in Experimental Procedures, and loaded onto a phosphocellulose column in buffer A + 60 mM KCl. The flowthrough (FT) was collected, and the column was eluted with Buffer A + 0.6 M KCl as described in Experimental Procedures. 20,000 cpm (0.1 ng) of an MLP probe containing sequences from -108 to -62 was incubated in a standard CP1 binding reaction (lanes 1-10) containing either 12 μ g of CP1A-containing fraction [AB] (lanes 1, 2, 6, and 9), 4 μ g of CP1B-containing fraction [CB] (lanes 1, 3, 7, and 10), 1 μ l of affinity-purified CP1 (lane 4), 1 μ l of the affinity-purified CP1 phosphocellulose flowthrough (P-Cell FT) fraction (lanes 5-7), or 1 μ l of the affinity-purified CP1 phosphocellulose 0.6 M KCl (P-Cell 0.6) eluate (lanes 8-10). Lane 1 contained a mixture of partially purified CP1A activity ([AB]) and partially purified CP1B activity ([CB]). Lanes 2 and 3 contained partially purified CP1A alone, and partially purified CP1B alone, respectively. Lane 4 contained 1 μ l of affinity-purified CP1. In lanes 5-7, 1 μ l of the affinity-purified CP1 phosphocellulose flowthrough fraction was incubated either alone (lane 5), with partially purified CP1A (lane 6), or with partially purified CP1B (lane 7). In lanes 8-10, 1 μ l of the affinity-purified CP1 phosphocellulose 0.6 M eluate was incubated either alone (lane 8), with partially purified CP1A (lane 9), or with partially purified CP1B (lane 10). Binding reactions were analyzed by native gel electrophoresis as above.

CP1A and CP1B both required for DNA-binding activity, they are also each present in the CP1-DNA complex.

Discussion

We have identified three distinct proteins, CP1, CP2, and NF-I, which are present in extracts of HeLa cells and which bind to different subsets of promoter elements containing the sequence CCAAT. Each of these proteins has been shown to make critical contacts within the CCAAT pentanucleotide sequence by methylation interference analysis and by competition binding analysis with mutants harboring alterations in the CCAAT element. The three

CCAAT-binding proteins were distinct since first, each formed a specific protein-DNA complex of unique mobility and appearance during native gel electrophoresis; second, each made a unique set of contacts with its binding site as assayed by methylation interference; third, each bound with a unique pattern of affinities to a variety of promoters containing CCAAT elements; and fourth, each of the three CCAAT proteins could be separated chromatographically.

How do the CCAAT-binding proteins identified here relate to those that have previously been identified? CP1 binds with high affinity to promoters that include human α -globin, human hsp70, H-2K^b, the adenovirus MLP, and the MSV LTR. Based upon the consensus sequence for CP1 (Table 1) and published methylation interference patterns for other CCAAT-binding factors, CP1 may be similar or identical to CCAAT-binding factors that bind to the sea urchin sperm histone H2B promoter, the human thymidine kinase promoter, the human hsp70 promoter, the mouse α -globin promoter, and the MHC class II E_u promoter (Barberis et al., 1987; Knight et al., 1987; Wu et al., 1987; Cohen et al., 1986; Dorn et al., 1987).

It is more difficult to establish the relationship of CP2 to previously described CCAAT-binding factors. CP2 binds with high affinity to CCAAT elements present in the γ -FBG promoter, the H-2K^b promoter, and the MOPC-41 V_H promoter (see Table 1 for consensus sequence). Proteins binding to these elements have not been previously described.

A previously characterized high-affinity binding site was used to detect NF-I (Rosenfeld et al., 1987). This probe generated a heterogeneous pattern of specific protein-DNA complexes with HeLa cell extracts in the gel electrophoresis DNA-binding assay. This pattern is highly distinctive, and essentially identical to complexes generated by affinity-purified NF-I (Rosenfeld et al., 1987; Jones et al., 1987). As anticipated from previous results, the formation of all of these complexes was eliminated by a double point mutation in the NF-I binding site (Rosenfeld et al., 1987). Finally, the methylation interference pattern observed for the heterogeneous set of NF-I protein-DNA complexes was identical to that reported for the binding of purified NF-I (De Vries et al., 1987). On this basis, we conclude that the collection of proteins that specifically binds to the probe containing the consensus NF-I binding site is NF-I.

It has previously been reported that a single CCAAT-binding protein, CTF/NF-I, is both a promoter-specific transcription factor (CTF) for RNA polymerase II and an initiation factor for adenovirus DNA replication (NF-I) (Jones et al., 1987). NF-I is a cellular protein that binds with high-affinity to the adenovirus origin of replication and is required for the initiation of replication (Nagata et al., 1983; Leegwater et al., 1985; Rosenfeld et al., 1987). CTF has been defined as a cellular DNA-binding protein that specifically stimulates transcription from the human α -globin promoter, the HSV thymidine kinase promoter, and the human hsp70 promoter by selective recognition of CCAAT elements in these promoters (Jones et al., 1985; Jones et al., 1987; Morgan et al., 1987). CTF has also been reported

to bind to CCAAT elements in the MSV LTR and the human β -globin promoter (Jones et al., 1985; Jones et al., 1987). Interestingly, these five CCAAT elements bound by CTF/NF-I are also bound by CP1 with virtually the same pattern of affinities, with the important exception that CP1 does not bind with appreciable affinity to the adenovirus origin of replication or to the NF-I consensus sequence. Moreover, the group of CCAAT-binding proteins that recognizes the NF-I consensus sequence probe does not bind with appreciable affinity to any of the CCAAT elements that are recognized by CP1. Thus, the high-affinity binding specificities of CP1 and NF-I are entirely nonoverlapping. However, the sum of the binding specificities of CP1 and NF-I are equivalent to the proposed binding specificity of the putative CTF/NF-I proteins.

Mutational analysis of the HSV tk and β -globin promoters suggests that CP1, and not NF-I, is likely to be responsible for the transcription stimulation mediated by the CCAAT elements in these promoters (Graves et al., 1986; Myers et al., 1986; Charnay et al., 1985). Several mutations that affect the transcription efficiency of these promoters do not alter sequences shared with the NF-I consensus sequence (Table 2; Rosenfeld et al., 1987; Leegwater et al., 1985). Each of these mutations, however, does alter sequences shared with the consensus sequence of CP1 (Table 1). Specifically, the positions of two mutations that decrease transcription from these two promoters (HSV tk -88 [A to C]; β -globin -72 [C to T]) and three mutations that increase transcription from these two promoters (HSV tk -87 [G to C]; HSV tk -80 [C to G]; β -globin -79 [G to A]) correspond to sites where base substitutions have no effect on the affinity of NF-I binding (Table 2; Graves et al., 1986; Myers et al., 1986; Rosenfeld et al., 1987). In contrast, nucleotides at each of these positions are conserved in high-affinity CP1 binding sites. For all five mutations, those which increase transcriptional efficiency increase the degree of similarity to the CP1 consensus sequence, while those that decrease transcriptional efficiency decrease the degree of similarity to the CP1 consensus sequence. Thus, the effect on transcriptional efficiency of a number of mutations in the CCAAT elements of the HSV tk and mouse β -globin promoters are more readily correlated with CP1 binding specificity than with NF-I binding specificity.

Surprisingly, each of the three CCAAT-binding proteins, CP1, CP2, and NF-I, appear to be composed of at least two different subunits. This unique property was investigated most thoroughly for CP1. CP1 binding activity in HeLa extracts was lost following phosphocellulose chromatography, and was only reconstituted when two partially purified fractions containing CP1A and CP1B were combined. Neither CP1A nor CP1B alone exhibited a significant amount of DNA-binding activity. However, DNA-affinity purified CP1 contained both CP1A and CP1B activities, demonstrating that both polypeptides were present in the CP1-DNA complex. Moreover, glycerol gradient sedimentation revealed that CP1A and CP1B form a stable complex in solution in the absence of a CP1 binding site. The simplest interpretation of these results is that CP1 is a multimeric protein composed of heterologous subunits.

Table 2. Mutational Analysis of the HSV tk and Mouse β -Globin Promoters

22	32	42	
T T T T G G A T T G A A G C C A A T A T G			Ad ORIGIN
A A A A C C T A A C T T C G G T T A T A C			
69	79	89	
G T G T T C G A A T T C G C C A A T G A C	G	C	HSV tk
C A C A A G C T T A A G C G G T T A C T G			
-90	-60	-70	
A C C C T G G T A A G G A G C C A A T C T G	A		MOUSE β -GLOBIN
T G G G A C C A T T C C C G G T T A G A C			

The sequences of the NF-I binding site in the adenovirus origin of replication, and of the HSV tk and mouse β -globin CCAAT-containing promoter regions were aligned according to the best fit to the NF-I motif, GCCAA. Positive numbers (NF-I) indicate the nucleotide position relative to the left terminus of adenovirus. Negative numbers (HSV tk, β -globin) indicate the nucleotide position relative to the transcription initiation site. Arrows that point down represent nucleotides that decrease transcription initiation when mutated. Arrows that point up represent nucleotides that increase transcription initiation when mutated to the nucleotide shown above the arrow. Mutational analysis was derived from Graves et al. (1986), Myers et al. (1986), and Charnay et al. (1985).

The CCAAT-binding proteins CP2 and NF-I showed a two-factor dependence for DNA-binding which was very similar to that exhibited by CP1. Each required the mixing of two fractions in order to reconstitute maximum binding activity. It is potentially important to note that the two factor binding requirement was not as stringently observed for CP2 and NF-I as it was for CP1. High concentrations of CP2B or NF-IB alone, but not of CP1B alone, were capable of specific DNA-binding (see below). The chromatographic properties of the fractions containing the CP1A, CP2A, and NF-IA activities were remarkably similar. Nevertheless, CP1A, CP2A, and NF-IA could be shown to be chromatographically distinct and therefore most likely reflect the activities of different polypeptides. The complementary group of activities, CP1B, CP2B, and NF-IB, also behaved similarly during chromatography, but CP1B was shown to be chromatographically distinct from CP2B and NF-IB. Further evidence also suggested that CP2B and NF-IB were distinct. Although we favor the model that CP2 and NF-I are also multisubunit proteins composed of heterologous subunits, this must be considered a hypothesis at present. However, this possibility is strengthened by the extensive similarity between these three CCAAT-binding factors in their recognition sequences, their apparent two-factor binding dependence, and the similar chromatographic behavior of each of their components. Thus we suggest that CP1, CP2, and NF-I are members of a family of evolutionarily related, heteromeric CCAAT-binding proteins.

Recently, the CCAAT-binding protein CP1 was shown to be highly homologous to a yeast CCAAT-binding protein

that activates transcription from the iso-1-cytochrome C (CYC1) promoter (Chodosh et al., 1988). This yeast protein has been shown by genetic and biochemical analysis to be a multisubunit complex composed of the gene products of the *HAP2* and *HAP3* loci (Guarente et al., 1984; Pinkham and Guarente, 1985; Olesen et al., 1987; Hahn et al., 1988; Hahn and Guarente, 1988). In fact, the yeast and human CCAAT-binding proteins are so similar that CP1A will functionally substitute for the *HAP3* gene product in a specific binding reaction (Chodosh et al., 1988). Similarly, CP1B will functionally substitute for the *HAP2* gene product in a specific binding reaction (Chodosh et al., 1988). Thus the conclusion that CP1 is a heteromeric CCAAT protein is supported not only by the above biochemical evidence but also, by extension, by genetic evidence in yeast. Although a mammalian heteromeric DNA-binding protein has not previously been described, a possible example is the specific interaction between the ovalbumin-specific transcription factor, COUP, and a non-DNA-binding factor, S300II (Tsai et al., 1987). Although S300II does not bind DNA, it is absolutely required for transcription of the ovalbumin promoter, but not for the SV40 early promoter. It appears that S300II is able to stabilize COUP-promoter complexes by slowing their rate of dissociation.

The factors, CP1A, CP2A, and NF-1A flow through negatively charged columns while the factors CP1B, CP2B, and NF-1B bind to DNA-cellulose and to other negatively charged columns. This suggests that the latter set may possess a degree of intrinsic DNA-binding activity. Specific DNA-protein complexes were observed with fractions containing only CP2B or NF-1B. Although these complexes could reflect the intrinsic binding of the single B factor, it is impossible to exclude the potential presence of contaminating A factor. The methylation interference pattern of the complex generated by CP2B alone was similar to that generated by CP2, but methylation at many more sites disrupted the CP2B-DNA complex. This would be expected if the affinity of the CP2B-DNA complex was lower than that of the CP2-DNA complex. By extension, we would predict that NF-1B could also be footprinted on DNA in the absence of NF-1A. Therefore, polypeptides encoded by the cDNA clones of either CP2B or NF-1B would specifically bind to their respective recognition sites. However, we would predict that the affinity of the cloned protein would be lower than that of the complete heteromer, and that the heteromer is the active form in vivo.

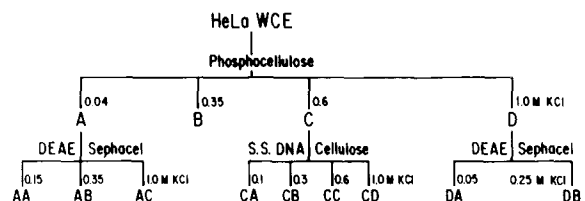
What is the functional significance of heteromeric CCAAT-binding proteins? A greater variety of regulatory options is available to a single protein that is composed of two or more subunits, each of which may be regulated independently in response to tissue-specific, temporal or environmental signals. In addition, the interchange of subunits among such proteins would provide an additional level of combinatorial complexity to gene regulation. If one of the polypeptides in the heteromer contains the DNA-binding specificity, then exchange of the second polypeptide with another factor could play an important role in the transcriptional regulation of a particular promoter. For example, CP1A-like factors may affect transcrip-

tional activation by conferring a particular specificity on, or increasing the efficiency of transcriptional activation of, CP1B-like factors. This model would readily help explain how some molecules which appear to lack intrinsic DNA-binding activity are able to modulate the transcription activity of DNA-binding proteins (Kingston et al., 1985). Thus, the *trans*-activator protein E1A of adenovirus may be a CP1A-like molecule. Conversely, this model could also help explain how DNA-binding proteins of apparently limited specificity are able to specifically activate transcription. The cellular proto-oncogene products *c-myc* and *c-fos* may function by this mechanism and may therefore be CP1B-like molecules.

Experimental Procedures

Plasmids

Plasmid pRMC was prepared by subcloning a 127 bp EcoRI-HaeIII fragment from pLP into EcoRI-SmaI-digested pUC13. pLP, whose construction has previously been described (Chodosh et al., 1986), contains Ad 2 MLP sequences from -174 to +33. The 5' deletion/linker scanning mutant, pΔ108-MLP, was prepared by digesting the 161 bp EcoRI-HindIII fragment of pRMC with MboI, blunting the ends with T4 DNA polymerase, and digesting the resulting MboI-HindIII fragment with BamHI. The MboI-BamHI fragment was subcloned into HincII-BamHI-digested pRW, a plasmid containing Ad 2 MLP sequences from -51 to +33 whose construction has previously been described (Chodosh et al., 1986). The resulting plasmid, pΔ108-MLP, contains Ad 2 MLP sequences from -108 to +33, with a linker-scanning mutation of the MLTF-binding site that replaces sequences from -62 to -53 with a pUC13 polylinker sequence containing a BamHI restriction site. Plasmid pαGBCCAAT contains human α-globin sequences from -81 to -58, and was prepared by inserting a synthetic oligonucleotide of sequence CTAGACTCCGCGCCAGCCAATGAGCGCCGT into the XbaI site of pUC18. Plasmid pαGBCCAAT-dm is identical to pαGBCCAAT with the exception that the inserted oligonucleotide differs at two positions, and contains the sequence CTAGACTCCGCGCCAGCGAGTGAGCGCCGT. Plasmid pγFBG-95, whose construction has previously been described (Chodosh et al., 1987), contains rat γ-fibrinogen sequences from -95 to +35. The plasmids pID-62/-54, pID-65/-54, and pID-74/-54, which were kind gifts from J. Morgan, are internal deletion mutants of the rat γ-FBG promoter and have been described previously (Chodosh et al., 1987; Morgan et al., 1988). Plasmid pNF-I contains a high-affinity NF-I binding site derived by mutating a single position of the adenovirus origin of replication. This mutant has previously been characterized as pUpm28 (Rosenfeld et al., 1987). pNF-I was constructed by inserting a synthetic oligonucleotide of sequence GATCATTTTGCTTGAAGCCAATATGA into the BamHI site of pUC13. The plasmid pNF-I-dm was prepared identically to pNF-I, with the exception that the inserted oligonucleotide differed at two positions, and contained the sequence GATCATTTTGCTTGAAGCAAATATGA. Plasmid pR1B contains 84 to 100 map units of Ad 5 cloned into the EcoRI site of pBR322 and has been previously described (Fire et al., 1981). The plasmid pH-2K^Δ15 (A. Baldwin, unpublished results) was derived from p365H2KCAT (Baldwin and Sharp, 1987), and contains sequences from -65 to +3 of the mouse class I major histocompatibility H-2K^b gene subcloned into pUC-CAT (gift from M. Gilman). The plasmid pH2KCCAAT-72 contains sequences from the MHC H-2K^b promoter between -90 and -64 and was prepared by inserting a



synthetic oligonucleotide of sequence GATCCGCGCCACCCAATG-GGGGTAAAGACTG into the BamHI site of pUC13. The plasmid p328-pUC, a gift from D. Galson, contains mouse β -globin sequences from -344 to -15 and has been described previously (Galson and Housman, 1988). The plasmid pSPiG_{V_k} gene was a gift from N. Speck and contains the 250 bp StaNI promoter fragment of the MOPC-41 V_k gene encompassing sequences upstream of -47 (Singh et al., 1986). All plasmids constructed by inserting synthetic oligonucleotides into pUC13 or pUC18 were sequenced to confirm single copy insertion and proper sequence.

Probes for Gel Electrophoresis DNA-Binding Assay

All probes were prepared by end-labeling DNA fragments with Klenow and either α -[³²P]dATP or α -[³²P]dCTP, and were isolated by agarose gel electrophoresis onto NA-45 DEAE-membrane followed by elution in a buffer consisting of 10 mM Tris-HCl (pH 7.9), 1 mM EDTA, and 1.0 M NaCl according to the supplier's recommendations (Schleicher & Schuell). The 161 bp MLP-containing probe RMC was prepared by end-labeling an EcoRI-HindIII digest of pRMC. The 74 bp MLP-containing probe, Δ 108, which contains MLP sequences from -108 to -62, was prepared by end-labeling a BamHI-HindIII digest of p Δ 108-MLP. The 90 bp γ -fibrinogen-containing probe, γ -FBG, which contains γ -FBG sequences from -95 to -39, was prepared by end-labeling an EcoRI/HindIII digest of p γ FBG-95. The 77 bp probe, NF-I, which contains sequences corresponding to +21 to +45 of the left terminus of Ad 2 with a point mutation at +28, was prepared by end-labeling an EcoRI/HindIII digest of pNF-I.

Competitor DNA Fragments for Gel Electrophoresis DNA-Binding Assay

The competitor DNA fragment, MLP (MLP-108), was prepared by digesting p Δ 108-MLP with BamHI and HindIII, and isolating the 74 bp fragment (-108 to -62 relative to the MLP cap site). The competitor fragment, MLP-53, was prepared by digesting p Δ -108-MLP with EcoRI and BamHI, and isolating the 98 bp fragment (-53 to +33). The competitor fragment, γ -FBG, was prepared by digesting p γ FBG-95 with EcoRI, and isolating the 90 bp fragment (-95 to -39). The competitor fragments, ID-62/-54, ID-65/-54, and ID-74/-54, were prepared by digesting the plasmid corresponding to each with BstEII and BstXI, and isolating the 122, 119, or 110 bp fragment, respectively (-95 to +35 with internal deletion of -62 to -54, -65 to -54, or -74 to -54, respectively). The competitor fragment, NF-I, was prepared by digesting pNF-I with EcoRI and HindIII, and isolating the 77 bp fragment (corresponding to +21 to +45 of the left terminus of Ad 2 with a point mutation at +28). The fragment, NF-I-dm, was prepared by digesting pNF-I-dm with EcoRI and HindIII, and isolating the 77 bp fragment. The fragment, Ad origin, was prepared by digesting pRIB with SmaI and EcoRI, and isolating the 580 bp fragment (-329 to +246). The competitor fragment, α -globin, was prepared by digesting α GBCCAAT with EcoRI and HindIII, and isolating the 80 bp fragment (-81 to -58). The competitor fragment, α -globin-dm, was prepared by digesting α -GBCCAAT-dm with EcoRI and HindIII, and isolating the 80 bp fragment. The competitor fragment, H-2K-51, was prepared by digesting pH-2K^b Δ 15 with XhoI and EcoRI, and isolating the 335 bp fragment (-65 to +3). The competitor fragment, H-2K-72, was prepared by digesting pH2KCCAAT-72 with EcoRI and HindIII, and isolating the 81 bp fragment (-90 to -64). The HSV tk competitor fragment was prepared by annealing the gel-purified 31 nucleotide synthetic oligonucleotide of sequence CTAGATGTTCGAATTCGCCAATGACAA-GACT to the complementary oligonucleotide CTAGATCTCTTGTCATTGGCGAATTCGAACAT (-94 to -70). The competitor fragment, β -globin, was prepared by digesting p328pUC with BamHI and HindIII and isolating the 330 bp fragment (-344 to -15). The competitor fragment, V_k, was prepared by digesting pSPiG_{V_k} with EcoRI and HindIII, and isolating the 300 bp fragment (-47 to -297). The competitor fragment, pUC-18, was prepared by digesting pUC18 plasmid with EcoRI and HindIII, and isolating the 50 bp polylinker. All fragments were purified by agarose gel electrophoresis onto NA-45 as above. DNA fragment concentrations were estimated by comparison with E. coli DNA digested with HincII using the ethidium bromide dot method (Maniatis et al., 1982).

Gel Electrophoresis DNA-Binding Assay

Protein-DNA complexes were resolved on low ionic strength native polyacrylamide gels as described (Carthew et al., 1985), with several modifications. Protein samples were incubated with approximately 20,000 cpm (0.1-0.3 ng) of an end-labeled double-stranded DNA fragment in the presence of 2.0 μ g poly(dI-dC):poly(dI-dC) in a final volume of 10 μ l. Incubations were carried out at 30°C for 10-30 min. All binding reactions contained 12% glycerol, 12 mM HEPES-NaOH (pH 7.9), 4 mM Tris-HCl (pH 7.9), 1 mM EDTA, 0.6 mM dithiothreitol, and 300 μ g/ml bovine serum albumin. In addition, CP1 and CP2 binding reactions contained 60 mM KCl, whereas NF-I binding reactions contained 150 mM KCl and 5 mM MgCl₂ except as indicated. Following electrophoresis, gels were dried onto Whatmann 3MM and visualized by autoradiography on Kodak XAR-5 film.

Methylation Interference Assay

The procedure of Siebenlist et al. (1980) was used with several modifications. Probes for methylation interference assays were prepared as above, except that each was uniquely end-labeled. In order to label the coding strand, p Δ 108-MLP, p γ FBG, and pNF-I were first digested with BamHI, EcoRI, and EcoRI, respectively. In order to label the noncoding strand, p Δ 108-MLP, p γ FBG, and pNF-I were first digested with HindIII, BamHI, and HindIII, respectively. Digested plasmids were end-labeled with Klenow and either α -[³²P]dATP or α -[³²P]dCTP. Following heat inactivation, plasmids p Δ 108-MLP, p γ FBG, and pNF-I labeled on the coding strand were redigested with HindIII, BamHI, and HindIII, respectively. Plasmids p Δ 108-MLP, p γ FBG, and pNF-I labeled on the noncoding strand were redigested with BamHI, EcoRI, and EcoRI, respectively. Probes were partially methylated with dimethylsulfate (Maxam and Gilbert, 1980), purified, and used as substrate in a standard binding reaction as above, with the exception that reactions were scaled up to 50 μ l, and contained 200,000 cpm of the appropriate probe in addition to 10 μ g poly(dI-dC)-(dI-dC), 180 μ g fraction [AB], and 58 μ g fraction [CB]. A mock binding reaction lacking fractions [AB] or [CB] was performed in parallel. After incubation at 30°C for 30 min, binding reactions were loaded onto native polyacrylamide gels and electrophoresed as above. Brief autoradiography at 4°C identified free and bound DNA-containing complexes which were excised from the native polyacrylamide gel and embedded in an agarose gel. Labeled DNA was electrophoresed from the polyacrylamide gel slices, through the agarose matrix, and onto NA-45 DEAE-membrane. DNA bound to the NA-45 membrane was eluted as above, purified, piperidine-cleaved, electrophoresed on a 10% polyacrylamide-8 M urea 0.3 mm thickness gel, and visualized by autoradiography. Gels from which complexes had been excised were revisualized by autoradiography to confirm that the appropriate complexes had been accurately excised.

Affinity Purification of CP1

CP1 was affinity-purified using a modification of the previously described streptavidin-biotin technique (Chodosh et al., 1986). The HincII-EcoRI fragment of pRMC was isolated and biotinylated with biotin-dUTP and Klenow fragment (Chodosh et al., 1986). One hundred and forty-four nanograms of the 145 bp biotinylated fragment was incubated with 100 μ g poly(dI-dC)-(dI-dC), 250 μ g/ml BSA, 2.4 mg fraction [AB], and 0.87 mg fraction [CB] in a final volume of 367 μ l under standard CP1 binding conditions at 30°C for 60 min. Eight hundred and fifty nanograms of streptavidin (Apel) were added, and the reaction continued for an additional 15 min. Reactions were chilled on ice, added to 30 μ l of pre-equilibrated biotin-cellulose (Pierce), resuspended, and gently mixed at 4°C for 30 min. The reaction was centrifuged at 2000 rpm \times 1 min and the biotin-cellulose pellet was washed five times with 10 ml each of cold binding buffer, transferred to a clean tube, poured into a P1000 Eppendorf pipet tip plugged with siliconized glass wool, and eluted with 200 μ l of Buffer A + 1.0 M KCl + 200 μ g/ml BSA. Fractions containing CP1 activity were pooled and dialyzed against Buffer A + 0.06 M KCl.

Extracts and Chromatography

HeLa whole-cell extracts were prepared as previously described (Manley et al., 1980). Chromatography of HeLa whole cell extract to generate fractions [A], [AB], [C], and [CB] was as described (Samuels et al.,

1982). Linear salt-gradient elution of [A] activity from DEAE-Sephacel was performed by loading 400 ml of phosphocellulose fraction [A] adjusted to 0.15 M KCl onto a 120 ml (2.5 × 24 cm) DEAE-Sephacel column. The column was washed with 3 column volumes of Buffer A + 0.1 M KCl. Buffer A consisted of 20% glycerol, 20 mM HEPES-NaOH (pH 7.9), 1 mM EDTA, and 1 mM dithiothreitol. The column was eluted with a 6 column volume linear gradient from 0.15 M–0.4 M KCl in Buffer A. Eleven milliliter fractions were collected, and dialyzed against Buffer A + 0.1 M KCl.

Linear gradient elution of [C] activity from DNA-cellulose (M. Samuels, unpublished data) was performed as follows: 90 ml of fraction [C], derived from 70 ml of HeLa WCE, were adjusted to 0.21 M KCl and applied to a 120 ml (2.5 × 24 cm) DEAE-Sephacel column equilibrated in Buffer A + 0.1 M KCl (Samuels et al., 1982). The flow-through, which contained CP1, CP2, and NF-I activities, was applied to a 40 ml (2.5 × 8 cm) DNA-cellulose column equilibrated in Buffer A + 0.1 M KCl + 5 mM MgCl₂. The column was washed with 5 column volumes of equilibration buffer and eluted with a 5 column volume linear gradient from 0.1 M to 0.8 M KCl in Buffer A + 5 mM MgCl₂. Ten milliliter fractions were collected and dialyzed against Buffer A + 0.1 M KCl.

Phosphocellulose chromatography of affinity-purified CP1 was carried out on a 50 μl phosphocellulose column poured and equilibrated with A + 0.1 in an Eppendorf P1000 pipet tip plugged with siliconized glass wool. The column was sequentially prewashed with A + 0.1 M KCl + 200 μg/ml BSA and A + 1.0 M KCl and equilibrated in A + 0.06 M KCl. Fifty microliters of affinity-purified CP1 was loaded onto the column at a flow rate of 6 column volumes per hour. The column was washed with 3 column volumes of A + 0.06 M KCl + 200 μg/ml BSA and eluted with 3 column volumes of A + 0.6 M KCl + 200 μg/ml BSA. Twenty-five microliter fractions were collected, and dialyzed against A + 0.1 M KCl.

Glycerol Gradient Sedimentation

Aliquots of fractions [AB] and [C] were dialyzed against 20 mM HEPES-NaOH (pH 7.9), 1 mM EDTA, 1 mM DTT, 60 mM KCl, and 8% glycerol. Fifty microliters of dialyzed fraction [AB] and 50 μl of dialyzed fraction [C] were incubated alone or in combination for 10 min at 30°C. Reactions were layered onto 10%–30% glycerol gradients and sedimented at 45,000 rpm for 19 hr in a SW60 rotor at 4°C. Fourteen fractions (275 μl each) were taken from each gradient, and 7 μl of each fraction were tested in a standard CP1 binding assay containing the complementary CP1 fraction and 0.4 ng of the MLP probe, RMC. In a parallel gradient, 50 μg each of ovalbumin, bovine serum albumin, carbonic anhydrase, cytochrome C, and catalase were sedimented, and portions of each fraction analyzed by SDS-polyacrylamide gel electrophoresis.

Acknowledgments

We thank members of the Sharp laboratory for numerous helpful discussions. M. Garcia-Blanco, J. LeBowitz, H. Singh, and S. Hahn for critical reading of the manuscript, S. Scaringe for synthesizing DNA oligonucleotides, M. Samuels for preparation of some chromatographic fractions, and D. Galson, N. Speck, M. Gilman, J. Morgan, and G. Crabtree for some of the plasmids used in this study. We thank M. Siatfaca for her patience and grace under pressure. L. A. C. is supported by the Harvard Medical School M.D.–Ph.D. program by Public Health Service National Research Service Award 5T 32 GM07753-09 from the National Institute of General Medical Sciences, and by the M.I.T. Department of Biology. A. S. B. is supported by a Leukemia Society Special Fellow Award. This work was supported by U. S. Public Health Service grant PO1-CA42063 from the National Institutes of Health, #CDR-8500003 from the National Science Foundation, and, partially, by National Cancer Institute Cancer Center core grant P30-CA14051 to P. A. S.

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.

References

- Arriza, J. L., Weinberger, C., Cerelli, G., Glaser, T. M., Handelin, B. L., Housman, D. E., and Evans, R. M. (1987). Cloning of the human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor. *Science* 237, 268–274.
- Baldwin, A. S., and Sharp, P. A. (1987). Binding of a nuclear factor to a regulatory sequence in the promoter of the mouse H-2K^b class I major histocompatibility gene. *Mol. Cell. Biol.* 7, 305–313.
- Barberis, A., Superti-Furga, G., and Busslinger, M. (1987). Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. *Cell* 50, 347–359.
- Benoist, C., O'Hare, K., Breathnach, R., and Chambon, P. (1980). The ovalbumin gene: sequence of putative control regions. *Nucl. Acids Res.* 8, 127–142.
- Carthew, R. W., Chodosh, L. A., and Sharp, P. A. (1985). An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. *Cell* 43, 439–448.
- Charnay, P., Mellon, P., and Maniatis, T. (1985). Linker scanning mutagenesis of the 5'-flanking region of the mouse β-major-globin gene: sequence requirements for transcription in erythroid and nonerythroid cells. *Mol. Cell Biol.* 5, 1498–1511.
- Chodosh, L. A., Carthew, R. W., and Sharp, P. A. (1986). A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell Biol.* 6, 4723–4733.
- Chodosh, L. A., Carthew, R. W., Morgan, J. G., Crabtree, G. R., and Sharp, P. A. (1987). The adenovirus major late transcription factor activates the rat γ-fibrinogen promoter. *Science* 238, 684–688.
- Chodosh, L. A., Olesen, J., Hahn, S., Baldwin, A. S., Guarente, L., and Sharp, P. A. (1988). A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. *Cell* 53, 25–35.
- Cohen, R. B., Sheffery, M., and Kim, C. G. (1986). Partial purification of a nuclear protein that binds to the CCAAT box of the mouse α₁-globin gene. *Mol. Cell Biol.* 6, 821–832.
- De Vries, E., Van Driel, W., Van den Heuvel, S. J., and Van der Vliet, P. C. (1987). Contactpoint analysis of the HeLa nuclear factor I recognition site reveals symmetrical binding at one side of the DNA helix. *EMBO J.* 6, 161–168.
- Dorn, A., Bollekens, J., Staub, A., Benoist, C., and Mathis, D. (1987). A multiplicity of CCAAT box-binding proteins. *Cell* 50, 863–872.
- Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissmann, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Sholders, C. C., and Proudfoot, N. J. (1980). The structure and evolution of the human β-globin gene family. *Cell* 21, 653–668.
- Fire, A., Baker, C. C., Manley, J. L., Ziff, E. B., and Sharp, P. A. (1981). *In vitro* transcription of adenovirus. *J. Virol.* 40, 703–719.
- Fried, M., and Crothers, D. M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl. Acids Res.* 9, 6505–6525.
- Galson, D., and Housman, D. E. (1988). Detection of two tissue-specific DNA-binding proteins with affinity for sites in the mouse β-globin intervening sequence 2. *Mol. Cell Biol.* 8, 381–392.
- Garner, M. M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions. Applications to components of the E. coli lactose operon regulatory system. *Nucl. Acids Res.* 9, 3047–3060.
- Graves, B. J., Johnson, P. F., and McKnight, S. L. (1986). Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* 44, 565–576.
- Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.-M., Argos, P., and Chambon, P. (1986). Human estrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature* 320, 134–139.
- Guarente, L., Lalonde, B., Gifford, P., and Alani, E. (1984). Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* 36, 503–511.

- Hahn, S., and Guarente, L. (1988). Yeast HAP2 and HAP3: transcriptional activators in a complex. *Science*, in press.
- Hahn, S., Pinkham, J., Wei, R., Miller, R., and Guarente, L. (1988). The HAP3 regulatory locus of *Saccharomyces cerevisiae* encodes divergent overlapping transcripts. *Mol. Cell. Biol.* 8, 655–663.
- Hatamochi, A., Paterson, B., and de Crombrugge, B. (1986). Differential binding of a CCAAT DNA binding factor to the promoters of the mouse $\alpha 2(I)$ and $\alpha 1(III)$ collagen genes. *J. Biol. Chem.* 261, 11310–11314.
- Hollenberg, S. M., Weinberger, C., Ong, E. S., Cerelli, G., Oro, A., Lebo, R., Thompson, E. B., Rosenfeld, M. G., and Evans, R. M. (1985). Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* 318, 635–641.
- Jones, K. A., Yamamoto, K. R., and Tjian, R. (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. *Cell* 42, 559–572.
- Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J., and Tjian, R. (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* 48, 79–89.
- Kingston, R. E., Baldwin, A. S., and Sharp, P. A. (1985). Transcription control by oncogenes. *Cell* 41, 3–5.
- Klock, G., Strahle, U., and Schutz, G. (1987). Estrogen and glucocorticoid responsive elements are closely related but distinct. *Nature* 329, 734–736.
- Knight, G. B., Gudas, J. M., and Pardee, A. B. (1987). Cell-cycle-specific interaction of nuclear DNA-binding proteins with a CCAAT sequence from the human thymidine kinase gene. *Proc. Natl. Acad. Sci. USA* 84, 8350–8354.
- Leegwater, P. A., Van Driel, W., and Van der Vliet, P. C. (1985). Recognition site of nuclear factor I, a sequence-specific DNA-binding protein from HeLa cells that stimulates adenovirus DNA replication. *EMBO J.* 4, 1515–1521.
- Leegwater, P. A., Van der Vliet, P. C., Rupp, R. A., Nowock, J., and Sippel, A. E. (1986). Functional homology between the sequence-specific DNA-binding proteins nuclear factor I from HeLa cells and the TGGCA protein from chicken liver. *EMBO J.* 5, 381–386.
- Loosfelt, H., Atger, M., Misrahi, M., Guiochon-Mantel, A., Meriel, C., Logeat, F., Benarous, R., and Milgrom, E. (1986). Cloning and sequence analysis of rabbit progesterone-receptor complementary DNA. *Proc. Natl. Acad. Sci. USA* 83, 9045–9049.
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). *Molecular Cloning. A Laboratory Manual*. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).
- Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987). Regulation of inducible and tissue-specific gene expression. *Science* 236, 1237–1244.
- Manley, J. L., Fire, A., Cano, A., Sharp, P. A., and Geffer, M. L. (1980). DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. USA* 77, 3855–3859.
- Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Meth. Enzymol.* 65, 499–560.
- McDonnell, D. P., Mangelsdorf, D. J., Pike, J. W., Haussler, M. R., and O'Malley, B. W. (1987). Molecular cloning of complementary DNA encoding the avian receptor for vitamin D. *Science* 235, 1214–1217.
- McKnight, S., and Tjian, R. (1986). Transcriptional selectivity of viral genes in mammalian cells. *Cell* 46, 795–805.
- Mellon, P., Parker, V., Gluzman, Y., and Maniatis, T. (1981). Identification of DNA sequences required for transcription of the human $\alpha 1$ -globin gene in a new SV-40 host-vector system. *Cell* 27, 279–288.
- Miksicek, R., Borgmeyer, U., and Nowock, J. (1987). Interaction of the TGGCA-binding protein with upstream sequences is required for efficient transcription of mouse mammary tumor virus. *EMBO J.* 6, 1355–1360.
- Morgan, W. D., Williams, G. T., Morimoto, R. I., Greene, J., Kingston, R. E., and Tjian, R. (1987). Two transcriptional activators, CCAAT-box-binding transcription factor and heat shock transcription factor, interact with a human hsp70 gene promoter. *Mol. Cell Biol.* 7, 1129–1138.
- Morgan, J. G., Courtois, G., Fourel, G., Chodosh, L. A., Campbell, L., Evans, E., and Crabtree, G. R. (1988). SP1, a CAAT-binding factor and the adenovirus major late promoter transcription factor interact with functional regions of the γ -fibrinogen promoter. *Mol. Cell Biol.*, in press.
- Myers, R. M., Tilly, K., and Maniatis, T. (1986). Fine structure genetic analysis of a β -globin promoter. *Science* 232, 613–618.
- Nagata, K., Guggenheimer, R. A., and Hurwitz, J. (1983). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc. Natl. Acad. Sci. USA* 80, 6177–6181.
- Oikarinen, J., Hatamochi, A., and de Crombrugge, B. (1987). Separate binding sites for nuclear factor 1 and a CCAAT DNA binding factor in the mouse $\alpha 2(I)$ collagen promoter. *J. Biol. Chem.* 262, 11064–11070.
- Olesen, J., Hahn, S., and Guarente, L. (1987). Yeast HAP2 and HAP3 activators both bind to the *CYC1* upstream activation site, UAS2, in an interdependent manner. *Cell* 51, 953–961.
- Pelkovich, M., Brand, N. J., Krust, A., and Chambon, P. (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature* 330, 444–450.
- Pinkham, J. L., and Guarente, L. (1985). Cloning and molecular analysis of the HAP2 locus: a global regulator of respiratory genes in *S. cerevisiae*. *Mol. Cell. Biol.* 5, 3410–3416.
- Rosenfeld, P. J., and Kelly, T. J. (1986). Purification of nuclear factor I by DNA recognition site affinity chromatography. *J. Biol. Chem.* 261, 1398–1408.
- Rosenfeld, P. J., O'Neill, E. A., Wides, R. J., and Kelly, T. J. (1987). Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. *Mol. Cell. Biol.* 7, 875–886.
- Samuels, M., Fire, A., and Sharp, P. A. (1982). Separation and characterization of factors mediating accurate transcription by RNA polymerase II. *J. Biol. Chem.* 257, 14419–14427.
- Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980). *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* 20, 269–281.
- Singh, H., Sen, R., Baltimore, D., and Sharp, P. A. (1986). A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature* 319, 154–158.
- Strahle, U., Klock, G., and Schutz, G. (1987). A DNA sequence of 15 base pairs is sufficient to mediate both glucocorticoid and progesterone induction of gene expression. *Proc. Natl. Acad. Sci. USA* 84, 7871–7875.
- Tsai, S. Y., Sagami, I., Wang, H., Tsai, M. J., and O'Malley, B. W. (1987). Interactions between a DNA-binding transcription factor (COUP) and a non-DNA binding factor (S300-II). *Cell* 50, 701–709.
- Von der Ahe, D., Renoir, J.-M., Buchou, T., Baulieu, E.-E., and Beato, M. (1986). Receptors for glucocorticosteroid and progesterone recognize distinct features of a DNA regulatory element. *Proc. Natl. Acad. Sci. USA* 83, 2817–2721.
- Weinberger, C., Hollenberg, S. M., Rosenfeld, M. G., and Evans, R. M. (1985). Domain structure of human glucocorticoid receptor and its relationship to the *v-erb-A* oncogene product. *Nature* 318, 670–672.
- Weinberger, C., Thompson, C. C., Ong, E. S., Lebo, R., Gruol, D. J., and Evans, R. M. (1986). The *c-erb-A* gene encodes a thyroid hormone receptor. *Nature* 324, 641–646.
- Wu, B. J., Williams, G. T., and Morimoto, R. I. (1987). Detection of three protein binding sites in the serum-regulated promoter of the human gene encoding the 70-kDa heat shock protein. *Proc. Natl. Acad. Sci. USA* 84, 2203–2207.
- Yamamoto, K. R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.* 19, 209–252.