Human Proto-Oncogene c-*jun* Encodes a DNA Binding Protein with Structural and Functional Properties of Transcription Factor AP-1

DIRK BOHMANN, TIMOTHY J. BOS, ARIE ADMON, TETSUJI NISHIMURA, PETER K. VOGT, ROBERT TJIAN*

Nuclear oncogene products have the potential to induce alterations in gene regulation leading to the genesis of cancer. The biochemical mechanisms by which nuclear oncoproteins act remain unknown. Recently, an oncogene, v-*jun*, was found to share homology with the DNA binding domain of a yeast transcription factor, GCN4. Furthermore, GCN4 and the phorbol ester–inducible enhancer binding protein, AP-1, recognize very similar DNA sequences. The human proto-oncogene c-*jun* has now been isolated, and the deduced amino acid sequence indicates more than 80 percent identity with v-*jun*. Expression of cloned c-*jun* in bacteria produced a protein with sequence-specific DNA binding properties identical to AP-1. Antibodies raised against two distinct peptides derived from v-*jun* reacted specifically with human AP-1. In addition, partial amino acid sequence of purified AP-1 revealed tryptic peptides in common with the c-*jun* protein. The structural and functional similarities between the c-*jun* product and the enhancer binding protein suggest that AP-1 may be encoded by c-*jun*. These findings demonstrate that the proto-oncogene product of c-*jun* interacts directly with specific target DNA sequences to regulate gene expression, and therefore it may now be possible to identify genes under the control of c-*jun* that affect cell growth and neoplasia.

A GROWING NUMBER OF VIRAL AND CELLULAR GENES HAVE been identified as potential cancer genes, collectively referred to as oncogenes (*1*). The cellular homologs of viral (v) oncogenes, the proto-oncogenes or c-oncogenes, act in the control of cell growth and differentiation or mediate intracellular signaling systems (*2, 3*). Implicit in the growth regulatory functions of proto-oncogenes is the potential to induce abnormal growth and cancer as a result of qualitative or quantitative alterations in proto-oncogene expression. Retroviruses can activate the oncogenic potential of proto-oncogenes either by transducing a truncated and mutant form of the gene or by increasing transcription of the cellular gene as a result of integrating viral promoter and enhancer sequences in its vicinity. The products of oncogenes can be classified according to their cellular location into secreted, surface, cytoplasmic, and nuclear oncoproteins (*4*). Secreted, surface and cytoplasmic proteins encompass growth factors and their receptors, G proteins, and cytoplasmic protein kinases. By contrast, a small but distinct subset of proto-oncogenes express proteins that are targeted to the cell nucleus where at least some are thought to act directly as specific trans-activators and regulators of RNA and DNA synthesis (*5*). Although several nuclear oncogenes (such as *myc, ski, myb*, and *fos*) have been isolated and characterized, the biochemical mechanisms by which the product of these genes function to regulate gene expression are not well understood.

We now describe the properties of the oncogene *jun*, and suggest a specific nuclear function for the *jun* product. The *jun* gene is a cell-derived sequence that has been identified as the presumed oncogene of avian sarcoma virus 17 (ASV-17), which induces fibrosarcomas in chickens and transforms chick embryo fibroblasts into spindle-shaped neoplastic cells (*6, 7*). Unlike the oncoproteins of all other avian sarcoma viruses, the *jun* product is not a protein kinase. At the amino acid level, the product *jun* shows an unexpected homology with the yeast transcription factor, GCN4, which regulates the expression of genes involved in amino acid biosynthesis (*8, 9*). The conserved region is restricted to the carboxyl terminal portion of v-*jun*, which has 44 percent identity with the DNA binding domain of GCN4. Thus, it has been proposed that v-*jun* might encode a sequence-specific DNA binding protein (*10*). An important clue that v-*jun* might have a normal cellular counterpart encoding a sequence-specific DNA binding factor, came with the discovery that the core consensus DNA sequence, TGACTCA, recognized by GCN4, is very similar to the binding site of a newly identified human trans-activator protein, termed AP-1 (*11, 12*).

AP-1 was first identified as a transcription factor that binds selectively to enhancer elements in the cis control regions of SV40 and human metallothionein IIA. Recently, a group of polypeptides (ranging in size from 40 to 47 kD) with a major species of 47 kD was purified to 95 percent homogeneity from HeLa cells and shown in vitro to have the DNA binding properties of AP-1 (*13*). In addition, purified AP-1 is capable of stimulating the transcription of SV40 and MTIIA in an in vitro reconstituted reaction (*13*). Interestingly, an analysis of cellular and viral genes responsive to AP-1 revealed that, in each case, the expression of these genes was induced by treatment of cells with phorbol ester tumor promoters, such as TPA (12-*O*-tetradecanoyl phorbol-13-acetate) (*13, 14*). Multiple synthetic copies of the consensus AP-1 binding site are sufficient to confer TPA-inducible enhancer activity, and cells treated with TPA appear to have elevated levels of AP-1 activity.

D. Bohmann, A. Admon and R. Tjian are in the Howard Hughes Medical Institute, Department of Biochemistry, University of California, Berkeley, CA 94720. T. Bos, T. Nishimura, and P. Vogt are in the Department of Microbiology, University of Southern California, School of Medicine, 2011 Zonal Avenue, Los Angeles, CA 90033.

*To whom correspondence should be addressed.

The challenge has been to identify specific biochemical functions associated with nuclear oncogene products. At present, there is no direct evidence indicating a structural or functional relation between AP-1 and *jun*. Therefore, we have carried out a molecular genetic analysis of c-*jun* and a biochemical characterization of the c-*jun* product to test whether or not there is a direct link between *jun* and AP-1. Here, we report the molecular cloning and characterization of c-*jun* from human cells. After DNA sequence analysis, a partial complementary DNA (cDNA) clone and genomic DNA of human c-*jun* have been compared with ASV-17 v-*jun*. A protein encoded by human c-*jun* has been expressed in bacteria, and its sequence-specific DNA binding properties have been characterized and directly compared to the DNA recognition properties of AP-1. Antibodies to v-*jun* protein have been used to determine antigenic cross-reactivity with both mammalian AP-1 and the bacterial protein. Moreover, partial amino acid sequence of tryptic peptides derived from AP-1 have been determined. Our findings suggest that c-*jun* may encode the enhancer binding protein AP-1. The functional and structural analyses of c-*jun* and AP-1 provide evidence for a proto-oncogene with clearly defined biochemical activities involved in gene regulation.

Antibodies to v-*jun* oncoprotein cross-react with purified AP-1. The possibility that v-*jun* encodes a product structurally related to the enhancer binding protein AP-1 was first tested by determining

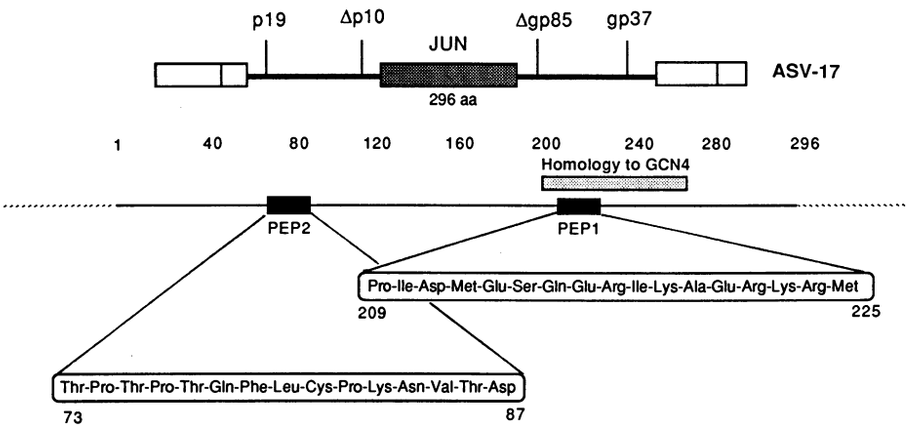

Fig. 1. Physical map of v-*jun* sequences within the ASV-17 genome. (Top) Location of v-*jun* coding sequences (heavily stippled rectangle) in ASV-17. The position of viral genes encoding gag (p19 and Δp10) and viral glycoproteins (Δgp85 and gp37) are demarcated by vertical lines. The open rectangles represent the long terminal report (LTR) sequences of ASV-17. The horizontal line at the center represents an expanded diagram of the v-*jun* coding region (296 amino acids) with the relative positions of the peptides used for raising antibodies to v-*jun* protein indicated by filled-in rectangles (PEP1 and PEP2). The 15–amino acid and 17–amino acid peptide sequences corresponding to PEP1 and PEP2, respectively, are shown at the bottom. The lightly stippled rectangle positioned over amino acid residues 208 to 260 represents the COOH-terminal domain of v-*jun* protein that shows similarity to the DNA binding domain of the yeast transactivator protein, GCN4.

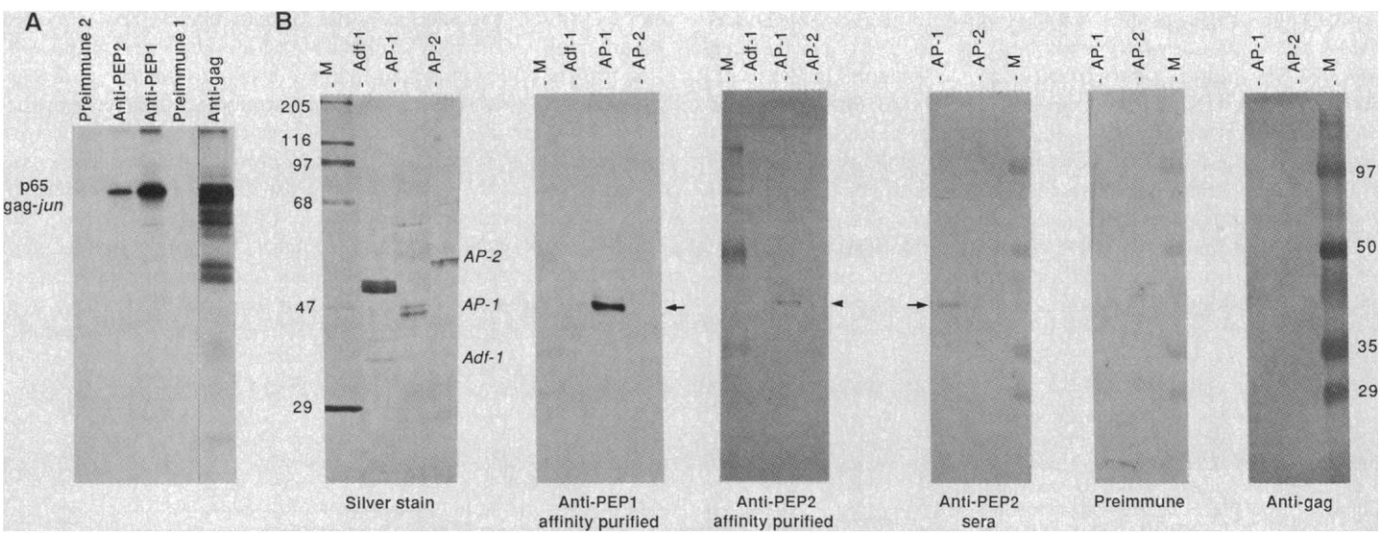

Fig. 2. Antibodies to peptides of v-*jun* cross-react with AP-1. Rabbit antisera to two chemically synthesized peptides derived from v-*jun* (PEP1 and PEP2) (Fig. 1) as described (*15*). (**A**) Antisera to PEP1, PEP2, or the Prague C strain of Rous sarcoma virus (anti-gag), or preimmune sera were used to immunoprecipitate [^{35}S]methionine labeled p65 *gag-jun* fusion proteins synthesized by in vitro transcription and translation of v-*jun* as described (*15*). Immunoprecipitation was done by preincubation of the in vitro translation extracts with preimmune sera and subsequent treatment with *Staphylococcus aureus* cells (pansorbin). The preabsorbed extracts were then incubated with test antisera or preimmune sera, and the specific antigen-antibody complexes were precipitated with pansorbin. The immunoprecipitates were subjected to SDS–polyacrylamide gel electrophoresis (PAGE) and the labeled proteins specifically bound by antibodies were visualized by autoradiography. (**B**) Either crude antisera or peptide affinity chromatography–purified monospecific antibodies to PEP1 and PEP2 of v-*jun* were used to detect cross-reactivity with the enhancer binding protein, AP-1. Samples of AP-1, AP-2 [another enhancer binding protein isolated from HeLa cells that is distinct from AP-1 (*17*)] and Adf-1 [a *Drosophila* DNA binding factor (*18*)] that were partially purified by sequence-specific DNA affinity chromatography (*19*) were subjected to SDS-PAGE, and either the gels were stained or the separated polypeptides were transferred to nitrocellulose by electroblotting. The pattern of polypeptides visualized by silver staining (left) revealed that AP-1 migrates as a doublet of 45 to 47 kD, AP-2 appears as a protein of 52 kD, and Adf-1 as a portion of 35 kD. Similar samples consisting of these different, partially purified DNA binding proteins were transferred to nitrocellulose by electroblotting and treated with various antisera. Crude antisera as well as affinity purified monospecific antibodies to PEP1 and PEP2 were used in the immunoblot assays. The two right panels depict control immunoblot reactions with preimmune sera and antisera to the gag probes. Specific antibody reactivity to AP-1 was visualized by treatment of immunoblot filters with alkaline phosphatase conjugated antibody to rabbit IgG. The lanes marked (M) show the pattern of protein markers (kilodaltons). The markers on the left panel were not prestained and represent a collection of polypeptides visualized by silver staining.

the cross-reactivity of polyclonal antibodies to the v-*jun* oncoprotein to purified human AP-1. Antisera were produced by immunization of rabbits with synthetic peptide antigens derived from two distinct hydrophilic domains of the v-*jun* oncoprotein. One synthetic peptide antigen consisted of a 17–amino acid sequence from the putative DNA binding domain (PEP1) located toward the carboxyl terminal portion of v-*jun* protein (Fig. 1). A second peptide antigen of 15 amino acid residues (PEP2) was generated from a stretch of hydrophilic sequence located near the NH_2-terminal domain of the v-*jun* coding sequence. The specificity of the two antisera to the peptides (anti-PEP1 and anti-PEP2) was tested by immunoprecipitation of v-*jun* synthesized by in vitro transcription and translation (Fig. 2A) and by indirect immunofluorescence. As would be expected, antibodies to these two separate antigenic determinants of v-*jun* specifically immunoprecipitate a protein with an apparent molecular size of 65 kD, which is also efficiently immunoprecipitated by antibodies to the gag protein (Fig. 2A). We have designated this fusion product of *gag-jun* as $P^{65\ gag\text{-}jun}$ even though its calculated size is close to 55 kD. In immunoprecipitation assays and immunoblot reactions, anti-PEP1 consistently displayed a tenfold higher titer than anti-PEP2.

In order to determine whether these two antibodies cross-react with transcription factor AP-1 from human cells, we performed immunoblot analysis with purified or partially purified preparations of AP-1 (Fig. 2B). Briefly, AP-1 was subjected to SDS gel electrophoresis, transferred to nitrocellulose, and reacted with either whole sera or affinity-purified monospecific antibodies directed against the two distinct v-*jun* peptides. Antigenic cross-reactivity was visualized by treatment of the immobilized antibody-antigen complex with alkaline phosphatase–conjugated goat antiserum to rabbit antibody. In the samples containing either partially purified AP-1 or homogeneous AP-1 (*13*), a protein with an apparent size of 47 kD was specifically immunoreactive with both antibodies to *jun*-encoded peptides. By contrast, two control samples containing other sequence-specific DNA binding proteins (AP-2 and Adf-1) that had been partially purified by the same DNA affinity procedures (*17–19*) used for AP-1 did not cross-react with either of the two antibodies to *jun* protein. Preimmune sera and antibodies to gag did not react with AP-1 or control proteins. These data indicate that AP-1 and the v-*jun* oncoprotein, $P^{65\ gag\text{-}jun}$, share at least two distinct antigenic determinants, including an epitope overlapping the carboxyl terminal domain containing the putative DNA binding region and an epitope located toward the NH_2-terminal protein of v-*jun*. Thus, the structural homology between enhancer binding protein AP-1 and the v-*jun* oncogene product is not limited to the putative DNA binding domain, which is homologous to GCN4. The extensive homology suggests the possibility that the cellular counterpart of v-*jun* in humans cells may encode AP-1 or a closely related protein. A molecular genetic characterization of the human c-*jun* gene should help establish a structural and functional relation between v-*jun* and AP-1.

Isolation of the human proto-oncogene c-*jun*. Human DNA sequences containing regions of homology to v-*jun* were isolated by screening a genomic DNA library of recombinant lambda phage with a hybridization probe consisting of a DNA fragment extending from nucleotide 720 to 1601 derived from the v-*jun* clone (*7*). Three lambda recombinants, each bearing approximately 15 kb of human DNA sequences, were isolated, and a partial restriction endonuclease cleavage map of the putative human c-*jun* genomic DNA was determined (Fig. 3A). The regions of the recombinant clones containing sequences homologous to v-*jun* were identified by DNA blot hybridization of human DNA fragments cleaved by different restriction enzymes (Fig. 3A).

In order to determine the extent of homology between v-*jun* and the putative c-*jun* DNA, the nucleotide sequence of a 1.5-kb stretch of cloned human genomic DNA that spanned the v-*jun* homology was determined by the M13 dideoxy sequencing technique. Conceptual translation of these DNA sequences revealed a single, long open reading frame of approximately 340 codons, that displayed more than 80 percent identity at the amino acid level when aligned with the sequence of v-*jun* protein (Fig. 4). A schematic representation of the amino acid sequence homology between v-*jun* and human c-*jun* proteins is shown in Fig. 3B. The NH_2-terminal amino acid residue of the c-*jun* product has not been identified. Within the first 149–amino acid sequence, the human cellular DNA product contains an insertion of 27 amino acid residues (between amino acid 31 and 57) and 18 amino acid changes relative to that of v-*jun*. The highly conserved COOH-terminal 118–amino acid sequence of the human gene contains only two alterations when compared to v-*jun*. The central 73 amino acids are the least homologous between v-*jun* and the human DNA, showing approximately 53 percent identity with the corresponding 59 amino acids of v-*jun*, compared to >90 percent identity in the rest of the sequence. This central region of c-*jun* contained several short insertions and amino acid changes relative to v-*jun*. A survey of amino acids between residues 207 and 260 indicates that approximately 50 percent are Gln or Pro residues.

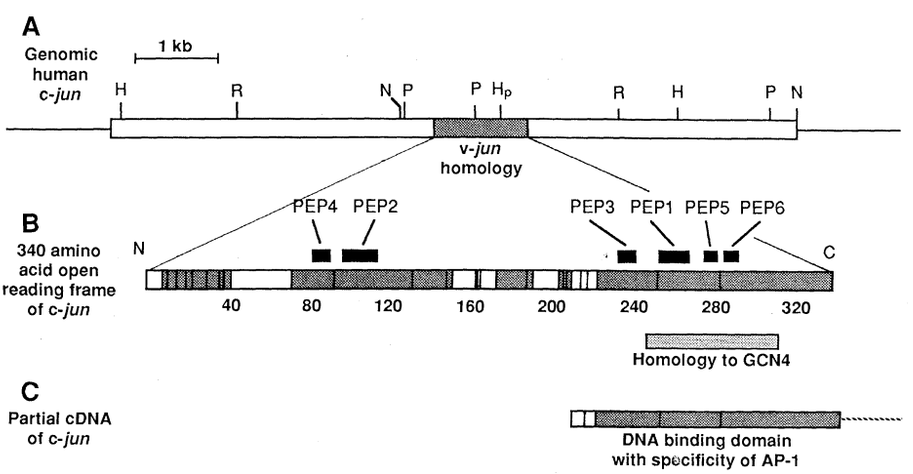

Fig. 3. Physical map of the human c-*jun* clone. (**A**) A partial restriction endonuclease map of approximately 9 kb of human genomic DNA containing the region of homology to v-*jun* (dark stippled rectangle). Restriction endonuclease cleavage sites for Eco RI (R), Hind III (H), Nco I (N), Pst I(P), and Hpa I (Hp) are shown. (**B**) A schematic representation of the c-*jun* clone containing 340 amino acids of open reading frame that shows extensive homology to v-*jun*. The darkly stippled regions represent amino acid identities between v-*jun* and c-*jun*, verticle lines represent single amino acid changes, and open rectangles indicate regions of patchy similarity or insertions in c-*jun* relative to v-*jun*. The positions of PEP1 and PEP2 sequences used to raise antibodies against v-*jun* and the sequence of AP-1 tryptic peptides PEP3, PEP4, PEP5, and PEP6 are indicated by closed rectangular boxes. The region of homology of GCN4 is shown as a lightly stippled rectangle. The sequences of c-*jun* contained within a partial cDNA clone have been aligned with the genomic DNA and are shown in (**C**). The heavily dashed line in (C) represents the noncoding 3′ end of the cDNA that was sequenced and determined to contain a stop codon for c-*jun*. The genomic library was generated with DNA isolated from a human T cell line. The cDNA library was derived from HeLa mRNA.

A similar stretch of sequences containing a high content of Pro and Gln residues has been identified between the DNA binding and the steroid binding domains of the mineralocorticoid receptor (*20*), where it is thought to serve as an intramolecular hinge. Similar flexible hinge regions have been proposed for the glucocorticoid (*21*) and estrogen (*22*) receptors.

A comparison of the human sequence to a chicken cellular homolog of v-*jun* indicates that both cellular c-*jun* sequences contain the 27–amino acid insertion as well as the two changes in the highly conserved COOH-terminal domain relative to v-*jun*. Inspection of the v-*jun* sequences when aligned with both human and chicken c-*jun* suggests that at least 1 kb of the genomic DNA region that has been sequenced represents an uninterrupted coding domain. The only possible intron present within this region would be the 81-bp insertion corresponding to the 27 amino acids present in c-*jun* but not v-*jun* protein. However, this extra DNA in the genomic sequences is unlikely to be an intron because it is a highly conserved open reading frame in both chicken and man. Moreover, there are no compelling splice acceptor or donor consensus sequences that would account for an in-frame splice junction. The sequence analysis of c-*jun* provided convincing evidence that a cellular homolog of v-*jun* had been isolated, and that the product of this human proto-oncogene was structurally closely related to the v-*jun* oncoprotein. The next step is to obtain direct evidence for a structural and functional link between human c-*jun* and the transcription factor, AP-1.

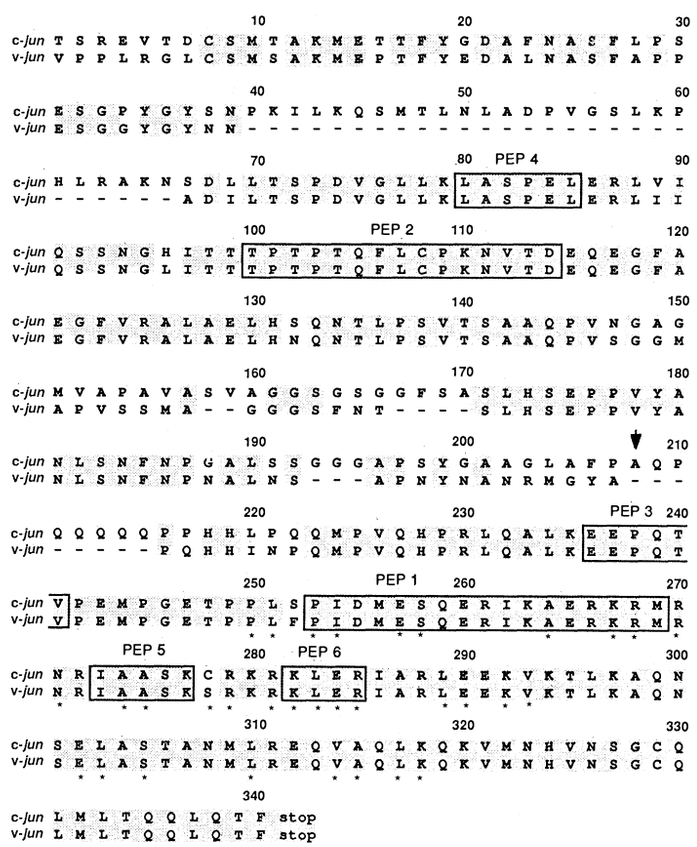

Fig. 4. Comparison of the predicted amino acid sequence of the human c-*jun* protein with that of v-*jun* protein and alignment of partial amino acid sequence of AP-1 peptides showing 340 amino acids of c-*jun* protein open reading frame and 296 amino acids of v-*jun* protein. Amino acids that are identical between v-*jun* and c-*jun* have been shaded with light stippling. The 15– and 17–amino acid residue peptides that are boxed represent PEP2 and PEP1, respectively. Partial amino acid sequence of the tryptic peptides derived from purified AP-1 are indicated by boxes residues marked PEP3, PEP4, PEP5 and PEP6. Amino acids marked with an asterisk represent common residues in v-*jun* and the yeast transcription factor, GCN4. The first nucleotide of the 5′ end of the cDNA clone is demarcated by a vertical arrow.

Partial amino acid sequence of AP-1. The ability of monospecific antibodies to PEP1 and PEP2 of v-*jun* product (Figs. 1 and 2) to recognize AP-1 polypeptides suggested that the product of c-*jun* shared significant amino acid sequence identity with AP-1. As an independent test of the structural identity between AP-1 and c-*jun* protein, we have determined the partial amino acid sequence of AP-1 purified from human cells. Attempts to obtain NH_2-terminal sequence of AP-1 failed, perhaps because of covalent modification and blockage of the terminal residue. In order to circumvent this problem, we generated tryptic peptide fragments of AP-1 and obtained short stretches of amino acid sequences from purified peptides derived from AP-1 (*23*). Two sequences of six residues (EEPQTV is PEP3; LASPEL is PEP4 (*24*) from AP-1 corresponds to amino acids 236 to 241 located toward the COOH-terminal region of c-*jun* product and amino acids 80 to 85 which lie toward the NH_2-terminal portion of c-*jun* protein. Additional AP-1 peptide sequences of five and four residues (IAASK is PEP5; KLER is PEP6) were identical to amino acids sequences in the putative DNA binding domain of c-*jun*. The position of these AP-1 peptides relative to the sequence of c-*jun* and v-*jun* are indicated in Figs. 3B and 4. Each of these peptides is preceded by a lysine or arginine residue in the c-*jun* sequences, which is consistent with the cleavage specificity of trypsin used to generate the AP-1 peptides. These findings, taken together with the antibody cross-reactivity data (Fig. 2), provide direct evidence for a close structural relationship between AP-1 polypeptides and the product of the proto-oncogene c-*jun*.

Isolation of a partial human c-*jun* complementary DNA. Several cDNA clones were isolated from a λgt10 recombinant phage library, constructed with cDNA synthesized from HeLa messenger RNA (mRNA). Analysis of endogenous human mRNA revealed a species of 2.7 kb that hybridized specifically to probe DNA consisting of a restriction fragment derived from c-*jun*. This same hybridization probe was used to screen 1.5×10^6 recombinant phage of the cDNA library, and three positive clones were isolated. The longest cDNA clone characterized thus far contained approximately 1 kb of human DNA. The nucleotide sequence of this partial cDNA was determined and the amino acid sequence derived from its only extended open reading frame corresponded to the carboxyl terminal 133 amino acid residues of c-*jun*. The position of the cDNA relative to the genomic map is shown in Figs. 3C and 4. The entire cDNA fragment, bounded by endonuclease Eco RI cleavage sites, was excised from the recombinant phage and inserted into a bacterial plasmid vector so that the open reading frame would be expressed as a hybrid protein containing 133 amino acids of c-*jun* product fused to 39 amino acids of vector sequences at the NH_2-terminus. Bacteria transformed with this plasmid construct (pHJ19) are expected to produce a fusion protein with an apparent molecular weight of 19 kD.

Purification and DNA binding properties of the c-*jun* fusion protein. Extracts prepared from bacteria transformed with pHJ19 or a control extract derived from bacteria carrying the pBSSK vector alone were prepared and subjected to DNA affinity chromatography (*19*). The DNA affinity resin contained tandem copies of a 17-bp AP-1 recognition sequence covalently attached to Sepharose. The bacterial extracts were first mixed with an appropriate amount of nonspecific competitor DNA and then applied to the affinity resin under salt and buffer conditions that allow nonspecific DNA binding proteins and other bacterial proteins to flow through the sequence-specific affinity resin. After the bound protein was eluted with buffer containing 1*M* KCl, the purified DNA binding protein was analyzed by SDS gel electrophoresis and immunoblotting (Fig.

5). A predominant polypeptide species of 19 kD was recovered from DNA affinity purification of the pHJ19 extract. In addition, a minor, nonspecific DNA binding protein (~66 kD) was detected after chromatography of both pHJ19 and control pBSSK extracts (Fig. 5A). Immunoblot analysis revealed that the 19-kD protein is specifically recognized by affinity purified antibody directed against the COOH-terminal DNA binding domain of v-*jun* protein (Fig. 5B). These results unambiguously identify the 19-kD DNA binding protein as the product of c-*jun* and establish its structural relatedness to both the v-*jun* oncoprotein and the transcription factor, AP-1. We refer to it as P^{19jun}.

A direct comparison of the DNA binding specificity of P^{19jun} and AP-1 was carried out with purified proteins in deoxyribonuclease I (DNase I) footprint protection experiments. We used 5′ end-labeled DNA fragments derived from the SV40 72-bp enhancer region as probes. Both AP-1 and P^{19jun} protected precisely the same recognition sequence within the SV40 enhancer (Fig. 6). Samples from control reactions containing the nonspecific DNA binding protein eluted from the affinity resin showed no activity in the DNase protection. The DNase I footprint patterns observed with AP-1 and P^{19jun} were virtually indistinguishable with probes labeled on either the coding or noncoding strands of the SV40 DNA. Only subtle differences were observed in enhanced cleavage sites that most likely reflect differences in the size of the two proteins leading to differential interaction with DNase I. The specificity of binding by AP-1 and P^{19jun} was further refined by assaying a series of mutant templates bearing single base substitutions that had been shown to influence the binding site affinity for both AP-1 and the yeast transcription factor GCN4. DNase I footprint analysis of these altered templates indicates that P^{19jun} behaved in precisely the same manner as AP-1 (Table 1). In particular, a single base substitution that led to a higher affinity binding site for AP-1 also bound P^{19jun} more efficiently, whereas base changes that decreased binding of AP-1 concomitantly reduced binding of P^{19jun}. These results provide direct evidence that human c-*jun* encodes a DNA binding protein with recognition properties indistinguishable from transcription factor AP-1.

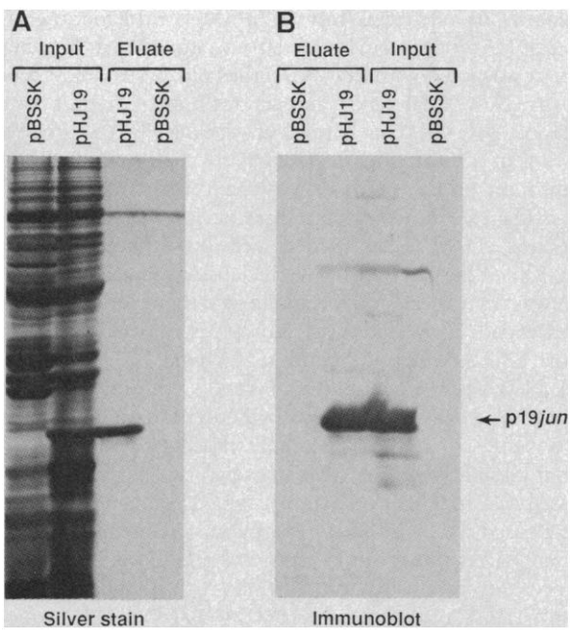

Fig. 5. Purification of P^{19jun} by DNA affinity chromatography. (**A**) Lysozyme extracts (0.8 ml, that is 23 mg of protein) from pHJ19 and pBSSK transformed bacteria (*35, 38, 40*) where mixed with 1 absorbancy unit of sonicated poly(dI-dC) in 100 μl of 100 m*M* NaCl, 10 m*M* tris (*p*H 7.5), 1 m*M* EDTA each and incubated on ice for 10 minutes. The material was centrifuged at 100,000*g* at 4°C for 20 minutes in the TL100.3 rotor of the Beckman tabletop ultracentrifuge. The supernatant fractions were processed on AP-1–specific DNA affinity columns (*13, 19*). Washing and elution was performed as described (*13*), except that bound protein was eluted from the column with 1*M* KCl. Samples, 3 μl of the first milliliter of eluate and 2 μl of starting material ("load"), were analyzed by gel electrophoresis on a 10 percent SDS polyacrylamide gel and silver staining. (**B**) A gel similar to the one shown in (A) was prepared and the proteins were transferred onto a nitrocellulose filter by electroblotting. The immobilized proteins were analyzed for cross-reactivity with anti-PEP1 as described (Fig. 2).

Does c-*jun* encode AP-1? Several lines of evidence establish that the human c-*jun* proto-oncogene encodes a protein closely related to the enhancer binding protein, AP-1, in both structure and function. First, monospecific antibodies against two distinct regions of v-*jun* cross-react specifically with AP-1 polypeptides (40 to 47 kD) purified from human cells. Sequence analysis confirms that both hydrophilic peptides of v-*jun* chosen for antibody production are completely conserved in c-*jun*. Partial amino acid sequence of peptide fragments generated by trypsin digestion of purified AP-1 correspond to peptides deduced from the nucleotide sequence of c-*jun*. Moreover, the product of a partial c-*jun* cDNA clone expressed in bacteria displays the sequence-specific DNA binding properties of AP-1. These findings, taken together, suggest that AP-1 may itself be encoded by c-*jun*. However, there are several issues that need to be addressed before such a conclusion can be drawn.

AP-1 has been defined as a cellular sequence-specific DNA binding protein that recognizes the enhancer element, TGACTCA (*11, 13, 14*). DNA affinity purified AP-1 consists predominantly of a 47-kD species that has been unambiguously identified as an active DNA binding species (*13*). In addition, highly purified AP-1 preparations also contain minor polypeptides of 60 kD, 45 kD, and 40 kD. Although these minor protein species copurify with the 47-kD AP-1 polypeptide, they have not been tested and shown individually to bind DNA in a sequence-specific manner. However, they are consistently found in DNA affinity purified samples, suggesting that they bind DNA with the same specificity as the 47-kD species. Multiple polypeptides present in highly purified preparations of AP-1 may be the result of proteolytic degradation or covalent modification of a single species. Alternatively, some of these forms may represent distinct cellular proteins encoded by different mRNA's or genes (or both), but each one displaying the same or similar DNA binding specificity. This latter possibility is of interest because if additional gene products in the human cell recognize and bind the TGACTCA enhancer element, then the DNA binding properties of AP-1 cannot be used as a strict functional criteria to define the product of a single gene.

AP-1 constitutes a family of closely related DNA binding proteins. Evidence has accumulated suggesting that multiple distinct factors can recognize the same cis regulatory enhancer or promoter elements. For example, there are at least two biochemically distinct "octamer" binding proteins, one that is ubiquitous in mammalian cells and one that is found in B lymphoid cell types (*25, 26*). Similarly, there is evidence suggesting that the "core" enhancer element of SV40 can be recognized by more than one cellular factor (*17, 27, 28*). Therefore, it seems prudent to anticipate that AP-1, as it is defined at present, may actually represent different proteins with indistinguishable DNA binding properties. If so, it is possible that there exists a small gene family encoding related proteins, all displaying the same DNA binding specificity of AP-1. There is little doubt that c-*jun* encodes a protein that would be a member of such an AP-1 gene family of enhancer binding proteins. However, it is less clear whether there are in fact additional members of this hypothetical gene family in human cells. To test directly such a possibility, we performed Southern blot analysis of total human

Table 1. Point mutants of the AP-1 recognition sequence have the same effect on binding of AP-1 and P^{19jun}. DNase protection analysis was carried out with wild-type and mutant probes derived from the *Saccharomyces cerevisiae his*3 promoter region bearing the AP-1 recognition site (*8, 12*). Relative binding efficiency of purified AP-1 and P^{19jun} was determined under conditions that allow half-maximal protection of the wild-type binding site. The numbers indicate the extent of protection against nuclease digestion of the different mutant probes relative to wild type as determined by densitometric scanning of the autoradiographs.

Mutant	Sequence	Binding	
		AP-1	P^{jun19}
WT	ATGACTCTT	1.0	1.0
193	**C**TGACTCTT	1.0	0.9
165	A**A**GACTCTT	<0.1	<0.1
166	AT**T**ACTCTT	<0.1	<0.1
168	ATGAC**G**CTT	1.0	0.8
189	ATGACTC**A**T	1.7	1.5
171	ATGACTCT**C**	1.2	1.3

genomic DNA and used restriction fragments of c-*jun* (Fig. 7) and v-*jun* as hybridization probes and, under stringent hybridization conditions, only the c-*jun* cognate sequences are detected. However, hybridization under moderate or low stringency conditions detected at least one additional gene sequence distinct from c-*jun*. These hybridization experiments confirm that the human sequence we have identified as c-*jun* are indeed the most closely related to v-*jun*. However, these results also indicate that c-*jun* may not be the only human gene sequence that shares homology with v-*jun*. There is at least one other c-*jun*–like sequence, albeit less homologous to v-*jun* than c-*jun* in the human genome which is contained within a 7.3-kb Eco RI fragment. A gene encoding a serum-induced product has been isolated and characterized from mouse cells (*29*). A comparison of the amino acid sequence derived from the serum-induced gene revealed a high degree of sequence similarity to v-*jun* particularly in the putative DNA binding domain. DNA blot analysis indicates that this new v-*jun*–related sequence is contained within a 12-kb Eco RI restriction fragment of human DNA rather than the 5-kb c-*jun* fragment (*29*). These findings, together with our results, suggest that there are additional genes in mammalian cells encoding proteins that share structural similarities with the DNA binding region of v-*jun*. Thus, it is likely that there are multiple distinct cellular factors that can bind to the AP-1 recognition site. It will be of interest to study the DNA binding and transcriptional activation properties of the other members of the AP-1 family to determine whether or not some forms of AP-1 are distributed, expressed, and utilized differentially in a tissue-specific or developmentally stage-dependent manner.

It is not entirely surprising that there is more than one gene in human cells encoding a protein with DNA binding properties of AP-1. For example, the very distantly related GCN4 protein in yeast recognizes precisely the same binding sequences as the human transcription factor, AP-1 (*8, 11, 12*). In addition, it has been reported that the DNA binding domain of GCN4 can be replaced by the carboxyl terminal region of v-*jun* protein and retain functional binding specificity in yeast (*30*). In contrast to AP-1, the structural homology between GCN4 and *jun* is more limited with virtually no identity in the NH_2-terminal 250–amino acid residues and only patches of homology in the more conserved COOH-terminal 60–amino acid DNA binding domain. The transcriptional activation domain of GCN4, which consists of a highly acidic amino acid sequence is not conserved in *jun* protein. Thus, although both proteins bind DNA with similar specificity and are involved in transcriptional regulation, there is no evidence that they govern similar control pathways in the cell or operate by analogous mechanisms. The conservation of DNA binding domains between GCN4 and AP-1 is perhaps more reminiscent of homeo box (*31*) and Zn finger (*32*) containing proteins that have been observed in *Drosophila*, amphibians and mammals. However, individual members of each of these two groups are likely to be involved in quite disparate physiological functions, linked by the usage of a common DNA binding structural motif.

The region of *jun* or AP-1 that is responsible for DNA recognition does not resemble any of the well-defined DNA binding structural motifs such as the lambda repressor type helix-turn-helix (*33*), the TFIIIA Zn fingers (*32*) or the Eco RI helix bundle (*34*). Two other genes encoding mammalian RNA polymerase II transcription factors were isolated and the primary amino acid sequence deduced from nucleotide sequence indicates that one of them, Sp1, contains three Zn fingers (*35*), but the other factor, CTF (*36*), contains a structure that is apparently distinct from either AP-1(*jun*), Sp1, or any of the other known structural domains associated with sequence-specific DNA binding. As more sequence-specific transcriptional regulatory factors such as AP-1(*jun*) are characterized, it may be possible to distinguish many different types of protein structures involved in DNA recognition and thus gain new insights concerning the molecular basis of productive transactions between regulatory proteins and DNA.

The neoplastic potential of v-*jun* in avian cells has been well documented (*6, 7*), but the effects of the *jun* oncogene on growth regulation and transformation in mammalian cells remains to be determined. It is also important to discover how the expression or function of the proto-oncogene c-*jun* can be distorted to become an activated oncogene. Preliminary evidence suggests that the sequence-specific DNA binding properties of v-*jun* protein are not drastically altered from that of c-*jun* protein and AP-1 (*37*). Thus, other aspects of c-*jun* function or expression may trigger its oncogenic potential. For example, AP-1(*jun*) is likely to interact

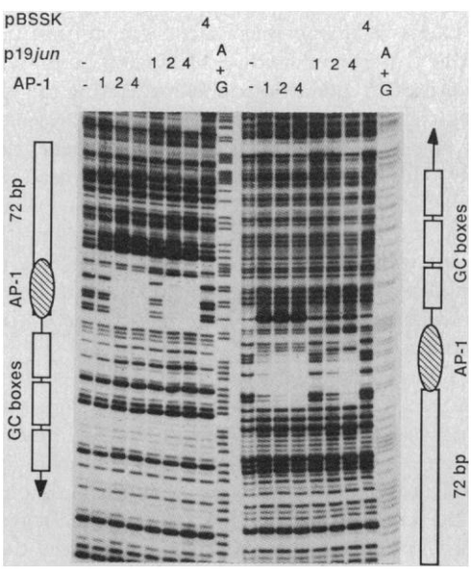

Fig. 6. Comparison of AP-1 and P^{19jun} binding to the SV40 enhancer by DNase I protection analysis. The 5′ labeled fragments used were the Nco I*–Asp718 fragment of pSVGCO (*41*) (SV40 coordinates 37 to 294, +strand) and the Hind III–Asp718* fragment (coordinates 5171 to 294, −strand). Increasing amounts of purified AP-1 from HeLa cells and of purified P^{19jun} (approximately 10, 20, and 40 ng) were used for the binding assay. Control reactions with either no added protein or an amount of the pBSSK eluate equivalent to P^{19jun} used for lane 4 are shown. Partial purine-specific cleavage products of the probes (A+G) are run along with the DNase I digestion products.

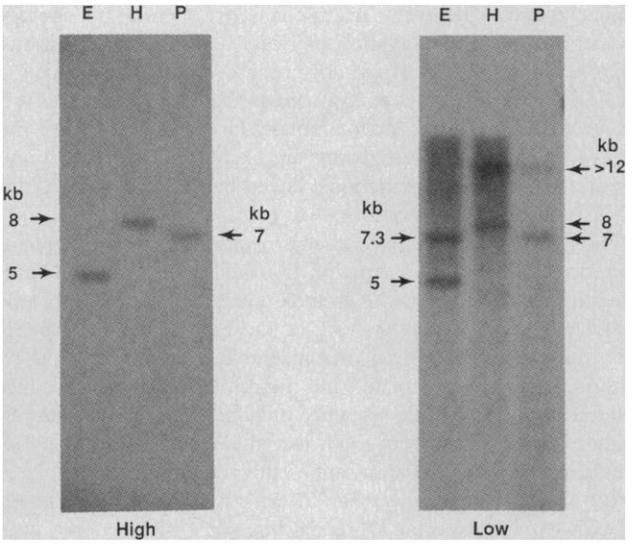

Fig. 7. High and low stringency DNA blot analysis of total human genomic DNA. Human placenta DNA was cleaved with Eco RI (E), Hind III (H), and Pst I (P). Two samples of each reaction were separated on a 1 percent agarose gel and transferred onto GeneScreen so that two identical blots were obtained. The immobilized DNA was analyzed by hybridization under high or low stringency conditions with a radioactive DNA probe that was prepared from the 1-kb cDNA fragment described in Fig. 3. High stringency hybridization was carried out in 0.2*M* sodium phosphate (*p*H 7.2), 1 m*M* EDTA, 7 percent SDS, 1 percent bovine serum albumin, and 15 percent formamide at 65°C. Low stringency hybridization was performed under the same conditions except that no formamide was present. Stringent washing conditions consisted of 0.2× SST [20× SST is 3*M* NaCl, 0.3*M* tris (*p*H 7.5), 50 m*M* EDTA] and 0.1 percent SDS at 65°C. For low stringency washing, 2× SST and 0.1 percent SDS at 50°C was used. The sizes of the genomic DNA fragments hybridizing to c-*jun* are indicated in kilobase pairs.

with specific enhancer binding sites that may in turn lead to protein-protein contacts with other components of the transcriptional apparatus, such as RNA polymerase and general initiation factors. The oncogenic change could have a direct effect on the specificity of these protein-protein interactions, thereby altering the spectrum of target genes that c-*jun* or AP-1 governs. At present, we have identified a small collection of unrelated cellular genes that contain AP-1 binding sites such as metallothionein, collagenase, α_1-antitrypsin, transthyretin, and MHC H-2D^d (*28*). There is no evidence that any of the known AP-1 responsive genes is involved in the pathway leading to oncogenesis. However, it may be possible to identify other potential target genes that influence the growth state and neoplastic potential of cells as a result of altered patterns of expression induced by *jun*-related oncoproteins.

Oncogenes in general appear to encode cellular constituents that regulate growth. When such growth regulatory signals are altered, specific changes in gene expression may result and induce neoplastic transformation. The most direct effectors of change in gene expression are likely to be transcription factors and promoter-selective DNA binding proteins. The discovery of c-*jun* and its relationship to AP-1 provides strong support for the notion that nuclear oncoproteins may in fact directly govern the expression and transcription of certain target cellular genes by binding to enhancer and promoter sequences. Treatment of cells with phorbol ester tumor promoters selectively activates certain promoters and enhancers carrying AP-1 binding sites (*13, 14*). Thus, AP-1 regulates transcription directly, and is itself influenced by cellular signals involved in growth control and neoplasia. The finding that transcription factor activities may be modulated through changes induced by effector molecules such as TPA suggests a complex and potentially important relation between nuclear oncogenes and cellular growth regulatory signals.

REFERENCES AND NOTES

1. J. M. Bishop, *Annu. Rev. Biochem.* **52**, 301 (1983).
2. H. Varmus, *Annu. Rev. Genet.* **18**, 353 (1984).
3. J. M. Bishop, *Science* **235**, 305 (1987).
4. R. A. Weinberg, *ibid.* **230**, 770 (1985).
5. R. E. Kingston, A. S. Baldwin, P. A. Sharp, *Cell* **41**, 3 (1985).
6. F. Cavalieri *et al.*, *Virology* **143**, 680 (1985).
7. Y. Maki, T. J. Bos, C. Davis, M. Starbuck, P. K. Vogt, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2848 (1987).
8. I. A. Hope and K. Struhl, *Cell* **43**, 177 (1985).
9. K. Arndt and G. Fink, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8516 (1986).
10. P. K. Vogt, T. J. Bos, R. F. Doolittle, *ibid.* **84**, 3316 (1987).
11. W. Lee, A. Haslinger, M. Karin, R. Tjian, *Nature (London)* **325**, 368 (1987).
12. D. Bohmann, K. Struhl, R. Tjian, unpublished observations.
13. W. Lee, P. Mitchell, R. Tjian, *Cell* **49**, 741 (1987).
14. P. Angel *et al.*, *ibid.*, p. 729.
15. T. Bos, H. Tsuchie, W. Uckert, P. K. Vogt, in preparation.
16. I. A. Hope and K. Struhl, *Cell* **43**, 177 (1985).
17. P. Mitchell, C. Wang, R. Tjian, *ibid.* **50**, 847 (1987).
18. U. Heberlein, B. England, R. Tjian, *ibid.* **41**, 965 (1985).
19. J. T. Kadonaga and R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5889 (1986).
20. J. L. Arriza *et al.*, *Science* **237**, 268 (1987).
21. V. Giguère, S. M. Hollenberg, M. G. Rosenfeld, R. M. Evans, *Cell* **46**, 645 (1986).
22. A. Krust *et al.*, *EMBO J.* **5**, 891 (1986).
23. Approximate 600 pmol of AP-1 was purified by successive passages over a sequence-specific DNA affinity column containing tandemly ligated copies of the AP-1 binding sequence

    GATCGTGACTCAGCGCG
    CACTGAGTCGCGCCTAG

    Further purification of the AP-1 polypeptides was achieved by C8 reversed-phase high-performance liquid chromotography. The highly purified AP-1 polypeptides were subjected to digestion with trypsin in the presence of 2*M* urea. The peptide fragments generated by trypsin digestion were applied to a C18 reversed-phase column to isolate individual peptides. The amino acid sequence of various purified AP-1 peptides was determined by sequential Edman degredation (Applied Biosystems, ABI, 477A pulsed liquid sequencer) and amino acids were identified ABI 120A PTH amino acid analyzer.
24. Abbreviations for the amino acid residues are: A, Ala; C, Cyś; D, Asp; E. Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
25. L. M. Staudt, H. Singh, R. Sen, T. Wirth, P. A. Sharp, D. Baltimore, *Nature (London)* **323**, 640 (1986).
26. T. Gerster, P. Matthias, M. Thali, J. Jiricny, W. Schaffner, *EMBO J.* **5**, 1323 (1987).
27. P. F. Johnson, W. H. Landschulz, B. J. Graves, S. L. McKnight, *Genes Devel.* **1**, 133 (1987).
28. N. Mermod, T. Williams, R. Tjian, unpublished results.
29. L. F. Lau and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1182 (1987); K. Ryder, L. F. Lau, D. Nathans, personal communication.
30. K. Struhl, *Cell* **50**, 841 (1987).
31. W. J. Gehring, *Science* **236**, 1245 (1987).
32. J. Miller, A. D. McLachlan, A. Klug, *EMBO J.* **4**, 1609 (1985).
33. C. P. Pabo and R. T. Sauer, *Annu. Rev. Biochem.* **53**, 293 (1984).
34. J. A. McClarin *et al.*, *Science* **234**, 1526 (1986).
35. J. T. Kadonaga, K. R. Carner, F. R. Masiarz, R. Tjian, *Cell*, in press.
36. C. Santoro, P. Andrews, R. Tjian, unpublished results.
37. T. Bos *et al.*, unpublished results.
38. *Escherichia coli* JK50 is hflA::Tn5 ΔlacX74 ara str r_k-m_k+/F′ lac I^QL_8Z::Tn10 (tps$^-$)*praAB* (J. T. Kadonaga and R. Tjian, unpublished data).
39. K. A. Jones, K. R. Yamamoto, R. Tjian, *Cell* **42**, 559 (1985).
40. C. Desplan, J. Theis, P. H. O'Farrell, *Nature (London)* **318**, 630 (1985).
41. D. Gidoni *et al.*, *Science* **230**, 511 (1985).

42. We thank C. Hauser for providing the human cDNA (λgt10) library; C. Croce for a genomic library from a human T cell line, K. Struhl for the *His* 3-GCN4 binding mutants; P. Mitchell, C. Santoro, and J. Kadonaga for cloning expertise, various reagents, and many discussions. T. Williams, N. Mermod, K. Perkins, M. Simon, I. Mattaj, G. Rubin, S. Martin, and A. Levinson for comments and critical reading of the manuscript; and K. Ronan for preparation of the manuscript and figures. Supported in part by NIH grants CA25417 (R.T.) and CA42564 (P.K.V.), a grant from the Council for Tobacco Research (1951) (P.K.V.), and a Deutsche Forschungsgemeinschaft Fellowship (D.B.).

21 September 1987; accepted 2 November 1987