Purification and Biochemical Characterization Of the Promoter-Specific Transcription Factor, Sp1

MICHAEL R. BRIGGS, JAMES T. KADONAGA, STEPHEN P. BELL, ROBERT TJIAN

The biochemical analysis of cellular trans-activators involved in promoter recognition provides an important step toward understanding the mechanisms of gene expression in animal cells. The promoter selective transcription factor, Sp1, has been purified from human cells to more than 95 percent homogeneity by sequence-specific DNA affinity chromatography. Isolation and renaturation of proteins purified from sodium dodecyl sulfate polyacrylamide gels allowed the identification of two polypeptides (105 and 95 kilodaltons) as those responsible for recognizing and interacting specifically with the GC-box promoter elements characteristic of Sp1 binding sites.

IN HIGHER EUKARYOTIC CELLS, REGULATION OF GENE EXpression is often mediated by turning on and off RNA synthesis in a temporally ordered manner. Although the mechanisms of transcriptional regulation are as yet not fully understood, studies in vitro have shown that a number of auxiliary transcription factors are required to reconstitute accurate initiation of transcription. Biochemical and genetic analysis suggests that for synthesis of messenger RNA by RNA polymerase II, promoter selectivity involves the interaction of multiple cellular factors that recognize and bind to specific DNA sequences located within the promoter regions of eukaryotic genes. It is likely that both temporal regulation and tissue specificity of transcription are governed by the action of these cellular DNA binding proteins (*1*).

The promoter-specific transcription factor, Sp1, was first detected in HeLa cells on the basis of its ability to activate the SV40 early promoter (*2*). Subsequently it was shown that Sp1 recognizes and binds selectively to a GC-rich decanucleotide sequence (GC box) present as six tandem copies in the SV40 early promoter (*3*). Sp1 interacts with its recognition sites by forming protein-DNA contacts in the major groove of the DNA. In the SV40 promoter, each binding site occurs once per turn of the DNA helix, and the five highest affinity sites are occupied by protomers of Sp1 aligned on the same face of the DNA (*4*). Most important, a detailed analysis of its DNA binding and in vitro transcription properties, made with wild-type and mutant SV40 DNA templates, demonstrated a functional link between Sp1 binding and transcriptional activation (*5*).

Although Sp1 was originally discovered by study of the SV40 promoter, recent work indicates that it can also activate transcription from other viral and cellular genes that have one or more GC-box recognition elements within the 5′ flanking promoter sequences (*6–8*). In addition, several other cellular factors that bind selectively to promoter elements and potentiate transcription have been identified through similar biochemical and genetic analyses (*8, 9–11*). Of particular interest was the finding that certain promoters, such as HSV-TK (herpes simplex virus–thymidine kinase), human metallothionein II_A, and human heat shock hsp70 promoters, require multiple distinct DNA binding factors working in conjunction with each other to modulate the level of transcription (*12, 13*). Thus, it has become increasingly clear that in higher eukaryotes, transcriptional specificity and regulation is dependent on the combined action of multiple specific DNA binding factors interacting with distinct control elements in a selective fashion to activate RNA synthesis.

Although the DNA binding properties of auxiliary transcription factors such as Sp1 provide an important clue to their mode of action, the molecular mechanisms of transcriptional activation by Sp1 and other related proteins remain unknown. The initial studies that allowed detection of Sp1 and other cellular transcription factors had been carried out with either crude or partially purified preparations. However, continued progress and the success of future studies will depend largely on characterization of the biochemical properties of transcription factors. A major experimental barrier has been the purification and identification of transcription proteins because they are generally present in extremely low levels in the cell. Here we report the purification of Sp1 to apparent homogeneity. Our primary objective was to identify the polypeptide or polypeptides responsible for Sp1 DNA binding and transcriptional activation and thus define the polypeptide species comprising Sp1. To purify Sp1, we have developed a rapid and efficient sequence-specific DNA affinity chromatography procedure. We have identified Sp1 proteins by renaturing the DNA binding activity of polypeptides isolated from an SDS polyacrylamide gel. Purified Sp1 has subsequently been used to investigate the protein-DNA interactions at its tandem recognition sequences within the SV40 early promoter. In addition, we have characterized the transcriptional properties of Sp1 in reconstituted in vitro reactions.

Purification of Sp1. The purification of Sp1 from HeLa cells is summarized in Table 1. The presence of Sp1 DNA binding activity was assayed by deoxyribonuclease (DNase) I protection experiments (footprinting) with a fragment of SV40 DNA that contains the six GC-box recognition sites for Sp1. The HeLa cells were processed by hypotonic Dounce homogenization and low speed centrifugation, the soluble cytoplasmic fraction was discarded, and the nuclear pellet was extracted with buffer containing 0.42*M* KCl, as described (*14*). This crude nuclear extract was adjusted to 53 percent ammonium sulfate, and the precipitated proteins were separated by sequential

The authors are in the Department of Biochemistry, University of California, Berkeley, CA 94720.

column chromatography. The first chromatographic step was Sephacryl S-300 gel filtration (Fig. 1A). A symmetrical and highly reproducible peak of Sp1 DNA binding activity eluted between fractions 73 and 85 (Fig. 1A). Thus, under our chromatographic conditions, Sp1 appears to migrate as a native species of approximately 500 kD. Samples containing Sp1 activity were pooled and subsequently applied to DEAE Sepharose CL-6B in buffer containing 0.1*M* KCl. Under these conditions, Sp1 passes through the DEAE column and contaminating nucleic acids, phosphatases, and polymerase are retained. The proteins that flow through DEAE Sepharose at 0.1*M* KCl were directly applied to a heparin-agarose column and eluted sequentially with buffer containing 0.2, 0.3, and 1*M* KCl. Most of the protein flowed through the heparin-agarose column at 0.1*M* KCl, and the Sp1 activity was present in the 0.3*M* fraction (Fig. 1B). This relatively concentrated and highly active Sp1 fraction was diluted to 60 m*M* KCl and was further purified on a high-performance liquid chromatography (HPLC) cation exchange column. (FPLC Mono S; Pharmacia). A linear gradient of KCl (0.06 to 0.4 *M*) was applied and Sp1 eluted between 0.23 and 0.27*M* KCl (Fig. 1C). At this stage of the purification, a roughly 500-fold increase in specific activity had been achieved relative to whole cell extracts (Table 1). We estimate from the total amount of DNA binding activity present in the pooled Mono S fraction that Sp1 represents approximately 1 to 2 percent of the protein in this preparation. Moreover, analysis of the polypeptides present in this fraction by SDS gel electrophoresis reveals a complex mixture of

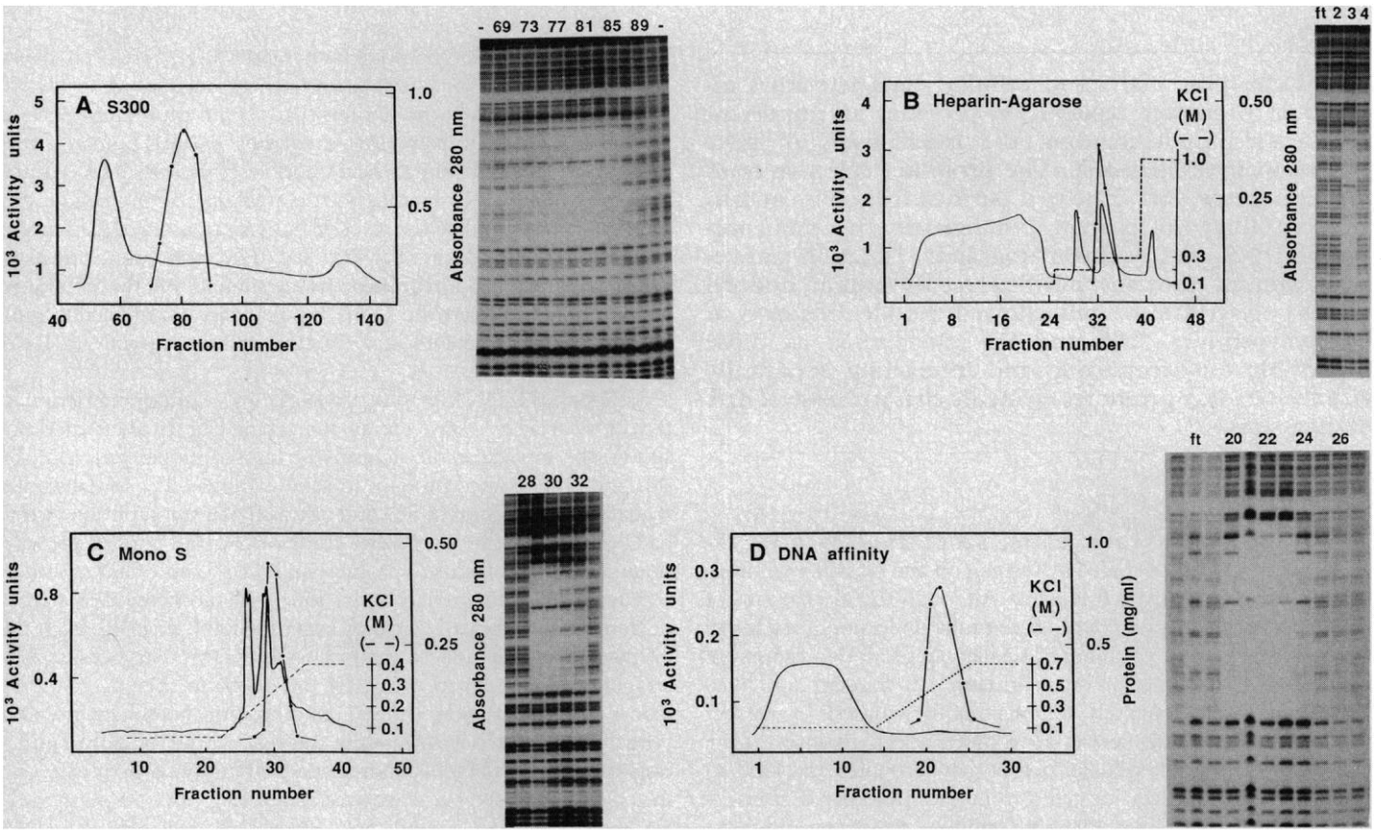

Fig. 1. Column chromatography of Sp1. Nuclear extract from 60 g of HeLa cells was prepared by the procedure of Dignam *et al.* (*14*), except that KCl rather than NaCl was used. The nuclear extract was precipitated by 53 percent saturated ammonium sulfate and centrifuged at 35,000*g* for 15 minutes; the pellet was resuspended in TM buffer (50 m*M* tris-HCl, *p*H 7.9, containing 12.5 m*M* $MgCl_2$, 1 m*M* EDTA, 1 m*M* DTT, and 20 percent glycerol) to a final protein concentration of about 30 mg/ml. (A) The soluble protein extract was applied to a Sephacryl S-300 column (5 by 51 cm) equilibrated with TM buffer containing 0.1*M* KCl. Protein elution was monitored by absorbance (*A*) at 280 nm, and Sp1 activity was determined by deoxyribonuclease protection (footprint) analysis of column fractions. The probe DNA was a 5′ end-labeled fragment generated from plasmid pSV07 (*28*) by cleavage with endonucleases Hind III and Eco RI. Footprinting was conducted as described in the legend to Table 1. The ratio of elution volume to void volume (V_e/V_o) for Sp1 was approximately 1.4. Fractions containing Sp1 were pooled (40 mg of protein in 40 to 60 ml) and applied to a 10-ml DEAE–Sepharose CL-6B column equilibrated with TM buffer containing 0.1*M* KCl. The Sp1 activity flowed through the DEAE column and was collected in one fraction containing approximately 30 mg of protein. (B) The heparin agarose resin was prepared as described (*29*). The material that flowed through the DEAE Sepharose was applied to a heparin agarose column (6 ml), equilibrated with TM buffer containing 0.1*M* KCl, and was eluted with a TM buffer series containing, successively, 0.2*M*, 0.3*M*, and 1.0*M* KCl. The Sp1 activity eluting at 0.3*M* KCl was pooled and subjected to further chromatography. (C) The heparin agarose eluate at 0.2 to 0.3*M* KCl, which contained Sp1 activity, was diluted to 60 m*M* KCl with TM buffer (or HM buffer, which contained 25 m*M* Hepes, *p*H 7.6, instead of tris) and applied to a 1-ml FPLC Mono S column. Sp1 was eluted with a 20-ml linear gradient of KCl (60 to 400 m*M*). Fractions (0.5 ml each) were collected, and the peak of Sp1 activity was detected in fractions 29 and 30 (eluted at approximately 0.26*M* KCl), as determined by DNase footprint analysis. (D) DNA affinity chromatography was carried out as described (*15*). The Sp1 eluate from the FPLC Mono S column (1.5 mg of protein in 2 ml of TM buffer containing 0.25*M* KCl) was mixed with sonicated calf thymus DNA (220 μg in 100 μl) and diluted to 0.1*M* KCl with buffer Z [25 m*M* Hepes (K^+), *p*H 7.6, containing 12.5 m*M* $MgCl_2$, 1 m*M* DTT, 20 percent (by volume) glycerol, and 0.1 percent Nonidet P-40]. This mixture was allowed to stand at 4°C for 10 minutes; it was then applied to a 1-ml DNA affinity column that was prepared by coupling concatamers (average length ~10) of annealed synthetic oligodeoxynucleotides containing the high affinity Sp1 recognition sequence, 5′-GGGGCGGGGC-3′ to CNBr-activated Sepharose CL-2B. The column was washed with buffer Z containing 0.1*M* KCl, and protein was eluted by addition of successive portions (1 ml each) of buffer Z containing 0.1*M* increments of KCl. The peak of Sp1 DNA binding activity was eluted in fractions 21 to 23 (0.4 to 0.6*M* KCl). The relative activity recovered from each of the chromatographic steps is given as the total number of DNase footprint units per column fraction as defined in the legend of Table 1.

many proteins (Fig. 2B), confirming that further purification of Sp1 is necessary to achieve homogeneity.

We subsequently tried other chromatographic columns, including FPLC Mono Q, FPLC Superose 6, Biorex-70 (Bio-Rad), hydroxylapatite, and various hydrophobic resins, but none of these provided any further purification. In most cases a gain in specific activity was achieved, but the recovery of Sp1 was poor (less than 10 percent). Consequently, we were prompted to develop less conventional yet more efficient chromatographic methods that would allow a substantially higher degree of purification.

Sequence-specific DNA affinity chromatography. The final step in the purification of Sp1 was achieved by applying the pooled Mono S fractions to a DNA affinity column prepared by covalently linking tandemly ligated synthetic oligodeoxyribonucleotide fragments containing high affinity Sp1 recognition sequences to Sepharose (see legend to Fig. 1D) (*15*). The partially purified Mono S preparation was first mixed with an appropriate amount of nonspecific competitor DNA, such as sonicated calf thymus DNA or synthetic poly d(I · C), and then binding of Sp1 to the affinity resin was carried out in a buffer containing 0.1*M* KCl. After washing the affinity resin with buffer containing 0.1*M* KCl, Sp1 was eluted with a multiple step gradient (0.2 to 1.0*M* KCl). Individual fractions eluted from the DNA affinity column were then assayed by DNase footprint protection (Fig. 1D). A peak of Sp1 DNA binding activity eluted between 0.3 and 0.6*M* KCl in fractions 20 to 24.

The proteins in these fractions were visualized by SDS polyacrylamide gel electrophoresis and silver staining (Fig. 2A). Comparison of the protein pattern before and after chromatography (Fig. 2B) showed that the sequence-specific DNA affinity column, which gives more than 50 percent recovery of Sp1, affords a remarkably efficient purification step. We estimate from the SDS gel that the affinity fraction contains a total of 10 to 15 μg of protein, whereas the input Mono S fraction contained approximately 1.5 mg of protein. Thus, the DNA affinity step provides further purification of Sp1, up to 60- to 100-fold, and the cumulative gain in specific activity throughout the purification scheme (Table 1) is approximately 30,000-fold.

The strategy that we describe for the purification of Sp1 has also been successfully applied to the isolation of several other transcription factors. Using different sequence-specific DNA affinity resins, we have purified CAAT-binding transcription factor (CTF) and activator proteins 1 and 2 (AP1 and AP2) that interact with enhancer elements, to more than 90 percent homogeneity (*13, 16*). It is likely that other sequence-specific DNA binding proteins can be isolated in a similar manner, and thus, this affinity procedure may significantly facilitate the purification of cellular DNA binding proteins. Moreover, nuclear factor I has been purified by a different DNA affinity chromatographic procedure with plasmid DNA containing multiple binding sites adsorbed to cellulose (*17*).

Analysis of purified Sp1 by SDS gel electrophoresis. The most striking feature of the protein pattern seen in the SDS gel are two prominent polypeptide species, one 105,000 daltons and the other 95,000 daltons, that are consistently present in the fractions containing Sp1 activity (Fig. 2, A and B). These two polypeptides constitute more than 95 percent of the protein in the purified samples, and they coelute with the peak of Sp1 DNA binding activity (Figs. 1D and 2A). Inspection of the SDS gels (Fig. 2) indicates also that there are several other polypeptides detectable in the fractions containing Sp1 activity, including minor species at 190, 115, and 110 kD, and trace amounts of polypeptides smaller than 90 kD. The 190-kD species is only seen when the protein samples are subjected to SDS gel electrophoresis in the absence of reducing agents such as dithiothreitol (DTT) or β-mercaptoethanol (compare Fig. 2, A and B), which suggests that the 190-kD species is a dimer of 95-kD polypeptides linked by intermolecular disulfide bonds. An estimate of Sp1 DNA binding activity present in the DNA affinity column fractions suggests that the 105- and 95-kD species are the only proteins present in sufficient quantities to bind stoichiometrically the Sp1 recognition sites of the probe DNA used in the footprint assays. Consequently, these two polypeptides represent the best candidates for Sp1.

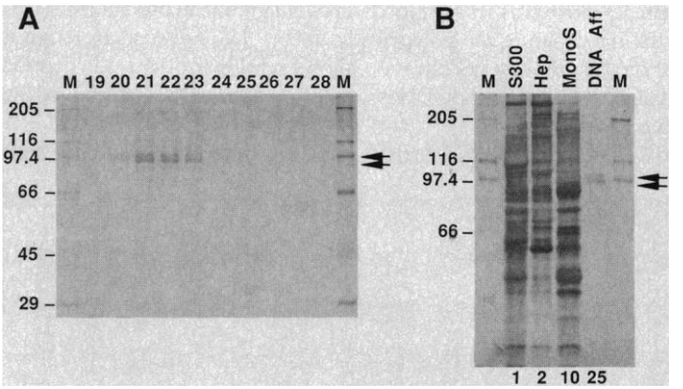

Fig. 2. Polypeptide composition of Sp1 containing fractions at different stages of purification. SDS polyacrylamide gels were prepared and subjected to electrophoresis by the method of Laemmli (*30*), and proteins were visualized by staining with silver by a modification of (*31*). (A) Samples (40 μl) of each DNA affinity fraction were precipitated with trichloroacetic acid, resuspended in SDS sample buffer, and subjected to electrophoresis on an 8 percent SDS polyacrylamide gel. Lane 19 represents a sample from the 0.2*M* KCl step; lanes 20 to 28 are samples eluted with incremental steps of KCl from 0.3 to 1.0*M*. This gel was run under nonreducing conditions. (B) Samples of pooled fractions at various stages of purification. Numbers at the bottom refer to the amount of Sp1 DNA binding (activity units) displayed in each lane. Molecular size marker lanes are designated M, molecular sizes are given in kilodaltons. The arrows designate the 105- and 95-kD Sp1 polypeptides. This 8 percent SDS polyacrylamide gel was run under reducing conditions.

Table 1. Purification of Sp1. One unit of activity is defined as the amount of Sp1 required to bind the entire region of the SV40 early promoter that encompasses the six GC boxes of the 21-bp repeats and protect these sequences from digestion by pancreatic DNase I. The DNase I footprint reaction contains 10 fmole of a 237-bp 5′ end-labeled DNA fragment generated by Eco RI and Hind III digestion of plasmid pSV07 (*28*) in a 50-μl reaction mixture containing 25 m*M* tris-HCl, *p*H 7.9, 6.25 m*M* $MgCl_2$, 0.5 m*M* EDTA, 0.5 m*M* DTT, 50 m*M* KCl, 10 percent glycerol, 2 percent polyvinyl alcohol, 0.3 μg of poly[d(I · C)], and the protein sample to be assayed. Binding was carried out for 10 minutes at 0°C, then 50 μl of a solution of 5 m*M* CaCl and 10 m*M* $MgCl_2$ was added at room temperature, the DNase I was added (the quantity was dependent on the purity of the protein fraction assayed), and mixed; the reaction mixture was incubated at room temperature for 1 minute and then stopped by the addition of an equal volume of 1 percent SDS, 20 m*M* EDTA, 200 m*M* NaCl, and 200 μg of RNA. The mixture was then extracted with phenol and chloroform, precipitated with ethanol, and analyzed on an 8 percent urea-acrylamide sequencing gel.

Procedure	Total protein (mg)	Activity		Purification factor	Cumulative yield (%)
		Total activity (units)	Specific activity (unit/mg)		
HeLa cells (60 g)	3,000				
Nuclear extract	600	10,000*	17	5×	100
Sephacryl S300	40	10,000	250	15×	100
DEAE Sepharose	30	7,500	250	1.0×	75
Heparin agarose	8	4,000	500	2×	40
FPLC Mono S	1.5	1,800	1,200	2.4×	18
DNA affinity	0.01	1,000	100,000	83×	10

*On the basis that there is 100 percent recovery of Sp1 at the Sephacryl S300 step; accurate determination of activity cannot be made on the nuclear extract.

Renaturation of Sp1 activity purified by SDS polyacrylamide gel electrophoresis. To provide conclusive evidence that the two major polypeptide species purified by the DNA affinity column represent Sp1, we have taken the pooled fractions and separated these proteins by preparative SDS gel electrophoresis. The region of the gel containing the 105- or 95-kD polypeptides (or both) was excised and eluted from the polyacrylamide gel. The SDS was removed from the protein by acetone precipitation, and the precipitated protein was resuspended in 6*M* guanidine hydrochloride (Gd-HCl) to denature the polypeptide chains (*18*). Subsequent removal of the Gd-HCl by gel filtration allows refolding of the polypeptide chains and recovery of DNA binding activities, as determined by two different assays.

Renaturation of Sp1 activity was tested by a gel mobility shift assay in which portions of renatured 105- and 95-kD polypeptides were allowed to bind a small fragment of DNA containing a single copy of an Sp1 recognition site (Fig. 3A). In the presence of native Sp1, protein-DNA complexes migrate with retarded mobility relative to unbound DNA on nondenaturing polyacrylamide gel electrophoresis (Fig. 3A, lanes 1, 3, 10, 11, and 12). The renatured 105- or 95-kD polypeptides also gave rise to the same pattern of slower migrating bands as the native Sp1 (Fig. 3A, lane 4). By contrast, a control sample containing excised and renatured proteins from other regions of the gel did not bind to nor form a specific complex with the GC-box DNA fragment (Fig. 3A, lane 6). A sample containing the 105- and 95-kD polypeptides that had been denatured but not subjected to the renaturation process also showed no DNA binding activity (Fig. 3A, lane 8). As was expected, all of the protein samples, including native Sp1 and the renatured 105- and 95-kD species, did not bind and retard a control DNA fragment that lacks a GC-box Sp1 recognition sequence (Fig. 3A, lanes 2, 5, 7, and 9). Even when a homogeneous preparation of Sp1 is used there are multiple species of slower migrating protein-DNA complexes. One possibility is that the different species represent dimers or oligomers of Sp1 bound to the single recognition site.

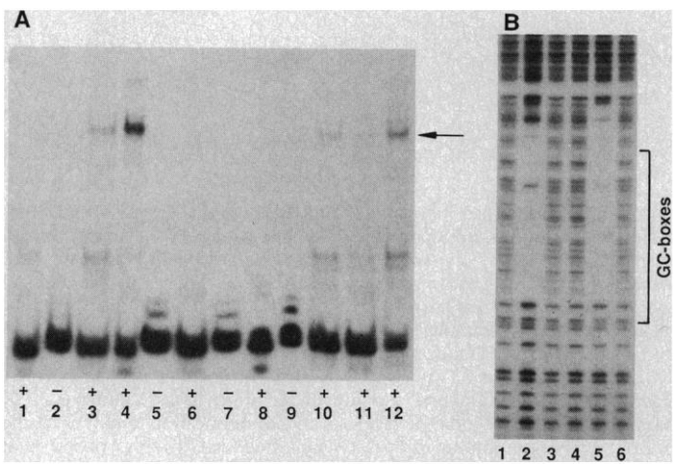

Fig. 3. Renaturation of Sp1 DNA binding activity. Approximately 2 μg of affinity purified Sp1 (>95 percent pure) was precipitated with trichloroacetic acid resuspended in SDS sample buffer, heated for 15 minutes at 68°C, and fractionated on a preparative SDS polyacrylamide gel. The region of the gel containing the 105- and 95-kD polypeptides was excised, transferred to a siliconized tissue homogenizer in 1 ml of elution buffer (50 m*M* tris-HCl, *p*H 7.9, buffer with 0.1 m*M* EDTA, 0.1 percent SDS, 5 m*M* DTT and 150 m*M* NaCl), crushed with a Teflon pestle, and incubated for 4 hours at room temperature. The acrylamide was removed by centrifugation, and the eluate was precipitated with five volumes of acetone at −20°C and centrifuged in a siliconized 15-ml Corex tube at 10,000*g* for 30 minutes at 4°C. The resulting protein pellet was resuspended in 50 μl of TM buffer containing 0.1*M* KCl, 0.1 percent NP-40, and 6*M* Gd-HCl and incubated for 30 minutes at room temperature. The Gd-HCl was removed by passing the solution through a 350-μl P-10 (Bio-Rad polyacrylamide gel filtration resin) column equilibrated with TM buffer containing 0.1*M* KCl, 0.1 percent NP-40, 1 m*M* DTT [adapted from (*18*)]. Typical recovery was 30 to 50 percent of the polypeptides from the gel in the elution step, and renaturation efficiency after removal of the Gd-HCl was approximately 10 percent, for a net yield of about 3 percent recovery of specific binding activity. (A) Gel mobility shift assay (*32*). Samples of native, denatured, and renatured Sp1 were added to a DNA binding reaction with ^{32}P-labeled probe containing one (+) or no (−) Sp1 binding site. The negative control probe used in this experiment had a contaminating DNA fragment that is evident particularly in lanes 5, 7, and 9. After protein-DNA complexes were allowed to form (under standard footprint conditions), the samples were subjected to electrophoresis on a 3.5 percent nondenaturing acrylamide gel. Lanes 1 and 11 contain 2 ng of native Sp1; lanes 2, 3, and 10 contain 4 ng of Sp1, and lane 12 contains 8 ng of Sp1. Lanes 4 and 5 contain renatured 105- and 95-kD polypeptides (~16 ng active for binding). Lanes 6 and 7 contain material extracted from a region of the gel that does not contain the 105- and 95-kD polypeptides and that was subjected to the renaturation process; lanes 8 and 9 contain eluted 105- and 95-kD polypeptides that were not subjected to the renaturation process. (B) DNase footprint protection with renatured Sp1. Lanes 1, 3, 4, and 6 are negative controls with no protein added. Lane 2 shows the pattern of protection observed with renatured 105- and 95-kD polypeptides and lane 5 displays the footprint pattern characteristic of native Sp1. DNase I footprint reactions were carried out with a fragment of SV40 DNA, as described in the legend to Table 1.

For a more discriminating assay, we have used both native Sp1 and the renatured 105- and 95-kD polypeptides to carry out DNase I footprint protection. The pattern of protected bands observed with the renatured polypeptides is indistinguishable from the footprint of native DNA affinity purified Sp1 (Fig. 3B). The results provide conclusive evidence that at least the 105- and 95-kD species are capable of recognizing and binding specifically to the GC-box motifs in a manner characteristic of Sp1.

Recently, we succeeded in separating the 105-kD species from the 95-kD species on a high resolution SDS gel and found that both polypeptides, after renaturation, are independently capable of binding selectively to the GC-box sequences of SV40. In addition, amino acid composition and preliminary HPLC analysis of peptide fragments generated by tryptic digestion of the 105- and 95-kD polypeptides suggests that they are structurally related to each other. These two species most likely represent different forms of the same protein, perhaps as a result of proteolytic degradation or covalent modification. It is also possible that the 95- and 105-kD forms of Sp1 are both proteolytic degradation products of a larger, full-length polypeptide. We are not at present able to distinguish between these two possibilities and further experiments are necessary to establish their functional relationship.

The more general applicability of renaturing specific DNA binding activity from SDS gels was shown by testing the recovery of binding activity of two other eukaryotic transcription factor, AP1 and Adf1 (*13*). In each case we found that it was possible to recover specific DNA binding activity from the renatured polypeptides. A number of other DNA binding factors have been isolated, and this technique may be a useful method for assigning function to specific polypeptides. Although renaturing DNA binding activity from SDS gels has not been widely used, a similar approach has been successfully applied to the identification of various bacterial transcription components and sigma factors (*18, 19*).

Transcriptional activation properties of purified Sp1. Throughout the purification described above, we relied solely on monitoring Sp1 DNA binding activity. It is therefore necessary to establish whether the nearly homogeneous preparations of Sp1 also retain the capacity to activate in vitro transcription. This is a particularly important point because all previous transcription studies have used relatively impure preparations of Sp1 (on the order of 1 to 2 percent pure), and thus, it was possible that the protein responsible for DNA binding is distinct from the transcriptional activator protein (*5, 7, 12*). We therefore used the most highly

purified preparations of Sp1 (more than 95 percent pure) to stimulate transcription in a reconstituted in vitro reaction directed by the SV40 early and human metallothionein I_A (MTI_A) promoters. In the presence of endogenous RNA polymerase II and a fraction (Sp2) that contains various general initiation factors—TFIIB, TFIID, and TFIIE (*20*)—required for RNA synthesis, only very low basal level transcription of the SV40 and MTI_A templates was observed (Fig. 4A, lanes 3 and 5). However, when purified Sp1 is added to the reaction, a 5- to 20-fold enhancement of transcription is observed (Fig. 4A, lanes 4 and 6). By contrast, transcription of a control template (adenovirus major late promoter) that does not contain an Sp1 recognition element showed no significant stimulation of transcription (Fig. 4A, lane 2). In addition to assaying for transcriptional activity of the most purified pooled Sp1 fractions, we also assayed its transcription in each fraction eluted from the DNA affinity column. Transcriptional activation coincides perfectly with both the DNA binding activities and the presence of the 105- and 95-kD polypeptides (Figs. 2A and 4B). These findings provide the most direct biochemical evidence that both DNA binding and transcriptional activities of Sp1 are carried out by the same protein. These data confirm that the 105- and 95-kD proteins are responsible for Sp1 activity.

DNA binding properties of purified Sp1. DNA binding studies were done to test whether there are any significant differences between the specificity and affinity of the highly purified Sp1 relative to more crude fractions that had been used in most of the prior studies on Sp1. In particular, we previously found that the pattern of Sp1 binding to the SV40 21-bp repeat elements was rather complex because there are six potential Sp1 recognition sites, but only five are generally occupied. Our results suggested that each of the GC-box elements in SV40 could interact independently with protomers of Sp1, and that there appeared to be a hierarchy of binding affinities. With increasing concentrations of the highly purified protein, DNase I footprint analysis of Sp1 and the SV40 21-bp repeats shows the characteristic pattern of protected and enhanced bands that was previously observed with crude Sp1 fractions (Fig. 5A). To examine the interaction of Sp1 with the 21-bp repeats at higher resolution, protection was also carried out with methidiumpropyl-EDTA · Fe(II) [MPE · Fe(II)] as the cleavage reagent (*21*) instead of DNase I. Interestingly, MPE · Fe(II) footprint analysis reveals that a region of approximately nine nucleotides at site IV is not efficiently protected by Sp1 (Fig. 5B). These results confirm dimethylsulfate methylation protection studies performed with partially purified Sp1 on a collection of clustered base substitution mutations spanning the Sp1 binding sites of SV40 (*5*). Thus, MPE protection shows that site IV is bound less well by Sp1 under conditions where the other sites are fully occupied. Analysis of mutants in site V suggests that the poor binding of Sp1 to site IV is at least in part due to steric constraints imposed by protomers of Sp1 bound at the adjacent high affinity site V. The physiological

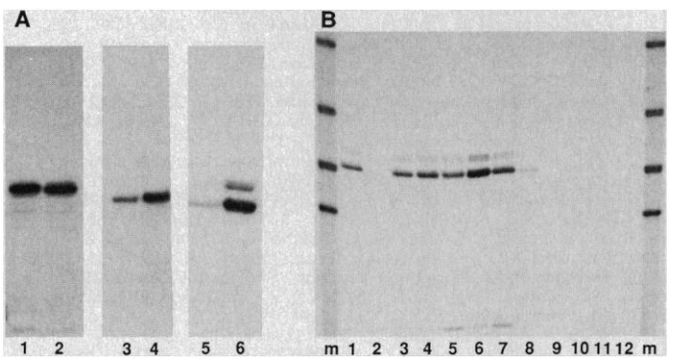

Fig. 4. Analysis of Sp1 directed transcription of the SV40 early promoter. In vitro reconstituted transcription reactions were assayed by primer extension analysis as described previously (*2*). (A) Characterization of Sp1 dependence of various eukaryotic promoters. Lanes 2, 4, and 6 contain Sp1 (20 ng) in the reaction, whereas lanes 1, 3, and 5 do not. Lanes 1 and 2 show the lack of a response to Sp1 characteristic of adenovirus 2 major late promoter. Lanes 3 and 4 show the human cellular gene metallothionein I_A and lanes 5 and 6 show the response of SV40 early promoter. (B) Transcription profile across the DNA affinity column (see Figs. 1D and 2A). Lane 1 represents the activity that flows through the resin; lane 2 is the 0.1*M* washing; and lanes 3 to 11 are the 0.1*M* incremental steps from 0.2*M* to 1.0*M* KCl. A 1-μl (~10 ng) sample of each fraction was assayed for Sp1 activity. Lane 12 shows the reaction products in the absence of Sp1. Each transcription reaction contained 100 ng of SV40 template DNA, a crude RNA polymerase II fraction (~25 μg of total protein), and the general transcription factor (or factors) Sp2 [~3 μg of total protein (*2, 3*)]. Purified Sp1 was added as indicated. The transcription reaction was allowed to proceed for 45 minutes at 30°C, and single-stranded, end-labeled specific primer was added, hybridized at 60°C for 1 hour, and extended with AMV reverse transcriptase (10 units, Life Sciences) for 1 hour at 37°C. The final products were subsequently analyzed on denaturing 8 percent polyacrylamide gels and visualized by autoradiography.

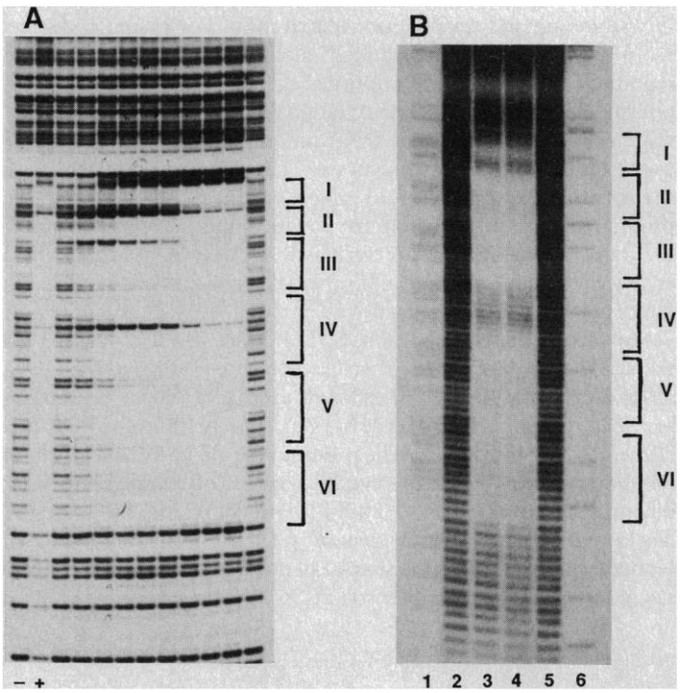

Fig. 5. DNA binding analysis with highly purified Sp1. (A) DNase I footprint analysis. Purified Sp1 was used to footprint SV40 DNA as described (*3*). The (+) lane is a control showing the footprint characteristic of a crude Sp1 protein fraction (heparin 0.3*M* KCl). The remainder of the gel shows the footprint obtained by increasing the concentration of homogeneous Sp1 from 0 to 20 ng (left to right). Outside lanes are no protein controls (−). (B) MPE · Fe (II) footprint analysis (*22*). Purified Sp1 (50 to 100 ng) in buffer P [10 μl; buffer P is 500 m*M* Pipes (K^+), *p*H 7.0, containing 50 m*M* KCl and 10% (v/v) glycerol] was combined with ^{32}P-labeled pSV07 probe (25 fmol in 6 μl H_2O) and incubated on ice for 15 minutes. A solution of 200 μ*M* $Fe(NH_4)(SO_4)_2$ (28 μl) was combined with 1.5 m*M* MPE (2 μl), and 2 μl of this solution was added to each sample. The digestion of DNA was initiated by the addition of 40 m*M* DTT (2 μl), and the samples were incubated for 5 minutes at 24°C. The reactions were stopped by the addition of 100 m*M* EDTA, *p*H 8, containing 10 m*M* thiourea, 150 m*M* sodium acetate, and 200 μl of transfer RNA at 0.15 μg/ml. The samples were extracted with a mixture of phenol and chloroform (1:1, by volume) and then precipitated. The resulting oil was resuspended in 70 percent ethanol (800 μl), chilled, and reprecipitated. The pellet was resuspended in sample buffer containing formamide and subjected to electrophoresis on an 8 percent urea–acrylamide gel. Lane 1 shows sequence markers from the cytosine-specific reaction on the probe DNA as described by Maxam and Gilbert (*33*). Lanes 2 and 5 are no protein controls; lane 3 is 50 ng Sp1; lane 4 is 100 ng; lane 6 is the guanine-specific sequence markers (*33*).

significance of this apparent negative cooperativity between site IV and V is not known.

Presence of Sp1 in different eukaryotic cells. We have surveyed a number of sources other than HeLa cells for Sp1 DNA binding activity. Thus far, we have detected Sp1 activity in nuclear extracts from cultured Chinese hamster ovary cells, human T cells, and human placenta tissue (*22*). Preliminary analysis of Sp1 purified from human placenta suggests that the active DNA binding species is a polypeptide of approximately 40 kD rather than the 105- and 95-kD proteins characteristic of HeLa Sp1. The behavior of the placental Sp1 suggests that it is most likely a proteolytic subfragment of the 105- and 95-kD species. Interestingly, the truncated placental protein displays the same affinity and specificity for binding to the Sp1 recognition elements of SV40 as the HeLa protein, but unlike the 105- and 95-kD species, the smaller form appears to be significantly less potent in transcriptional activity (*22*). Thus, it seems unlikely that DNA binding in itself will be directly responsible for potentiating transcription. It would, however, not be surprising to find that eukaryotic transcription factors such as Sp1 are multifunctional proteins with separate domains devoted to DNA binding and transcriptional activation. For example, distinct DNA binding and transcriptional activation domains have been reported for the yeast transcriptional activator GAL4 (*23*). Indeed, our findings suggest that it may be possible to isolate a proteolytic fragment of Sp1 bearing the DNA binding domain but lacking the transcriptional activation functions. They are also consistent with the idea that Sp1 may activate mRNA synthesis by binding to the promoter and interacting selectively with additional components of the transcriptional apparatus, such as RNA polymerase II subunits, enhancer binding proteins, and other transcription factors through direct protein-protein contact.

How many genes respond to Sp1? It is now evident that Sp1 is a cellular transcription factor that has the potential to activate transcription from a number of different viral and cellular promoters, provided that in each case, at least one properly positioned GC-box element is contained within the promoter element. A comparison of 36 different binding sites reveals a range of binding affinities, differing by at least 10- to 20-fold with individual binding sites displaying a remarkable degree of sequence variation (*24, 25*). Nonetheless, it has been possible to define a decanucleotide consensus sequence that is characteristic of Sp1 binding sites,

$$5' \begin{matrix} GG \\ TA \end{matrix} GGCG \begin{matrix} GGGC \\ TAAT \end{matrix} 3'$$

In all of the cases studied thus far, the binding of purified Sp1 to the promoter region can be correlated with in vitro transcriptional activation mediated by Sp1. More significantly, in many cases including SV40, HSV TK, and MT II_A, Sp1 appears to act in conjunction with other cellular transcription factors such as CTF, AP1, and AP2 to modulate the level of transcription. Perhaps the interplay of multiple interdigitated sequence-specific transcription factors will emerge as a characteristic feature common to most eukaryotic promoters. Examples of multiple factor interactions at eukaryotic promoters have recently been reported for transcription of human ribosomal RNA genes by RNA polymerase I (*26*), adenovirus 2 major late transcription by RNA polymerase II (*10*), and transfer RNA gene transcription by RNA polymerase III (*27*).

Sp1 appears to be an important component of the cellular transcriptional apparatus. Unfortunately, we cannot predict from the examples tested how many transcription units in the cell are capable of responding to Sp1, or how many different sequence-specific DNA binding proteins participate in the formation of a complex regulatory circuitry that is necessary to govern eukaryotic gene expression. However, with the purification of Sp1 and the identification of the active polypeptide species, it should now be possible to obtain a partial amino acid sequence and prepare appropriate DNA probes to isolate the gene encoding Sp1. In addition, purified Sp1 will be used to generate antibodies against this transcription factor. These experiments should facilitate analysis of the temporal and spatial expression of Sp1 as well as the study of the mechanism of transcriptional activation by promoter-specific factors.

REFERENCES AND NOTES

1. W. S. Dynan and R. Tjian, *Nature (London)* **316**, 774 (1985).
2. ———, *Cell* **32**, 669 (1983).
3. ———, *ibid.* **35**, 79 (1983).
4. D. Gidoni, W. S. Dynan, R. Tjian, *Nature (London)* **312**, 409 (1984).
5. D. Gidoni, J. T. Kadonaga, H. Barrera-Saldana, K. Takahashi, P. Chambon, R. Tjian, *Science* **230**, 511 (1985).
6. W. S. Dynan, J. D. Saffer, W. S. Lee, R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4915 (1985).
7. W. S. Dynan, S. Sazer, R. Tjian, R. T. Schimke, *Nature (London)* **319**, 246 (1985).
8. K. A. Jones and R. Tjian, *ibid.* **317**, 179 (1985).
9. C. S. Parker and J. Topol, *Cell* **37**, 273 (1984).
10. M. Sawadogo and R. G. Roeder, *ibid.* **43**, 165 (1985).
11. R. W. Carthew, L. A. Chodosh, P. A. Sharp, *ibid.*, p. 439.
12. K. A. Jones, K. R. Yamamoto, R. Tjian, *ibid.* **42**, 559 (1985).
13. W. Lee, B. England, R. Tjian, unpublished data.
14. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, *Nucleic Acids Res.* **11**, 1475 (1983).
15. J. T. Kadonaga and R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5889 (1986).
16. K. Jones, J. Kadonaga, P. Rosenfeld, T. Kelly, R. Tjian, in preparation.
17. P. J. Rosenfeld and T. J. Kelly, *J. Biol. Chem.* **261** (No. 3), 1398 (1986).
18. D. A. Hager and R. R. Burgess, *Anal. Biochem.* **109**, 76 (1980).
19. J. Pero, R. Tjian, J. Nelson, R. Losick, *Nature (London)* **257**, 248 (1975).
20. M. Sawadogo and R. G. Roeder, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4394 (1985).
21. P. Dervan, *Science* **232**, 464 (1986).
22. M. Briggs, J. Kadonaga, A. Axelrod, R. Tjian, unpublished observations.
23. R. Brent and M. Ptashne, *Cell* **43**, 729 (1985).
24. J. T. Kadonaga, K. A. Jones, R. Tjian, *Trends Biochem. Sci.* **11**, 20 (1986).
25. K. A. Jones, J. T. Kadonaga, P. A. Luciw, R. Tjian, *Science* **232**, 755 (1986).
26. R. M. Learned, T. K. Learned, M. M. Haltiner, R. T. Tjian, *Cell* **45**, 847 (1986).
27. S. Fuhrman, D. Engelke, E. P. Geiduschek, *J. Biol. Chem.* **259**, 1934 (1984).
28. R. M. Myers and R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6491 (1980).
29. B. L. Davison, T. Leighton, J. C. Rabinowitz, *J. Biol. Chem.* **254**, 9220 (1979).
30. U. K. Laemmli, *Nature (London)* **227**, 680 (1970).
31. J. H. Morrissey, *Anal. Biochem.* **117**, 307 (1981).
32. M. Fried and D. M. Crothers, *Nucleic Acids Res.* **9**, 6505 (1981).
33. A. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).

34. We thank B. Johnson and the tissue culture staff for HeLa cells, W. Lee for assistance with the adenovirus transcription, K. Jones for helpful suggestions and protein fractions; D. Rio, C. M. Kane, W. Morgan, C. Hauser, R. Klein for reading of the manuscript; K. Ronan for assistance in the preparation; and P. Dervan for providing MPE. Supported by NIH grant T32 ES07075 (M.B.); a fellowship from the Miller Institute for Basic Research; grants from the NIH (R.T.).

11 June 1986; accepted 29 August 1986