Method to identify genomic targets of DNA binding proteins

LAUREN SOMPAYRAC* AND KATHLEEN J. DANNA

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309-0347

Communicated by David M. Prescott, January 16, 1990 (received for review October 30, 1989)

ABSTRACT We have devised a cyclical immunoprecipitation protocol that can be used to identify and clone a specific DNA sequence that is recognized by a DNA binding protein, even if that sequence is present in only one copy in the genome of a mammal. As an example, we have used this procedure to purify mouse genomic sequences to which the simian virus 40 tumor (T) antigen binds.

Regulation of gene expression is one of the most intensely studied topics in modern biology. Recently, techniques such as DNase cleavage protection patterns (footprinting) and gel retardation assays have made it possible to identify anonymous DNA binding proteins (transcription factors) when their DNA targets are known or to confirm that a known DNA sequence is the target of a known DNA binding protein. However, these techniques cannot be used when the protein has been identified but its target in the mammalian cell is unknown. In fact, a general technique does not exist to answer the question, "Given a DNA binding protein, to which DNA sequences in the mammalian genome does it bind?" This type of information would be useful in many cases. For example, the binding of certain oncoproteins to specific sequences on cellular DNA may result in the abnormal expression of cellular genes, thereby eliciting a cancerous phenotype. To test this hypothesis, one would like to discover the cellular DNA sequences to which these oncoproteins bind. In multicellular organisms, regulatory proteins that bind to specific cellular DNA sequences can influence the expression of genes involved in development. This is another case in which it would be useful to identify the genomic targets of regulatory proteins.

McKay provided an example of how an unknown DNA binding site could be identified when he used an immunoprecipitation procedure to discover the simian virus 40 (SV40) sequences to which SV40 tumor (T) antigen binds (1). Although the binding site that McKay found was contained in the small SV40 genome [5×10^3 base pairs (bp)], we thought it might be possible to devise an immunoprecipitation technique that could discover the specific target of a given DNA binding protein, even if that sequence were represented only once in the 3×10^9 -bp genome of a mammalian cell. In this paper we describe such a procedure and use it to purify cellular sequences to which SV40 T antigen binds.

METHODS

Cells, Viruses, Plasmids, Antibodies, and Competitor DNA.

The 293 (transformed primary human embryonal kidney) cell line (2) and the Ad5-SVR111 virus (adenovirus type 5–SV40 hybrid virus further described in *Results*) (3) were kind gifts from Yasha Gluzman of the Cold Spring Harbor Laboratory. MAX Efficiency DH5a *Escherichia coli* were purchased from Bethesda Research Laboratories. To construct pTet, we removed most of the ampicillin-resistance gene from

pBR322 by deleting the sequences between the *Dra* I sites at nucleotides 3232 and 3943. pSVO10, a gift from Aleem Siddiqui of the University of Colorado Medical School, contains a 228-bp *Eco*RI–*Hind*III fragment of SV40 DNA that spans the origin of DNA replication (4). We constructed pBRori by using linkers to insert these origin sequences from pSVO10 into the *Eco*RI site of pBR322. We constructed pSP64ori by inserting the SV40 origin sequences from SV010 into the *Bam*HI site of pSP64 (Promega). Pab108 (5) and Pab419 (6) are monoclonal antibodies specific for SV40 T antigen. To prepare competitor DNA, we resuspended salmon sperm DNA (Sigma D1626) in SSC/10 (0.015 M NaCl/0.0015 M sodium citrate, pH 7) overnight at a concentration of 10 mg/ml and sonicated this DNA on ice five times for 30 sec each at 60% power in a model 300 Fisher Sonic Dismembrator. The average size of the sonicated DNA was about 600 bp.

Preparation of a Mouse Genomic Library. We partially digested genomic DNA from C3H/10T $\frac{1}{2}$ mouse cells (7) with *Sau*3AI enzyme and used gel electrophoresis to select DNA fragments that were between 2 and 7 kilobases (kb) long. We ligated these fragments into *Bam*HI-cut, phosphatase-digested SP64 plasmid DNA, used this ligated DNA to transfect competent DH5 bacteria, and purified amplified library DNA by CsCl/EtdBr centrifugation.

Preparation of *Staphylococcus aureus* Bacteria. We used the method of Kessler (8) to prepare formalin-fixed *Staphylococcus aureus* (Cowan I strain) and froze them in aliquots at -70°C . For each experiment, we thawed a fresh aliquot of cells, activated the bacteria for 20 min at room temperature at a 10% concentration (wt/vol) in STE buffer (100 mM NaCl/10 mM Tris, pH 7/1 mM EDTA) containing 0.5% Nonidet P-40 (NP-40), and then washed them with an equal volume of STE buffer containing 0.05% NP-40.

Protein Extracts. We infected subconfluent 293 cells growing on one 100-mm dish with Ad5-SVR111 virus at a multiplicity great enough to cause a cytopathic effect in 100% of the cells in about 22 hr. At that time, we harvested the cells by scraping, washed them twice with cold phosphate-buffered saline, and lysed them for 1 hr on ice in 400 μl of 150 mM Tris, pH 8.0/150 mM NaCl/1 mM EDTA/1 mM dithiothreitol/10% (vol/vol) glycerol/0.5% NP-40/100 μg of aprotinin (United States Biochemical; 6000 kallikrein inactivator units/mg per ml). We centrifuged the lysate for 5 min at $500 \times g$ to pellet nuclei, centrifuged the supernatant for 2 min in a Microfuge, and precleared an aliquot of the supernatant by adding an equal volume of 50% (wt/vol) *S. aureus* suspended in $1 \times$ binding buffer (0.01 M Pipes, pH 7/0.1 mM EDTA/1 mM dithiothreitol/0.05% NP-40/100 μg of aprotinin per ml/5% glycerol) containing 50 mM NaCl. After 15 min on ice, we centrifuged the staphylococcus–protein mixture in a Microfuge for 2 min and removed the supernatant to a new tube. We precleared this supernatant a second time by adding a volume of 50% *S. aureus* suspension equal to the original volume of the protein aliquot. After the mixture was on ice

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: SV40, simian virus 40; T antigen, tumor antigen; Ad5, adenovirus type 5; EtdBr, ethidium bromide.

*To whom reprint requests should be addressed.

for 15 min, we pelleted the bacteria, removed the supernatant to a new tube, centrifuged again, and used an aliquot of the supernatant for our experiments. All centrifuging was done with a precooled Microfuge.

Immunopurification Procedure. The following procedure was devised over a long period of trial and error. We strongly suggest that the protocol be followed as closely as possible, since small changes in the procedure can result in large increases in nonspecific background immunoprecipitation.

First cycle. On ice we mixed 50 μ l of 2 $\times$ binding buffer, 2 μ l of 5 M NaCl, 28 μ g of plasmid library DNA, and distilled water to a final volume of 100 μ l. To this we added 4 μ l of precleared protein extract (see above) and incubated the mixture overnight at 4°C to allow protein–DNA complexes to form. The next day, we resuspended a 10- μ l aliquot of activated, washed 10% (wt/vol) suspension of *S. aureus* in 100 μ l of solution A (1 $\times$ binding buffer containing 100 mM NaCl and 5 mg of sonicated salmon sperm DNA per ml) and incubated this mixture on ice for at least 30 min before use. To the protein–plasmid library DNA mixture, we added 35 μ l of solution B [25 μ l of Pab419 culture supernatant and 10 μ l of 1 $\times$ binding buffer] and incubated the mixture on ice for 30 min. We then added the 100 μ l of preincubated *S. aureus* suspension and continued the incubation for an additional 10 min. We centrifuged the immunoprecipitate for 1 min at full speed in an Eppendorf 5415 microcentrifuge in the cold room to pellet the DNA–T-antigen–*Staphylococcus* complexes and removed the supernatant with a Pasteur pipette. We washed the immunoprecipitate three times by adding 1 ml of solution C (10 mM Tris, pH 7/100 mM NaCl/0.1 mM EDTA/0.05% NP-40), mixing in a Vortex to resuspend the pellet, centrifuging for 1 min, and removing the supernatant. We then added 1 ml of solution C to the pellet, mixed in a Vortex, transferred the sample to a new tube, centrifuged again, and removed the final wash with a Pasteur pipette. At room temperature, we resuspended the bacterial pellet in 40 μ l of SSC/10 that contained 0.02 μ g of sonicated salmon sperm DNA, mixed the sample, added 10 μ l of solution D [5% sodium dodecyl sulfate (SDS)/50 mM Tris, pH 8/10 mM EDTA/2 mg of yeast RNA per ml], and incubated the sample for 10 min at 60°C. We centrifuged the sample for 2 min, removed the supernatant containing the eluted plasmid DNA into a fresh tube, centrifuged again to pellet any remaining bacteria, and precipitated the supernatant by adding sodium acetate and ethanol. We pelleted the DNA in a cold Microfuge, washed the pellet with 70% ethanol, dried the DNA under vacuum, resuspended it in 25 μ l of SSC/10, centrifuged once more for 2 min to pellet any debris, and used the supernatant to transfect 500 μ l of MAX Efficiency DH5 α competent cells. DNA from untreated cells was purified either by CsCl/EtdBr banding or by a plasmid miniprep procedure.

Second cycle. The second cycle of immunopurification was identical to the first with the following exceptions. Since the first cycle of purification resulted in a 1000-fold enrichment of library sequences with T antigen binding sites, for the second cycle we used a mixture of 0.028 μ g of DNA from the first cycle of immunopurification plus 20 μ g of pTet DNA as a nonspecific-binding competitor. Because our library was constructed in a plasmid vector that confers ampicillin resistance and since pTet does not, this competitor DNA will not survive amplification in bacteria grown in the presence of ampicillin. In the second cycle, we used monoclonal antibody Pab 108, which recognizes an epitope on T antigen that is different from the epitope recognized by Pab 419.

Third cycle. The third cycle of immunopurification was performed in the same way as the second except that 2.8 $\times$ 10⁻⁵ μ g of twice-purified library DNA was used together with 20 μ g of pTet.

Notes on the Immunopurification Protocol. (i) We have tried protein A attached to several different matrices, but *S. aureus* gives the best ratio of signal to background. (ii) At least two different monoclonal antibodies should be used in the course of immunopurification. Monoclonal antibodies can cross-react with cellular DNA binding proteins, resulting in the immunoprecipitation of unwanted cellular sequences. (iii) After several cycles of immunopurification, it is important to monitor whether library plasmids that remain have specific binding sites for the protein of interest. This can be accomplished by mixing the purified library DNA with a roughly equal amount of vector plasmid DNA (with no cellular DNA insert) and performing an additional cycle of immunopurification. Gel electrophoresis on the resulting DNA will reveal whether library plasmids that contain putative binding sites are more efficiently immunoprecipitated than the vector plasmid that does not. (iv) The background of nonspecific immunoprecipitation increases as the amount of *S. aureus* increases. Therefore, we used only 10 μ l of a 10% *S. aureus* suspension. We performed pilot immunoprecipitations with labeled T antigen to determine the maximum volume of antibody culture supernatant that could be efficiently bound by 10 μ l of *S. aureus* suspension and the maximum volume of protein extract that could be efficiently immunoprecipitated by this volume of antibody culture supernatant. (v) We decreased the amount of library DNA used to begin each cycle by the enrichment factor achieved in the previous cycle. From this, the reader might conclude that one must know the enrichment factor in advance. Fortunately, this is not the case. Because the molarity of T antigen in the binding reaction is much greater than that of the DNA binding sites, the immunopurification works well over a wide range of library DNA concentrations (compare the ratio of entries 5 and 3 in experiment 1 vs. the ratio of entries 7 and 6 in experiment 2 in Table 1).

RESULTS

Experimental Design. To be generally useful, an immunoprecipitation procedure should have the following properties: (i) it should be powerful enough to identify and clone a DNA binding site even if it were represented only once in the mammalian genome; (ii) protein used for these experiments should be derived from crude, unpurified cellular lysates because DNA binding activity might be lost during blind

Table 1. Single cycle of immunopurification

Exp.	DNA	Protein	Efficiency of immunoprecipitation
1	1. pTet	None	8×10^{-6}
	2. pTet	XO	1.4×10^{-5}
	3. pTet	T Ag	1.4×10^{-5}
	4. pBRori	XO	$<3.7 \times 10^{-4*}$
	5. pBRori	T Ag	$1.2 \times 10^{-2+}$
2	6. pTet	T Ag	2.0×10^{-5}
	7. pBRori	T Ag	$2.2 \times 10^{-2+}$

Ag, antigen. In experiment 1 we mixed 20 μ g of pTet DNA with 10⁻⁵ μ g of pBRori DNA and added either no protein, a protein extract from XO-infected 293 cells (XO; an adenovirus containing no gene for T antigen), or a protein extract from Ad5-SVR111-infected 293 cells (T Ag). We performed a single cycle of immunopurification, transfected *E. coli* with the resulting DNA, and evaluated the fraction of input DNA that survived immunoprecipitation by scoring transfected bacteria for tetracycline resistance (pTet plus pBRori) or for ampicillin resistance (pBRori). Experiment 2 is identical to experiment 1 except that we mixed 20 μ g of pTet with 10⁻³ μ g of pBRori. *No colonies were observed.

[†]The low efficiency of complex formation between pBRori DNA and T antigen reflects the rather large equilibrium dissociation constant of this complex, about 10⁻⁹ M (9).

purification; and (iii) binding and immunoprecipitation wash buffers should have a physiological pH and a moderate salt concentration because high pH and high salt concentrations can dissociate DNA-protein complexes. To aid us in devising an immunoprecipitation scheme with these features, we chose SV40 T antigen as a model protein. T antigen is a well-characterized DNA binding protein that binds not only to the SV40 origin of DNA replication but also to origin-like sequences in mouse and human DNA (10, 11).

Preliminary experiments with T antigen and SV40 origin DNA indicated that because of the large size of the mammalian genome and the inherent stickiness of DNA, it would be impossible to achieve the needed level of purification in a single cycle of immunoprecipitation. To overcome this problem, we used a protocol with several cycles of immunoprecipitation to purify the target binding site. Fig. 1 is a schematic representation of this purification scheme. The first step is to construct a library of cellular DNA sequences in a bacterial plasmid vector. Because extremely high transfection efficiencies can be achieved by using the Hanahan procedure (12) or bacterial electroporation (13), a genomic library can be constructed that has a high probability of containing a given cellular sequence. Next, plasmid DNA from this library is incubated with a crude cellular lysate that contains T antigen. When DNA-T antigen complexes have formed, a monoclonal antibody specific for T antigen is added, and the complexes are immunoprecipitated by the addition of *S. aureus*. The immunoprecipitates are washed extensively to remove as much uncomplexed plasmid DNA as possible, and plasmid DNA is recovered from the immunoprecipitated DNA-T antigen complexes. The result of this single cycle of purification is a library enriched in plasmids that contain T antigen binding sites. This library is then amplified in *E. coli* and used to begin another cycle of immunopurification. The cycling procedure is continued until plasmids that contain cellular DNA inserts that bind T antigen are purified away from plasmids that do not.

Conditions for a Single Cycle of Immunopurification. We first determined incubation and wash conditions that would maximize the binding of T antigen to origin DNA and that would minimize the immunoprecipitation of background plasmid DNA. We mixed DNA of a plasmid that confers ampicillin and tetracycline resistance and contains SV40 origin sequences (pBRori) with a 2×10^6 -fold excess of plasmid DNA that confers only tetracycline resistance and contains no SV40 origin sequences (pTet). We incubated this DNA mixture with a crude protein extract from Ad5-

SVR111-infected 293 cells. Ad5-SVR111 is an adenovirus-SV40 hybrid virus in which the SV40 T antigen coding sequence is under the control of the adenovirus major late promoter (3). To evaluate various incubation and wash conditions, we subjected the T antigen-DNA mixture to a single round of immunoprecipitation and used the immunoprecipitated DNA to transfect *E. coli*. We assayed the efficiency of immunoprecipitation of origin-containing plasmids (the signal) by plating the transfected bacteria on agar that contained ampicillin. Because the amount of pTet DNA in our mixture was so much greater than the amount of pBRori DNA, we could estimate the frequency with which non-origin-containing plasmids survived immunoprecipitation (the background) by plating the transfected bacteria on agar that contained tetracycline. Using this assay, we varied the immunoprecipitation conditions until the signal-to-background ratio was maximum. As a control, we did parallel immunoprecipitations in which we replaced the T antigen-containing extract with an extract from XO-infected 293 cells. XO is an adenovirus identical to Ad5-SVR111 except that it contains no gene for T antigen (14). Table 1 summarizes results of typical experiments using conditions we found to be optimal. For example, in experiment 1, 1.2×10^{-2} of the origin-containing plasmids survived immunoprecipitation compared with only 1.4×10^{-5} of the non-origin-containing plasmids, indicating that it should be possible to enrich for origin-containing sequences roughly 1000-fold in a single cycle of immunopurification.

Reconstruction Experiments. Next we determined whether, in multiple cycles of purification, a single T antigen binding site could be purified away from a genome equivalent of non-origin-containing DNA. A library of 2×10^6 plasmids, each containing a 4.5-kb insert of cellular DNA, would have about a 95% probability of including a given cellular DNA sequence (15). Therefore, in reconstruction experiments, we mixed plasmid DNA containing the SV40 origin of replication (pSP64ori) with a 2×10^6 -fold excess of plasmid DNA that contains no origin sequences (pSP64) and added a lysate from Ad5-SVR111-infected 293 cells. Complexes between pSP64ori DNA and T antigen were immunoprecipitated, and the resulting DNA was used to transfect *E. coli* to obtain once-purified DNA. Measured amounts of this DNA together with known amounts of pSP64ori DNA were subjected to Southern blot analysis. When we probed this blot with a radioactive oligonucleotide specific for the SV40 origin, we estimated that the origin-containing plasmids had been enriched about 1000-fold in this first cycle of immunoprecipitation (data not shown). After a second cycle of purification, gel electrophoresis and EthBr staining revealed that the ratio of pSP64 to pSP64ori DNAs was about 2:1, indicating that the mixture of DNAs had been enriched about 10^6 -fold in the origin-containing plasmid (Fig. 2, lanes 3 and 4). When, as a control, we replaced the T antigen-containing extract with an extract from XO-infected 293 cells, no purification was observed (Fig. 2, lane 2). After a third cycle of purification, only a single band corresponding to pure pSP64ori DNA could be detected by gel electrophoresis (Fig. 2, lane 5). This reconstruction experiment demonstrates that in three cycles of immunopurification, an origin-containing plasmid can be purified away from a genome equivalent of plasmid DNA that does not contain the SV40 origin.

Identification of Cellular DNA Sequences to Which T Antigen Binds. Since others have shown that mouse DNA contains origin-like sequences to which SV40 T antigen can bind, we decided to use the immunoprecipitation protocol to purify cellular binding sites for T antigen. Such an experiment would be a realistic test of the cyclical procedure. We constructed a mouse genomic library in a bacterial plasmid vector (pSP64LIB) and amplified this DNA in *E. coli*. We mixed DNA of a plasmid containing the SV40 origin of

FIG. 1. Outline of cyclical immunopurification scheme.

FIG. 2. Reconstruction experiment to test multiple cycles of immunopurification. We mixed 20 μg of pSP64 DNA with 10^{-5} μg of pSP64ori DNA, performed three cycles of immunopurification, and subjected the immunoprecipitated DNA to agarose gel electrophoresis. Lanes: 1, result of one cycle of immunopurification with an extract from Ad5SVR111-infected 293 cells that contains T antigen; 2, result of second cycle with extract from XO-infected 293 cells that contains no T antigen; 3, result of second cycle using T antigen-containing extract; 4, same as lane 3 except that DNA from the first cycle was purified by a miniprep procedure instead of CsCl/EtdBr centrifugation; 5, result of three cycles of immunopurification with T antigen-containing extract.

replication (pSP64ori) with a 2×10^6 -fold molar excess of pSP64LIB DNA to test whether the cyclical immunoprecipitation technique was powerful enough to pull out the SV40 binding site if it were present in an amount that corresponded to a single entry in a library large enough to represent the entire mouse genome. If this were true, then after several rounds of purification we would expect to purify not only those cellular sequences that had T antigen binding sites, but also the plasmid we added that contains the SV40 viral origin. Since inserts of cellular DNA in the library ranged from 2 to 7 kb, the 3.3-kb origin-containing plasmid would be easy to identify by gel electrophoresis. After we subjected the mixture of pSP64LIB and pSP64ori DNAs to one cycle of purification, Southern blot analysis revealed that the pSP64ori DNA had been enriched by roughly a factor of 2000, which, within experimental error, is consistent with the enrichments observed in the reconstruction experiments. This result indicated that the majority of the library plasmids did not contain sequences that bound T antigen and demonstrated that the pSP64ori plasmid, representing a single binding site equivalent of DNA, had survived the first cycle of purification.

When, after two cycles of purification, the resulting DNA was subjected to gel electrophoresis, EtdBr staining revealed a faint band corresponding to the pSP64ori plasmid below the

FIG. 3. Immunopurification of cellular DNA sequences that contain T antigen binding sites. We mixed 28 μg of pSP64LIB DNA with 7×10^{-6} μg of pSP64ori (as an internal control) and subjected the mixture to three cycles of immunopurification followed by gel electrophoresis. Lanes: 1, aliquot of starting mixture [the dim band at the bottom represents vector DNA (pSP64) that recircularized without inserts of cellular DNA when the library was constructed]; 2, result after one cycle of immunopurification; 3, result after two cycles; 4, result after 3 cycles.

FIG. 4. Immunopurified library DNA contains specific T antigen binding sites. We used *Ban* I to digest DNA from five library plasmids (lanes 1–5) that survived three cycles of immunopurification, end-labeled the DNA with ^{32}P , and subjected the fragments produced by this digestion to a single cycle of immunoprecipitation followed by agarose gel electrophoresis and autoradiography. Lanes: A, aliquot of digest before immunoprecipitation; B, immunoprecipitation with a T antigen-containing protein extract; C, immunoprecipitation with an extract from XO-infected 293 cells (contains no T antigen).

brighter smear that represents the partially purified pSPLIB DNA (Fig. 3, lane 3). By comparing the brightness of the pSP64ori band in the twice-purified DNA with that of known amounts of pSP64ori DNA, we estimated that the pSP64ori DNA in the mixture had been enriched by only about 50-fold during the second cycle of immunopurification. Such a decrease in enrichment in the second cycle relative to the first would be expected if there were entries in the cellular DNA library that contained binding sites for T antigen. After a third cycle of immunopurification, there was no significant further enrichment of pSP64ori relative to pSP64LIB, indicating that library plasmids that survived three cycles of immunopurification had inserts containing T antigen binding sites (Fig. 3, lane 4).

To confirm the presence of T antigen binding sites in the thrice-purified library plasmids, we picked individual colonies of bacteria from the third-cycle transfection and made DNA minipreps. We digested these DNAs with *Ban* I and subjected the DNA fragments to one cycle of immunopurification followed by gel electrophoresis. Fig. 4 shows that the plasmids do contain fragments, derived from mouse DNA inserts, that are specifically immunoprecipitated by T antigen. Since others have studied examples of cellular T antigen binding sites in detail (10, 11), we did not analyze these plasmids further.

DISCUSSION

We have described a general method for purifying cellular DNA sequences that are targets for DNA-binding proteins. This simple method involves multiple cycles of immunopurification, each of which can be completed in about 3 days. During each cycle, the genomic library is enriched in entries that contain specific binding sites. By performing multiple cycles of purification, library plasmids containing binding sites can be purified essentially to homogeneity.

Three pieces of evidence indicate that this method is powerful enough to purify a cellular DNA binding site, even if it occurs only once in the genome of a mammal. First, we showed that when plasmid DNA that contains a binding site for SV40 T antigen is mixed with a 2×10^6 -fold excess of plasmid DNA that does not, a single cycle of immunopurification results in about a 1000-fold enrichment in plasmids that contain binding sites (Table 1). This result suggested that

three such cycles of immunoprecipitation should be sufficient to purify a plasmid that contains a T antigen binding site away from a genome equivalent of plasmids that do not. In our second set of experiments (Fig. 2), we performed three cycles of enrichment and demonstrated that the plasmid containing the T antigen binding site was purified essentially to homogeneity. Finally, as a more realistic test, we mixed identifiable plasmid DNA containing a T antigen binding site with DNA of a mouse genomic library in proportions such that the marked plasmid represented a single binding site in a genome equivalent of library plasmid DNA. After three cycles of immunopurification, the marked, origin-containing plasmid could be recovered, even though many other genomic binding sites for T antigen were present in the genomic library (Figs. 3 and 4).

To use the cyclical procedure with crude protein extracts, one must have available a system that overproduces the protein of interest relative to other cellular proteins. The reason is not that a large amount of protein is needed; in our experiments we used an aliquot of cell lysate containing about 50 ng of T antigen. Rather, the need for overproduction arises because mammalian cells contain a large number of DNA binding proteins that can bind to library plasmids and can survive immunoprecipitation by sticking to *S. aureus*. This can result in the immunoprecipitation of library plasmids that do not have binding sites for the protein of interest (Table 1; compare entries 1 and 2). In our experiments, we used Ad5-SVR111-infected cells as a source of T antigen. However, a number of other mammalian, bacterial, and insect systems should work equally well.

Although we designed the protocol so that unpurified protein preparations could be used, the cyclical procedure also works well with purified proteins. In fact, when we used immunopurified T antigen in pilot experiments, we achieved a single-cycle enrichment that was even greater than that achieved with unpurified lysates (data not shown).

There are several limitations on the use of our protocol. First, the enrichment obtained in each cycle depends on the relative affinity of the protein of interest for its specific, high-affinity binding site versus its affinity for other, low-affinity binding sites or its nonspecific affinity for DNA in general. In our test experiments we used SV40 T antigen as our model protein. In a sense, this may represent a worst-case example because T antigen is a notoriously promiscuous DNA binding protein that even binds to specific low-affinity sites on plasmid DNA (16). Still, there may be DNA-binding proteins whose affinity for a specific DNA binding site is so low that the immunoprecipitation procedure could not be used.

Another limitation of this technique arises because protein-DNA complexes are not infinitely stable. There is a period of about 20 min in the immunopurification procedure during which complexes that dissociate will be lost. Thus, DNA-protein complexes with a dissociation half-life significantly shorter than this will not survive immunoprecipitation. To minimize this potential problem, wash steps are performed under conditions that help stabilize protein-DNA complexes, namely 4°C in the absence of magnesium (17). Fortunately, although data on dissociation half-lives of eukaryotic protein-DNA complexes are scanty, it appears that these complexes are relatively stable. For example, the

half-life of T-antigen-origin complexes is about 100 min (1), the half-life for the dissociation of the adenovirus major late transcription factor from its target sequence is greater than 20 hr (17), and the dissociation half-life of the *Drosophila* heat shock activator protein from its binding site is about 30 min (18).

We would like to emphasize that our approach is not unique. Others have used similar protocols to identify, in cases in which there were many binding sites per genome, a subset of these binding sites (10, 11, 19–21). Usually, these experiments were carried out with binding and wash buffers that were not physiological (10, 11, 19, 20). What we describe here are conditions that make the immunoprecipitation procedure powerful enough to identify, using physiological buffers, a DNA binding site that occurs only once in the mammalian genome. We believe this procedure should be especially useful for cloning rare targets of DNA binding proteins.

We thank Sharon Spencer for expert technical assistance, Dr. Bill Dynan for the oligonucleotide we used for Southern blotting, and Dr. Larry Gold for helpful suggestions on the manuscript. We are especially indebted to Dr. Yasha Gluzman, whose kind gifts of reagents made these experiments possible. This work was supported by U.S. Public Health Service Grant CA-34072 from the National Cancer Institute to L.S.

1. McKay, R. D. G. (1981) *J. Mol. Biol.* **145**, 471–488.
2. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. (1977) *J. Gen. Virol.* **36**, 59–74.
3. Gluzman, Y., Reichl, H. & Solnick, D. (1982) in *Eucaryotic Viral Vectors*, ed. Gluzman, Y. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 187–192.
4. Myers, R. M. & Tjian, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6491–6495.
5. Gurney, E. G., Tamowski, S. & Deppert, W. (1986) *J. Virol.* **57**, 1168–1172.
6. Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. (1981) *J. Virol.* **39**, 861–869.
7. Reznikoff, C. A., Brankow, D. W. & Heidelberger, C. (1973) *Cancer Res.* **33**, 3231–3238.
8. Kessler, S. W. (1975) *J. Immunol.* **115**, 1617–1624.
9. Hinzpeter, M., Fanning, E. & Deppert, W. (1986) *Virology* **148**, 159–167.
10. Lane, D. P., Simanis, V., Bartsch, R., Yewdell, J., Gannon, J., & Mole, S. (1985) *Proc. R. Soc. London Ser. B* **226**, 25–42.
11. Gruss, C., Wetzel, E., Baack, M., Mock, U. & Knippers, R. (1988) *Virology* **167**, 349–360.
12. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
13. Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988) *Nucleic Acids Res.* **16**, 6127–6145.
14. Sompayrac, L. & Danna, K. J. (1986) *Virology* **153**, 297–309.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Wright, P. J., DeLucia, A. L. & Tegtmeyer, P. (1984) *Mol. Cell. Biol.* **4**, 2631–2638.
17. Chodosh, L. A., Carthew, R. W. & Sharp, P. A. (1986) *Mol. Cell. Biol.* **6**, 4723–4733.
18. Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V. & Ueda, H. (1987) *Science* **238**, 1247–1252.
19. Oliphant, A. R., Brandl, C. J. & Struhl, K. (1989) *Mol. Cell. Biol.* **9**, 2944–2949.
20. Pollwein, P., Wagner, S. & Knippers, R. (1987) *Nucleic Acids Res.* **15**, 9741–9759.
21. Biedenkapp, H., Borgmeyer, U., Sippel, A. E. & Klempnauer, K. (1988) *Nature (London)* **335**, 835–837.