MOLECULAR AND CELLULAR BIOLOGY, Sept. 1985, p. 2307–2315
0270-7306/85/092307-09$02.00/0

Vol. 5, No. 9

Novel Method for Identifying Sequence-Specific DNA-Binding Proteins

DAVID LEVENS[1] AND PETER M. HOWLEY[2]*

Laboratory of Pathology[1] and Laboratory of Tumor Virus Biology,[2] National Cancer Institute, Bethesda, Maryland 20205

Received 10 April 1985/Accepted 14 June 1985

We developed a general method for the enrichment and identification of sequence-specific DNA-binding proteins. A well-characterized protein-DNA interaction is used to isolate from crude cellular extracts or fractions thereof proteins which bind to specific DNA sequences; the method is based solely on this binding property of the proteins. The DNA sequence of interest, cloned adjacent to the *lac* operator DNA segment is incubated with a *lac* repressor–beta-galactosidase fusion protein which retains full operator and inducer binding properties. The DNA fragment bound to the *lac* repressor–beta-galactosidase fusion protein is precipitated by the addition of affinity-purified anti-beta-galactosidase immobilized on beads. This forms an affinity matrix for any proteins which might interact specifically with the DNA sequence cloned adjacent to the *lac* operator. When incubated with cellular extracts in the presence of excess competitor DNA, any protein(s) which specifically binds to the cloned DNA sequence of interest can be cleanly precipitated. When isopropyl-β-D-thiogalactopyranoside is added, the *lac* repressor releases the bound DNA, and thus the protein-DNA complex consisting of the specific restriction fragment and any specific binding protein(s) is released, permitting the identification of the protein by standard biochemical techniques. We demonstrate the utility of this method with the lambda repressor, another well-characterized DNA-binding protein, as a model. In addition, with crude preparations of the yeast mitochondrial RNA polymerase, we identified a 70,000-molecular-weight peptide which binds specifically to the promoter region of the yeast mitochondrial 14S rRNA gene.

Sequence-specific DNA-binding proteins are the mediators of many essential cell regulatory processes including transcription, initiation of DNA replication, and site-specific recombination. Studies in phage and bacterial genetics have allowed the identification and characterization of a number of procaryotic sequence-specific DNA-binding proteins. However, relatively few such proteins have been identified in eucaryotes, where classical genetic manipulations are cumbersome and biochemical analyses, are often hindered by smaller quantities of starting materials and by the apparent low abundance of many sequence-specific DNA-binding proteins.

In contrast, recombinant DNA technology coupled with improved methods for introducing in vitro manipulated DNA into eucaryotic cells has permitted the rapid identification of specific sequences involved in the regulation of gene expression, in the replication of viral DNAs, or in site-specific recombination (such as immunoglobulin gene rearrangement). Techniques for in vivo and in vitro footprinting can define the precise DNA sequences protected by unidentified factors (12). In addition, the development of accurate systems for the in vitro transcription of cloned DNAs has helped define some of the template requirements necessary for faithful RNA synthesis (10, 17). We devised a general method which uses specific DNA regulatory sequences identified by standard molecular biological approaches to detect and enrich the protein components which specifically interact with those sequences. The method uses a well-characterized specific and avid DNA-protein interaction, the binding of the *lac* repressor to the *lac* operator, as the basis for isolating specific factors from complex mixtures of protein. The DNA sequence of interest is cloned adjacent to a *lac* operator, and a DNA fragment containing both the operator and regulatory sequence is precipitated by the sequential addition of both a *lac* repressor–beta-galactosidase fusion protein and immobilized antibodies directed against beta-galactosidase (Fig. 1A). The specifically bound DNA fragment is then incubated with competitor DNA and an extract or fraction containing the factor(s) which recognizes the regulatory sequence of interest (Fig. 1B). The immobilized DNA is recovered (Fig. 1C), and the *lac* operator-containing DNA fragment with bound proteins is specifically released by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) which allosterically induces the *lac* repressor (Fig. 1D). The eluate can then be analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

We demonstrated the specificity of this technique by isolating lambda repressor from whole cell extracts of *Escherichia coli* as a complex with the lambda operator region o_L. In a single step the repressor is enriched from a minor cell constituent to approximately 90% of the total protein. The applicability of this method to less well-characterized systems is shown by the identification of a 70,000-molecular-weight (MW) peptide which binds to the promoter region of the yeast mitochondrial 14S rRNA gene. Partially purified preparations of yeast mitochondrial RNA polymerase, which support faithful transcription in vitro of yeast mitochondrial genes, contain a prominent 45,000-MW peptide as well as many minor bands which can be separated by polyacrylamide gel electrophoresis and visualized with silver staining (9, 15). Although polyclonal antibodies directed against gel-purified 45,000-MW peptide inhibit polymerase activity, the enzyme was never purified to homogeneity, and there has been no definitive evidence that this 45,000-MW protein is an intrinsic component of the yeast mitochondrial transcriptional apparatus. Further analysis of the factors and conditions influencing specific and nonspecific transcription with conventional methods has been hampered by the low relative abundance and small absolute quantities of specific

* Corresponding author.

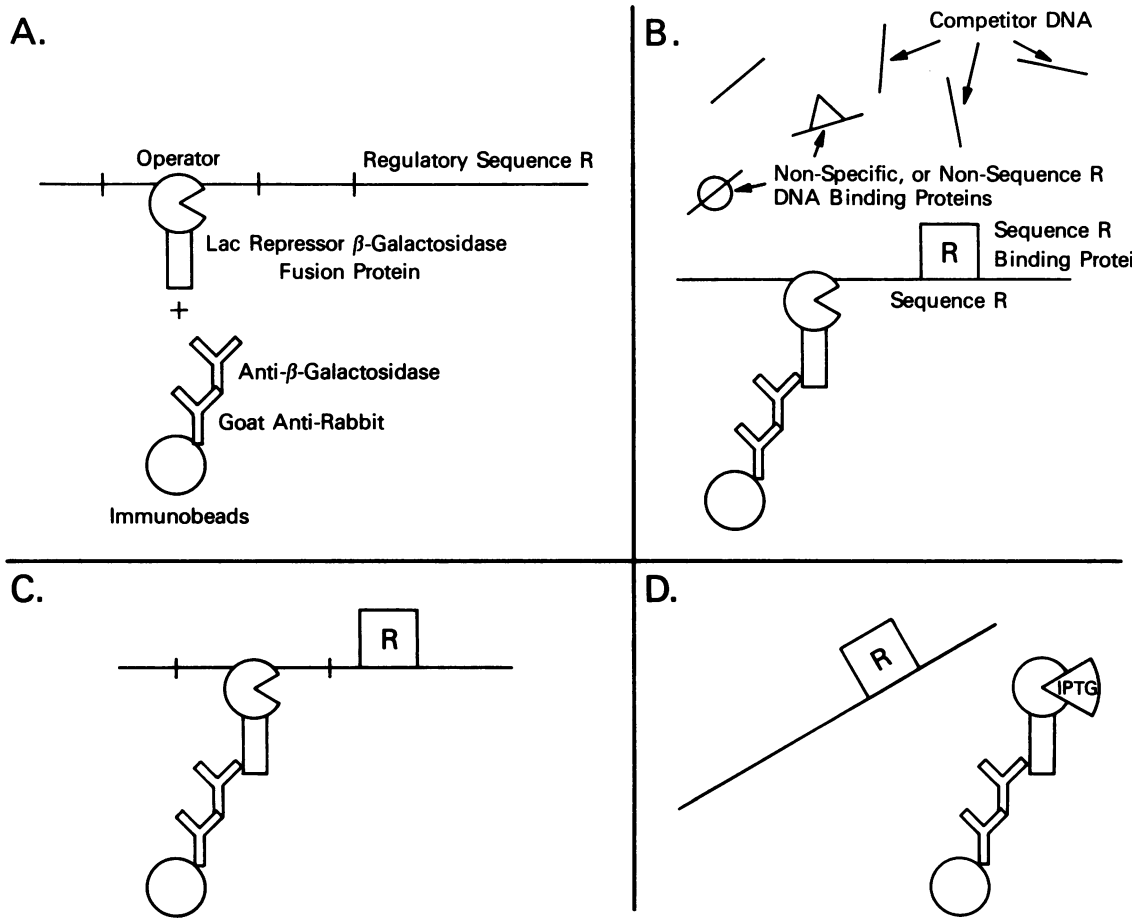

FIG. 1. Schematic summary of procedure for identifying sequence-specific DNA-binding proteins. A detailed protocol is described in the text.

proteins present in preparations of the mitochondrial RNA polymerase. With the knowledge that the sequences immediately containing and surrounding the initiation site of transcription are sufficient to promote RNA synthesis in vitro, we inferred that a promoter region-binding protein (perhaps the RNA polymerase itself) would exist and might be amenable to identification with the technique described above. A 70,000-MW peptide, a minor component of partially purified preparations of the yeast mitochondrial RNA polymerase, can be demonstrated to bind the 14S rRNA promoter region. Furthermore, preliminary evidence correlates the binding of this protein with the presence of specific transcription in vitro, although not with total RNA synthetic catalytic activity.

MATERIALS AND METHODS

Bacterial strains. BMH72-19-1, which contains the *lac* repressor–beta-galactosidase fusion gene, was the generous gift of M. Koenen and B. Muller-Hill, Universitat zu Köln, Cologne, Federal Republic of Germany. Bacterial strain N99 (harboring the plasmid pKB277 which encodes the lambda repressor *c*I gene under *lac* control) was the generous gift of M. Gottesman, National Cancer Institute, Bethesda, Md.

Plasmids. Six plasmids were used in these experiments. Experiments with the lambda repressor used pUC8 and a pUC8 derivative in which an *Hpa*I (nucleotide 35,261 from the left end of lambda DNA)–*Hin*dIII (nucleotide 36,845) fragment was cloned between the *Sma*I and *Hin*dIII sites in the polylinker. This fragment contains the leftward promoter (p_L) and its operator (o_L). The resulting plasmid is designated pUC8-o_L (Fig. 2). Cleavage of pUC8-o_L with restriction endonucleases *Pvu*II and *Bgl*II produced three fragments. The smallest fragment is 733 base pairs and contains the *lac* operon regulatory sequences as well as o_L and p_L. Cleavage of pUC8 with *Pvu*II produced two fragments; the smaller fragment is 301 base pairs (bp) and contains the *lac* operon regulatory sequences.

Four plasmids were used in the experiments identifying the mitochondrial promoter region DNA-binding protein (Fig. 3). Two plasmids, pTB3′+29 and pTB3′−31, contain segments of the yeast mitochondrial 14S rRNA gene with deletions extending from a site in the body of the gene, in a 5′ direction, to positions +29 and −31, respectively (2). The 5′ border of the mitochondrial sequences in these plasmids is the *Sac*II site at position −234 relative to the site of the initiation of transcription. The mitochondrial sequences are cloned between the *Bam*HI (5′) and *Eco*RI (3′) sites of pUR250 (26). pTB3′+29 contains the *lac* operator approximately 250 base pairs to the 5′ side of the initiation site of transcription. The mitochondrial insert and the *lac* operator can be excised on a 501- or 561-bp *Pvu*II restriction enzyme fragment from pTB3′−31 or pTB3′+29, respectively. In addition, the mitochondrial inserts from these plasmids were

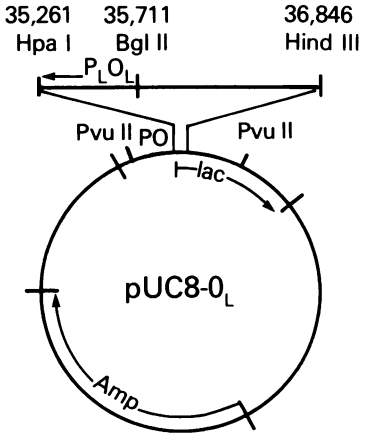

FIG. 2. Structure of pUC8-o_L. The *Hpa*I-*Hin*dIII fragment containing lambda DNA sequences 35,261 through 36,845 was excised from plasmid pKC30 and inserted between the *Sma*I and *Hin*dIII sites of pUC8. A 733-bp *Pvu*II-*Bgl*II fragment which contains the *lac* operator, as well as operator o_L and promotor p_L, can be cleaved from the resulting plasmid.

cloned as *Eco*RI/*Bam*HI fragments between the *Eco*RI and *Bam*HI sites of pBR322, generating pBR3′−31 and pBR3′+29, respectively; this step separated the mitochondrial sequences from the *lac* operator to prevent their precipitation in binding experiments.

High-MW DNA. Commercial salmon sperm DNA (Sigma Chemical Co., St. Louis, Mo.) was sonicated, extracted four times with phenol and once with chloroform, and precipitated with ethanol. *E. coli* DNA was prepared by the procedure of Marmur (18).

Bacterial extracts. Bacterial extracts were prepared by the procedure of Mauer et al. (19). Briefly, mid-log bacteria were harvested, suspended in 2 volumes of buffer, sonicated, centrifuged for 10 min at 13,000 × *g*, and dialyzed into 10 mM Tris hydrochloride (pH 7.4)–2 mM $CaCl_2$–0.1 mM EDTA–1 mM beta-mercaptoethanol–5% glycerol (WB) with 100 mM KCl.

Antibodies. Rabbit anti-beta-galactosidase was purchased from Cappel Laboratories, Cochranville, Pa., and affinity purified on beta-galactosidase coupled to CNBr-activated

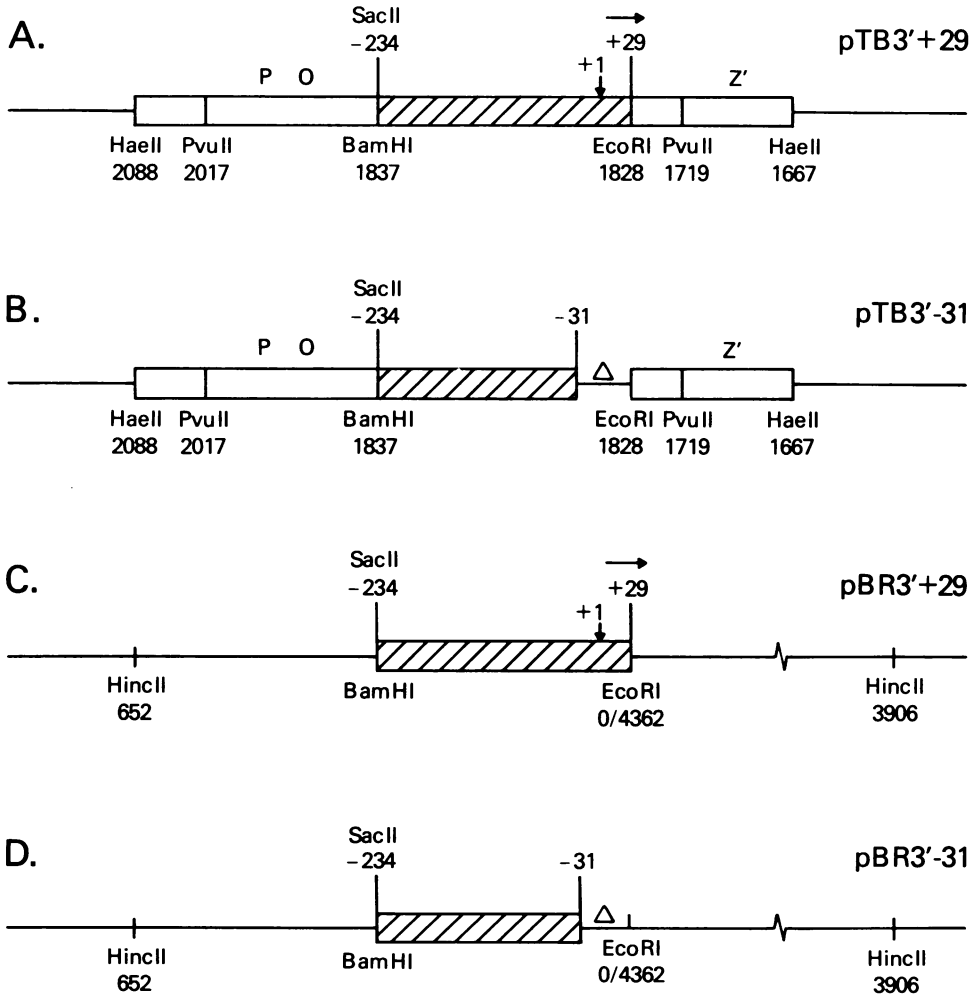

FIG. 3. Summary of plasmids. The hatched region represents mitochondrial sequences containing or immediately 5′ to the initiation site of transcription of the 14S rRNA gene. The small arrow shows the direction of transcription. (A) pTB3′+29 was constructed as described previously (9). The open bar shows the i gene and *lac* operon sequences present in pUR250. The thin lines indicate the remaining vector sequences. (B) pTB3′−31 was constructed as described previously (9). The thin line with the delta symbol indicates the sequences deleted in this plasmid compared with pTB3′+29. (C) pBR3′+29 was constructed by cloning the mitochondrial sequences of pTB3′+29 into pBR322. (D) pBR3′−31 was constructed by cloning the mitochondrial sequences of pTB3′−31 into pBR322.

Sepharose prepared by the specifications of the manufacturer (Pharmacia Fine Chemicals, Piscataway, N.J.). Immunobeads with covalently bound goat anti-rabbit were purchased from Bio-Rad Laboratories, Richmond, Calif. *lac* repressor–beta-galactosidase fusion protein was prepared by the procedure of Fowler and Zabin (11). Preliminary experiments were performed with fusion protein, which was the generous gift of I. Zabin.

Mitochondrial RNA polymerase. A partially purified preparation of yeast mitochondrial RNA polymerase (gift of J. Wettstein, University of Chicago, Chicago, Ill.) was prepared by the procedure of Levens et al. (15) through the phosphocellulose step with the omission of the Sepharose 4B column. Polyacrylamide gel electrophoresis of this enzyme revealed a complex mixture of peptides, ranging in MW from approximately 15,000 to 200,000.

Protein electrophoresis. Proteins were separated on SDS-polyacrylamide gels (10% for mitochondrial experiments and 12% for lambda repressor) and visualized with Kodavue stain (Eastman Kodak Co., Rochester, N.Y.). The staining procedure used did not visualize nucleic acids, in contrast to conventional silver staining procedures.

Binding of specific proteins. The binding procedure conveniently is divided into several steps. The *lac* repressor–beta-galactosidase fusion protein was added to a solution of restriction endonuclease-digested plasmid DNA. This fusion protein retains full operator-binding capacity, inducibility, and beta-galactosidase activity (3, 21). The restriction endonuclease chosen should leave both the *lac* operator and the DNA sequence of interest on the same fragment. Typically, 3 to 5 μg of digested plasmid DNA is dissolved in *lac* repressor binding buffer, and 1 to 2.5 μg of *lac* repressor–beta-galactosidase fusion protein is added. To minimize the binding of non-operator DNA, it is essential to use sufficient operator-containing DNA to saturate the *lac* repressor–beta-galactosidase fusion protein.

Immunobeads were suspended in phosphate-buffered saline according to the recommendations of the manufacturer, and affinity-purified anti-beta-galactosidase (approximately 10 μg/ml of suspension) and 100 μg of bovine serum albumin per ml were added. This suspension was then incubated for 3 to 12 h at 4°C and washed twice with phosphate-buffered saline and then twice with *lac* repressor binding buffer (10 mM Tris hydrochloride [pH 7.4], 10 mM $MgCl_2$, 20 mM NaCl, 0.1 mM dithiothreitol, 0.1 mM EDTA). The washed Immunobeads were suspended in 100 to 200 μl of binding buffer per ml of original suspension. Immunobeads from a total of 0.5 ml of the original suspension were added for each microgram of *lac* repressor–beta-galactosidase fusion protein. The beads were added in three to four portions at intervals of approximately 15 min to the repressor-operator binding reaction. The incubation was continued for another 60 min with periodic mixing, after which the beads were briefly centrifuged and washed twice in binding buffer. The washed beads with the bound restriction fragment next were suspended in the appropriate buffer along with an excess of competitor DNA. The protein extract or fraction was added last. The temperature and duration of the incubation was dependent upon the individual protein. In some experiments, after the binding reaction, additional competitor DNA was added to further absorb out non-sequence-specific DNA-binding proteins. The reaction was briefly centrifuged, washed two to four times in the buffer of choice, and eluted into 10 mM Tris hydrochloride–2 mM EDTA–2 mM IPTG—20 mM NaCl for 20 to 30 min. A portion of the sample may be analyzed by agarose gel electrophoresis for the presence of the *lac* operator-containing restriction fragment, and the remainder of the eluate may be concentrated by lyophilization and subjected to polyacrylamide gel electrophoresis for the analysis of protein components. Nonspecific competitor DNA must be included in the final wash to increase the efficiency of elution. Proteins and DNA not eluted with IPTG may be recovered by washing the beads in a solution containing 0.2% SDS. Typically, 50 to 90% of the bound restriction fragment is present in the IPTG eluate.

Transcription reactions. For in vitro transcription reactions, Immunobeads bearing immobilized restriction endonuclease fragments containing or deleted of mitochondrial 14S promoter region sequences were incubated with 40 μl of crude RNA polymerase and washed under the same conditions as described in the legend to Fig. 5. However, the second wash did not include salmon sperm DNA. The Immunobead pellets with bound protein were suspended in 50 μl containing 10 mM Tris hydrochloride (pH 7.4), 10 mM $MgCl_2$, 125 μM each of ATP, GTP, and CTP, 5% glycerol, 100 μg of bovine serum albumin per ml, 1 mM dithiothreitol, 1 mM IPTG, and 20 μM [α-^{32}P]UTP (60 Ci/mmol) as well as template DNA at 25 to 60 μg/ml. The template DNA was pBR3′+29 cleaved with either *Hin*cII (which generates a 485-nucleotide runoff transcript from the mitochondrial promoter in the presence of the mitochondrial RNA polymerase) or *Aat*II (which yields a 101-nucleotide runoff transcript). Specific transcription of the immobilized promoter-containing DNA pTB3′+29 results in a 138-nucleotide transcript (2). pTB3′−31 does not direct specific RNA synthesis (2). Transcription reactions were allowed to proceed for 20 min at 20°C with periodic mixing. The reactions were stopped by the addition of SDS and EDTA to final concentrations of 1% and 10 mM, respectively. After centrifugation, the supernatants were removed, proteinase K was added to 100 μg/ml, and the mixtures were incubated at 37°C for 30 min. The reactions were extracted with phenol, and the aqueous phase was centrifuged on a Bio-Gel P6 column as described previously (23). The excluded material was precipitated with ethanol, dried, counted for Cherenkov radiation, and analyzed by urea-polyacrylamide gel electrophoresis. Runoff transcripts synthesized by crude mitochondrial RNA polymerase and not subjected to the binding procedure were processed in a similar manner.

RESULTS

Binding of lambda repressor demonstrates the principle of this technique. We describe here a method designed to exploit knowledge of regulatory sequences, based on genetic or in vitro experiments, to isolate the factors which bind to the DNA regulatory elements. The feasibility of this approach is demonstrated by specifically binding and enriching lambda repressor from crude lysates of *E. coli*. Figure 4, lanes 1 and 2, shows lysates of *E. coli* containing plasmid pKB277 from cells grown in the presence or absence of 1 mM IPTG. The amount of protein run on the gel shown is equal to 1% of the input into the binding reactions. Examination of the crude lysates reveals, as expected, a complicated mixture of proteins. Comparison of the lysates in the region of lambda repressor monomers (26,000 MW) fails to reveal any prominent IPTG-induced band. In addition, comparison of these crude lysates with *E. coli* extracts from cells possessing no lambda repressor gene fails to reveal the presence of a new band at 26,000 MW (data not shown). Expression of the lambda *c*I gene under *lac* control has been reported to result in the synthesis of lambda repressor amounting to approximately 1% of cellular protein (1, 14).

The extracts shown were subjected to the binding procedure outlined above in the presence of *E. coli* competitor DNA according to the details in the legend of Fig. 4. Of the total IPTG eluate from the induced and uninduced extracts, 8% is shown in lanes 3 to 6. Induced and uninduced extracts reveal the binding of the expected 26,000-MW lambda repressor peptide to DNA fragments containing the *lac* regulatory seqeunces o_L and p_L (lanes 4 and 6) but not to the *Pvu*II fragment of pUC8 containing the *lac* sequences but no lambda sequences (lanes 3 and 5). From the intensity of staining in several gels, we estimate that in induced extracts the 26,000-MW band amounts to approximately 90% of the total bound and eluted protein; it also is clearly the major protein bound, even in uninduced cells. No similar band could be enriched in extracts from cells lacking a *c*I gene

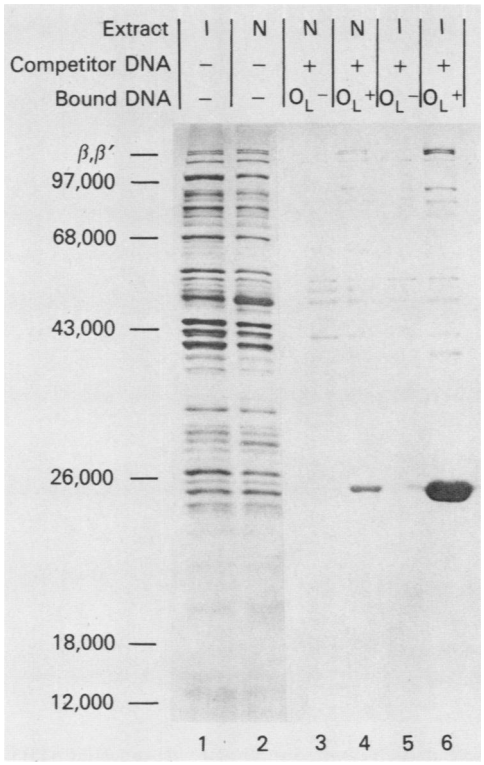

FIG. 4. Enrichment of lambda repressor from whole-cell extracts in one step. A 6-μg amount of pUC8 (o_L^-) cleaved with *Pvu*II and 7.5 μg of pUC8-o_L (o_L^+) cleaved with *Pvu*II and *Bgl*II were each incubated with 4 μg of repressor–beta-galactosidase fusion protein and Immunobeads saturated with anti-beta-galactosidase as described in the text. The beads were washed twice with binding buffer containing 20 mM KCl and suspended in 400 μl of WB with 30 mM KCl–10 mM $MgCl_2$. Each reaction was divided equally, and 100 μl of WB containing 170 μg of sonicated *E. coli* DNA(c) was added to each tube. Then 100 μl of an *E. coli* extract from cells expressing lambda repressor under *lac* control either induced with IPTG (I) or uninduced (N) was added to each reaction. The reactions were incubated for 30 min at 4°C, centrifuged, washed twice in WB with 30 mM KCL–10 mM $MgCL_2$, and eluted as described in the text. The eluates were lyophilized and analyzed on the 12% SDS-polyacrylamide gels shown here. Of the final eluates, 8% are displayed here. Lanes: 1, 1 μl of extract I containing 15 μg of total protein; 2, 1 μl of extract N containing 11 μg of total protein; 3, proteins from extract N bound to pUC8 DNA; 4, proteins from extract N bound to pUC8-o_L DNA; 5, proteins from extract I bound to pUC8; 6, proteins from extract I bound to pUC8-o_L. The positions of MW standards are indicated.

(data not shown). It is interestng that several minor bands amounting to a small fraction of the total protein present in the eluates also appear to be binding specifically to the lambda sequences. The two largest of these minor bands comigrate with the two largest bands present in the whole-cell extracts and hence can be identified as the beta and beta′ subunits of *E. coli* RNA polymerase (4).

Application of the binding technique to yeast mitochondria. Having demonstrated the efficacy of the method described above in a procaryotic model, we next sought to demonstrate its applicability to a less well-characterized system by identifying the protein(s) which binds to the promoter of the yeast mitochondrial 14S rRNA gene. Figure 3 shows the plasmids used in these experiments. pTB3′+29 contains sufficient sequences to direct the in vitro synthesis of a specific runoff transcript initiated at position +1 of the 14S rRNA gene. pTB3′−31 does not allow detectable specific transcription of the 14S rRNA gene (2). These plasmids differ only by 60 bp centered on position +1 and containing nucleotides −10 to +2 which are essential for accurate in vitro transcription (2). The mitochondrial sequences present in each plasmid consist of A-T-rich (approximately 95%) DNA with no other distinctive features (28). Cleavage of each of these plasmids with *Pvu*II generates a small fragment containing the entire mitochondrial insert and the *lac* operator; these fragments were immobilized on Immunobeads via the *lac* repressor–beta-galactosidase fusion protein and anti-beta-galactosidase as described above. To eliminate the isolation of mitochondrial proteins potentially binding to sequences upstream from position −31, which are not required for specific transcription in vitro (2), the competitor plasmids retained the same upstream mitochondrial DNA sequences through position −234 as those retained by pTB3′−31 and pTB3′+29. The resulting plasmids pBR3′+29 and pBR3′−31 do not contain *lac* operator sequences and hence are not precipitated in the binding reaction.

Identification of a 70,000-MW peptide which binds to the mitochondrial 14S rRNA promoter region. The experiments shown in Fig. 5 directly identify and enrich a 70,000-MW peptide which specifically binds to the promoter region of the 14S rRNA gene. This peptide binds to promoter-containing DNA immobilized on the Immunobeads in the presence of a 50-fold mass excess of competitor DNA deleted of promoter sequences (Fig. 5, lane 4). This peptide is not precipitated by immobilized promoter-deleted DNA (Fig. 5, lanes 2 and 3). In addition, the binding of the peptide is eliminated when the competitor DNA in excess contains functional promoter sequences (Fig. 5, lane 5). (In some experiments, trace amounts of an 85,000-MW sequence-specific DNA-binding protein are also detected [Fig. 5, lane 4].)

The specific binding of the 70,000-MW peptide to immobilized promoter sequences is dependent upon the stoichiometry of the immobilized and free promoter region sequences (Fig. 6). As the molar ratio of bound pTB3′+29/competitor pBR3′+29 is decreased from 1/0 to 1/8 to 1/25 (Fig. 6, lanes 1 to 3, respectively), the precipitation of the 70,000-MW peptide is progressively reduced. The total mass of competitor DNA in the experiments depicted in Fig. 5 and 6 was not varied, indicating that the disappearance of the 70,000-MW-peptide can only be attributed to specific competition by sequences within the 60-nucleotide region contained uniquely on the promoter-plus plasmid.

The presence or absence of the 70,000-MW peptide in various fractions cannot be explained by variation in the efficiency of elution of the protein-DNA complex from the

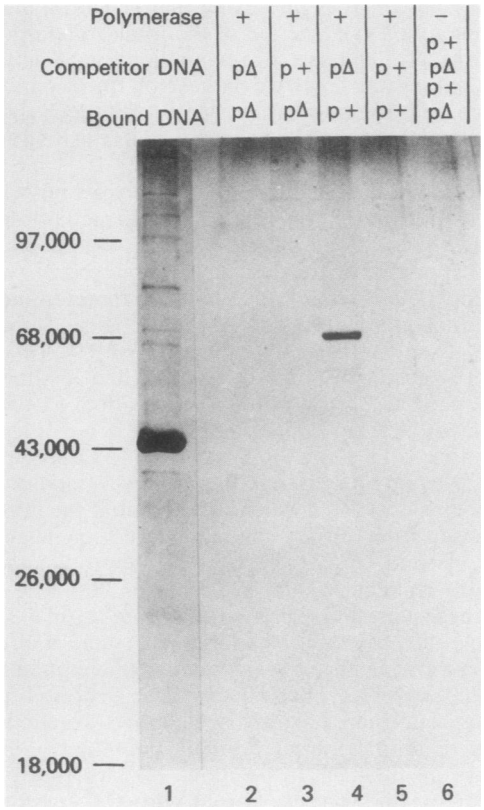

FIG. 5. A 70,000-MW peptide binds to the yeast mitochondrial 14S rRNA promoter region. A 15-μg amount each of pTB3′+29 (p+) and pTB3′−31 (pΔ) DNA cleaved with *Pvu*II was incubated with 10 μg of the *lac* repressor–beta-galactosidase fusion protein and Immunobeads saturated with anti-beta-galactosidase as described in the text. The Immunobeads were washed three times with binding buffer containing 20 mM KCl and suspended in 200 μl of binding buffer with 100 μM ATP–5% glycerol and no KCl or NaCl. The Immunobeads with bound DNA were divided into separate reactions containing the appropriate competitor DNA cleaved with *Hin*fI. RNA polymerase (60 μl) was added last. Final reaction volumes were 420 μl in binding buffer with 5% glycerol–100 μM ATP and approximately 40 to 50 mM KCl supplied by the addition of the RNA polymerase. The reactions were incubated at 15°C for 30 min, centrifuged, and washed once in 0.5 ml of binding buffer with 5% glycerol–100 μM ATP–30 mM KCl and then once with 0.5 ml of the same buffer containing 100 μg of salmon sperm DNA per ml. Finally, the reactions were eluted as described in the text. Samples were analyzed for DNA content (the rest was lyophilized), and 70% of each sample was subjected to electrophoresis on the 10% polyacrylamide gel shown here. The additions to each reaction follow. Lanes: 1, 2 μl of input RNA polymerase; 2, 80 μl of pΔ Immunobeads and 75 μg of pBR3′−31; 3, 80 μl of pΔ Immunobeads and 75 μg of pBR3′+29; 4, 80 μl of p+ Immunobeads and 75 μg of pBR3′−31; 5, 80 μl of p+ Immunobeads and 75 μg of pBR3′+29; 6, 40 μl each of p+ and pΔ Immunobeads and 37.5 μg each of pBR3′+29 and pBR3′−31. Positions of MW standards are indicated. Several faint additional bands seen in reactions without polymerase are artifacts of staining (22, 30).

Immunobeads. Figure 7 shows portions of the eluates from the samples shown in Fig. 5 analyzed for DNA on a 1.5% agarose gel. Equal amounts of *lac* operator-containing restriction fragments and a small quantity of residual salmon sperm DNA are present in each sample. Also shown (Fig. 7, lanes E and F) is the DNA not bound in the reaction which immobilized the *lac* operator-containing fragments. The presence of these fragments demonstrates saturation of the repressor with operator DNA and indicates the mobilities of the various restriction fragments.

Functional assessment of the specifically bound fraction. Recent experiments have shown that preparations of the mitochondrial RNA polymerase can be resolved into a selective fraction which supports accurate transcription in vitro as well as a nonselective catalytic fraction (B. Ticho and G. S. Getz, University of Chicago, personal communication). These results suggest that accurate transcription might require a specificity factor to confer selectivity on the catalytic component. To functionally assess with our method the role of the factors which are bound to DNA, Im-

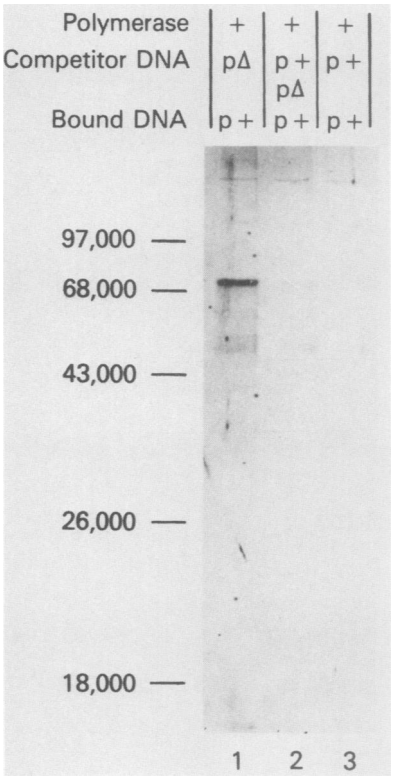

FIG. 6. For each point, 3.0 μg of *Pvu*II-digested pTB3′+29 was reacted with 0.7 μg of *lac* repressor–beta-galactosidase fusion protein, bound, and washed as described in the text. Approximately 50% of the *lac* operator-containing restriction fragment bound in each sample. The Immunobead pellets were suspended to a final reaction volume of 300 μl in binding buffer containing 5% glycerol, 40 mM KCl (supplied by the addition of RNA polymerase), 100 μM ATP, and the additions noted below, RNA polymerase being added last. The samples were incubated for 20 min at 20°C, at which time 40 μg of salmon sperm DNA as additional competitor was added to the samples. All samples were immediately centrifuged, washed once in 0.5 ml of binding buffer with 100 μM ATP and 20 μg of salmon sperm DNA, and then eluted into 300 μl of 10 mM Tris hydrochloride (pH 7.4)–5 mM EDTA–2 mM IPTG for 20 min. The eluates were lyophilized and analyzed on the 10% polyacrylamide gel shown here. The plasmid competitor DNA was digested with *Hin*fI extracted with phenol and then ethanol precipitated before being added to the binding reactions. The additions to each reaction follow. Lanes: 1, 50 μg of pBR3′−31 and 40 μl of RNA polymerase; 2, 15 μg of pBR3′+29, 35 μg of pBR3′−31, and 40 μl of RNA polymerase; 3, 50 μg of pBR3′+29 and 40 μl of RNA polymerase; p+ DNA contains a functional mitochondrial promoter. In pΔ DNA the functional promoter sequences have been deleted.

munobead pellets with proteins bound to the promoter region-containing DNA fragment were allowed to synthesize RNA directed from both bound and exogenously added templates. Faint specific runoff transcripts were produced from both the bound and added templates from binding reactions with immobilized promoter-containing DNA but not with immobilized promoter-deleted DNA (data not shown). Transcripts of 138 and 485 nucleotides, respectively, were directed by the bound *Pvu*II fragment of pTB3′+29 and by added pBR3′+29 cleaved with *Hin*cII. In addition, reactions containing pBR3′+29 cleaved with *Aat*II produced the expected 101-nucleotide transcript. Less than 0.5% of the total polymerizing activity added to the binding reaction was recovered in the pellets (the binding reagents do not inhibit transcription, data not shown), which indicates that catalytic activity is not enriched by this specific binding assay. However, specific transcription in vitro is reconstituted by the method described in this paper with the additions of a fraction containing nonselective activity and a bound fraction prominently enriched for a 70,000-MW peptide (data not shown; see below).

DISCUSSION

Binding of lambda repressor: a demonstration of the method. We developed a method to identify proteins which bind to known regulatory DNA sequences. To verify the utility of this approach, we applied the procedure to extracts of *E. coli* containing the lambda repressor. In a single step, the lambda repressor is brought to approximately 90% of the total protein in the purified sample. In addition to the lambda repressor, it appears that a much smaller amount of *E. coli* RNA polymerase is also specifically precipitated by the technique described here. This could be due to specific binding of *E. coli* RNA polymerase to the lambda phage DNA sequences in pUC8-o_L or to a weak association of polymerase with the lambda repressor bound to DNA. The increased amount of polymerase precipitated from extracts of induced cells containing the greatest amount of lambda repressor would favor the second possibility.

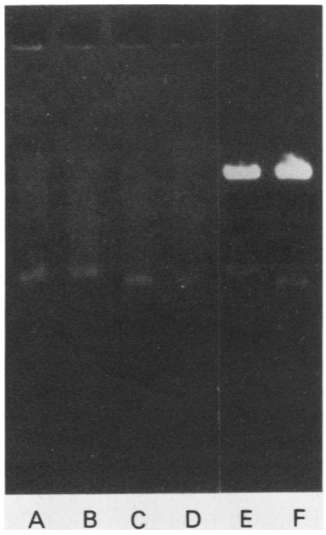

FIG. 7. Promoter-containing and promoter-deleted DNA are recovered equally from binding reactions. Of the total eluate recovered from samples shown in Fig. 5, 6% were analyzed on the 1.5% agarose gel shown here. Reaction additions are listed below. Competitor DNA was restricted with *Hin*fI. Lanes: A, 80 μl of p+ Immunobeads and 75 μg of pBR3′+29; B, 80 μl of p+ Immunobeads and 75 μg of pBR3′−31; C, 80 μl of pΔ Immunobeads and 75 μg of pBR3′+29; D, 80 μl of pΔ Immunobeads and 75 μg of pBR3′−31. Lanes E and F show a portion of the supernatant from the binding reaction which immobilized the *lac* operator-containing restriction fragment and demonstrates DNA excess in that reaction relative to *lac* repressor–beta-galactosidase fusion protein as well as indicating the mobilities of the restriction fragments used in the experiment.

Is the 70,000-MW protein a yeast mitochondrial transcription factor? We also applied this technique to a less well-characterized system in identifying a heretofore unrecognized protein which binds to a 60-bp promoter region of the yeast mitochondrial 14S rRNA gene. The identification and enrichment of this peptide have been made solely on the basis of its binding to a 60-bp region containing sequences essential for specific transcription in vitro. These sequences direct the accurate in vitro initiation of 14S rRNA transcription and have been analyzed extensively with a series of nuclease BAL 31-generated deletion mutants. By using nested deletions extending from upstream or downstream sites toward the initiation site (+1), the region from −10 to +2 was shown to provide all of the information necessary for the correct promotion of RNA synthesis in vitro. The necessary nucleotides for accurate transcription lie squarely in the middle of the 60-base region unique to the promoter-plus plasmids used in the current experiments. The binding of the 70,000-MW peptide to immobilized promoter-containing DNA but not to promoter-deleted DNA in the presence of a 50-fold mass excess of competitor demonstrates its specificity for this region. In addition, the stoichiometric titration of the binding of the 70,000-MW peptide from bound to soluble promoter sequences confirms this observation.

The 70,000-MW peptide which specifically binds to the 60-bp region is only a minor component of a crude, active fraction of yeast mitochondrial RNA polymerase and is distinct from the major 45,000-MW peptide. Although the 70,000-MW peptide is not enriched for catalytic activity, the binding of this factor correlates with a small amount of specific transcription in vitro. A simple model is that the 70,000-MW peptide is a factor which confers transcriptional specificity to unidentified catalytic components. Transcription by partially purified preparations of the yeast mitochondrial RNA polymerase has a second-order concentration dependence consistent with this model (9). The partial resolution by conventional chromatographic techniques of specific and nonspecific transcriptional activities for yeast mitochondria has revealed that selective activity is observed in an overlap region between the nonspecific activity and fractions enriched for a 70,000-MW peptide (Ticho and Getz, personal communication). Although direct addition of this 70,000-MW peptide-enriched fraction to a fraction with nonspecific activity fails to reconstitute specific transcription, binding of the 70,000-MW peptide-enriched fraction to immobilized DNA by the technique described in this paper, followed by the addition of nonspecific activity, restores accurate in vitro transcription and eliminates random RNA synthesis (unpublished data). This result, in combination with the direct demonstration of a 70,000-MW promoter region specific binding protein reported here, strongly suggests a role for the 70,000-MW peptide in the initiation of transcription.

The mechanisms through which promoter specificity is conferred upon RNA polymerases can vary. Some RNA

polymerases, such as those found in some bacteriophage, possess intrinsic capacity for promoter site selection (6). Most bacterial host RNA polymerases rely upon sigma factors which do not bind directly to DNA but do confer specificity upon core syzymes (5). In eucaryotes, transcription factors such as TF IIIA and Sp-1 bind directly to specific sequences independently of RNA polymerase (8, 10). If the 70,000-MW peptide proves to be a transcription factor, it would appear to be similar to other more well-characterized eucaryotic transcription factors in that it possesses specific DNA-binding properties in the absence of RNA polymerase. An alternative role for this 70,000-MW peptide in the regulation of 14S rRNA gene expression cannot be excluded at this time. It should be noted, however, that a similar consensus sequence is present at the initiation sites of transcription from a variety of mitochondrial genes (7).

Small amounts of other peptides also may be enriched from the partially purified yeast mitochondrial RNA polymerase fraction. For instance, an 85,000-MW peptide is also visualized (Fig. 5, lane 4). Because the 60-bp segment used for specific binding is potentially large enough to bind more than one factor, we cannot determine whether the binding of the 85,000-MW peptide is independent of, cooperative with, or complexed to the 70,000-MW peptide. Binding experiments with small, defined deletions in the promoter region may help to resolve this matter and could provide a general means for precisely localizing the specific DNA-binding sites of particular proteins.

Thermodynamic and kinetic constraints of the method. (i) The *lac* repressor–beta-galactosidase fusion protein is a tetramer of 150,000-MW monomers and possesses two identical DNA-binding sites (3, 4, 21). A typical concentration of *lac* repressor–beta-galactosidase protein used in an analytical experiment is approximately 1×10^{-8} M; hence, the maximal specific DNA concentration bound is approximately 2×10^{-8} M. Therefore, any protein with a sequence-specific K_d of 5×10^{-9} M or less may be significantly depleted after several washes. However, many specific sequence binding proteins characterized thus far bind tighter than a K_d of 10^{-9} M under optimal conditions, i.e., simian virus 40 large T antigen, lambda repressor, and cro protein (13, 20, 24, 29).

(ii) The extent of enrichment could be limited by the relative K_d's of a specific protein for the specific DNA sequence bound to the Immunobeads versus the nonspecific sequences present as competitor. In addition, the source and quality of the competitor DNA may influence the binding of the protein of interest.

(iii) The K_d for the binding of nonspecific sequences by the protein of interest as well as the quantity of competitor DNA determines kinetically the rate at which the protein of interest approaches equilibrium binding to its specific regulatory sequence. For example, the *lac* repressor has a K_d of approximately 0.5 μg/ml for non-operator-containing DNA and under the conditions used (approximately a 100-fold excess of competitor DNA, relative to operator, present at a final concentration of 300 μg/ml) should attain 90% equilibrium binding to operator in approximately 150 s (16). Proteins which bind tighter to nonspecific sequences may require decreased amounts of competitor DNA or extended incubations, or both, to approach equilibrium.

(iv) Because the *lac* repressor binds tightly to random DNA (K_d, 0.5 μg/ml) (16), it is important to use sufficient operator-containing DNA to saturate all DNA-binding sites in the repressor–beta-galactosidase fusion protein. This prevents significant amounts of competitor DNA carrying nonspecific proteins from precipitating with the protein of interest.

(v) IPTG increases the K_d of *lac* repressor for *lac* operator from 10^{-13} M to approximately 10^{-10} M, which is still very tight binding (25). Therefore, the presence of long pieces of competitor DNA in the final wash or the use of operator-containing competitor DNA facilitates elution with IPTG by binding to vacant repressor DNA-binding sites, effectively blocking rebinding of eluted DNA and thereby displacing the protein DNA complex of interest into the supernatant of the final centrifugation. Due to the tight binding of the *lac* repressor to the *lac* operator even in the presence of IPTG, fragments bearing multiple *lac* operators are to be avoided when attempting to identify and enrich DNA-binding proteins. Effective elution is prevented by multiple repressor-operator interactions on the same piece of DNA.

This procedure theoretically can be scaled up approximately 100- to 1,000-fold and still be performed on a convenient laboratory scale. In all likelihood, the amount of available extract will be the limiting factor, not the amount or expense of required reagents. All of the reagents are either commercially available or easily prepared in the laboratory.

Potential applications, advantages, and limitations. The method described in this paper has many potential applications. It can be used with ^{35}S-labeled cells for the radiographic identification of sequence-specific DNA-binding proteins from cell extracts or partially purified fractions. Much of the background staining seen in the results shown in this paper is from the added protein reagents and Kodavue staining and not from the crude mitochondrial RNA polymerase. Therefore, with radiolabeled extracts, this portion of the background becomes invisible.

Preliminary experiments with extracts of [^{35}S]methionine-labeled mammalian cells reveal the IPTG elution to be a very specific step which leaves nonspecifically absorbed proteins bound to the Immunobeads. If desired, the nonspecifically bound proteins can be eluted for electrophoretic analysis by the addition of SDS.

Factors recognizing transcriptional regulatory DNA sequences (possibly enhancers) should be identifiable with this method, even if specific for supercoiled DNA. Supercoiled plasmids can be readily and, in fact, preferentially precipitated with this method (data not shown). No covalent modification of the reversibly immobilized DNA is necessary with this technique; therefore, one can be certain of the integrity of the bound supercoiled molecules. Potentially, this technique can be coupled with the use of extrachromosomal vectors such as bovine papillomavirus (27) or yeast plasmids to rapidly identify regulatory proteins bound in vivo to DNA and to study chromatin structure. Additionally, the factors which regulate the expression of genes cloned into these extrachromosomal vectors may be isolated as part of a complex, formed in vivo, with their target sequences.

The technique described in this paper is designed to enrich and identify proteins solely on their ability to bind to specific DNA sequences. It provides no evidence for function other than DNA binding; hence, correlation with additional in vivo or in vitro evidence is required to assign a role to an enriched peptide in replication, recombination, or gene regulation. Because these experiments are performed in DNA excess to drive the equilibrium toward maximal specific DNA binding, techniques which require quantitative or near quantitative protection of specific sequences such as DNase footprinting cannot be performed directly on the eluted DNA-protein

complex. Additional analysis of the specific protein may require elution from its target sequence and removal of nucleic acid. Additional advantages and limitations of this technique will be defined by applications such as those outlined above.

ACKNOWLEDGMENTS

We thank T. Biswas, M. Gottesman, B. Ticho, J. Wettstein, and I. Zabin for materials and helpful advice. We are grateful to S. Mackem and B. Spalholz for critical readings of the manuscript and to N. Freas for its preparation.

LITERATURE CITED

1. **Backman, K., and M. Ptashne.** 1978. Maximizing gene expression on a plasmid using recombination *in vitro*. Cell **13**:65–71.
2. **Biswas, T. K., J. C. Edwards, M. Rabinowitz, and G. S. Getz.** 1985. Characterization of a yeast mitochondrial promoter by deletion mutagenesis. Proc. Natl. Acad. Sci. U.S.A. **82**:1954–1958.
3. **Brake, A. J., A. V. Fowler, I. Zabin, J. Kania, and B. Muller-Hill.** 1978. Beta-galactosidase chimeras: primary structure of a lac repressor-beta-galactosidase protein. Proc. Natl. Acad. Sci. U.S.A. **75**:4824–4827.
4. **Burgess, R. R.** 1976. Purification and physical properties of *E. coli* RNA polymerase, p. 69–100. *In* R. Losick and M. Chamberlin (ed.), RNA polymerase. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
5. **Chamberlin, M.** 1976. RNA polymerase-an overview, p. 17–67. *In* R. Losick and M. Chamberlin (ed.), RNA polymerase. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
6. **Chamberlin, M., and T. Ryan.** 1982. Bacteriophage DNA-dependent RNA polymerases, p. 87–104. *In* P. D. Boyer (ed.), Enzymes, vol. 15. Academic Press, Inc., New York.
7. **Christianson, T., and M. Rabinowitz.** 1983. Identification of multiple transcriptional initiation sites on the yeast mitochondrial genome by *in vitro* capping with guanylyltransferase. J. Biol. Chem. **258**:14025–14033.
8. **Dynan, W. S., and R. Tijan.** 1983. The promoter-specific transcription factor SP 1 binds to upstream sequences in the SV40 early promoter. Cell **35**:79–87.
9. **Edwards, J. C., D. Levens, and M. Rabinowitz.** 1982. Analysis of transcriptional initiation of yeast mitochondrial DNA in a homologous *in vitro* transcription system. Cell **31**:337–346.
10. **Engleke, D. R., S. Y. Ng, B. S. Shastry, and R. G. Roeder.** 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell **19**:717–728.
11. **Fowler, A. V., and I. Zabin.** 1983. Purification, structure, and properties of hybrid beta-galactosidase proteins. J. Biol. Chem. **258**:14354–14358.
12. **Galas, D., and A. Schmitz.** 1978. DNase footprinting: a simple method for detection of protein-DNA binding specificity. Nucleic Acids Res. **5**:3157–3170.
13. **Johnson, A., B. J. Meyer, and M. Ptashne.** 1979. Interactions between DNA-bound repressors govern regulation by the lambda repressor. Proc. Natl. Acad. Sci. U.S.A. **76**:5061–5065.
14. **Johnson, A. D., C. O. Pabo, and A. T. Sauer.** 1980. Bacteriophage lambda repressor and cro-protein: interactions with operator DNA. Methods Enzymol. **65**:839–856.
15. **Levens, D., A. Lustig, and M. Rabinowitz.** 1981. Purification of mitochondrial RNA polymerase from *Saccharomyces cerevisiae*. J. Biol. Chem. **256**:1474–1481.
16. **Lin, S.-Y., and A. D. Riggs.** 1972. *lac* repressor binding to non-operator DNA: detailed studies and a comparison of equilibrium and rate competition methods. J. Mol. Biol. **72**:671–690.
17. **Manley, J. L., A. Fire, P. A. Sharp, and M. L. Gefter.** 1980. DNA-dependent transacription of adenovirus genes in a soluble whole-cell extract. Proc. Natl. Acad. Sci. U.S.A. **77**:3855–3859.
18. **Marmur, J.** 1962. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. **3**:208–218.
19. **Mauer, R., B. J. Meyer, and M. Ptashne.** 1980. Gene regulation at the right operator (O_R) or bacteriophage lambda. J. Mol. Biol. **139**:147–161.
20. **McKay, R. D. G.** 1981. Binding of a simian virus 40 T-antigen-related protein to DNA. J. Mol. Biol. **145**:471–488.
21. **Muller-Hill, B., and J. Kania.** 1974. Lac repressor can be fused to beta-galactosidase. Nature (London) **249**:561–563.
22. **Ochs, D.** 1983. Protein contaminants of sodium dodecyl sulfate-polyacrylamide gels. Anal. Biochem. **135**:470–474.
23. **Penefsky, H. S.** 1977. Reversible binding of P_i by beef heart mitochondrial adenosine triphosphatase. J. Biol. Chem. **252**:2891–2899.
24. **Ptashne, M., K. Backman, M. Z. Humayan, A. Jeffrey, R. Mauer, B. Meyer, and R. T. Sauer.** 1976. Autoregulation and function of a repressor in bacteriophage lambda. Science **194**:156–161.
25. **Riggs, A. P., R. F. Newby, and S. Bourgeois.** 1970. *lac* repressor-operator interaction. IV. Effect of galactosides and other ligands. J. Mol. Biol. **51**:303–314.
26. **Ruther, U.** 1982. pUR250 allows rapid chemical sequencing of both DNA strands of its inserts. Nucleic Acids Res. **10**:5765–5772.
27. **Sarver, N., P. Gruss, M.-F. Law, G. Khoury, and P. M. Howley.** 1981. Bovine papilloma virus deoxyribonucleic acid: a novel eucaryotic cloning vector. Mol. Cell. Biol. **1**:486–496.
28. **Sor, F., and H. Fukuhara.** 1982. Nucleotide sequence of the small ribosomal RNA gene from the mitochondria of *Saccharomyces cerevisiae*, p. 255–262. *In* P. Slonimski, P. Borst, and G. Attardi (ed.), Mitochondrial genes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Takeda, Y., A. Folkmanis, and H. Echols.** 1977. Cro regulatory protein specified by bacteriophage lambda. J. Biol. Chem. **252**:6177–6183.
30. **Tasheza, B., and G. Desseu.** 1983. Artifacts in sodium dodecyl sulfate-polyacrylamide gel electrophoresis due to 2-mercaptoethanol. Anal. Biochem. **129**:98–102.