MOLECULAR AND CELLULAR BIOLOGY, Aug. 1988, p. 3406–3414
0270-7306/88/083406-09$02.00/0

Vol. 8, No. 8

Purification and Characterization of a High-Mobility-Group-Like DNA-Binding Protein That Stimulates rRNA Synthesis In Vitro

HSIN-FANG YANG-YEN AND LAWRENCE I. ROTHBLUM*

Department of Pharmacology, Baylor College of Medicine, Houston, Texas 77030

Received 1 February 1988/Accepted 2 May 1988

A 16,000-dalton, high-mobility-group-like (HMG-like) DNA-binding protein, referred to as p16, has been purified to homogeneity from Novikoff hepatoma ascites cells. p16 binds specifically to a portion of the 5′ flanking region of the rat rRNA gene (−620 to −417), which is part of the upstream activator sequence identified previously (B. G. Cassidy, H.-F. Yang-Yen, and L. I. Rothblum, Mol. Cell. Biol. 6:2766–2773, 1986). p16 also binds to a segment of the external transcribed spacer (+352 to +545). In vitro reconstituted transcription experiments demonstrated that the addition of p16 stimulated rRNA synthesis up to ca. fourfold. The stimulation was dose dependent and saturable. The effect of p16 on ribosomal gene transcription was also dependent on the presence of either the upstream or the downstream DNA-binding site, or both. The amino acid composition of p16 is very similar to that of HMG-I, suggesting that p16 may be a member of the HMG-I family of proteins. In this case, our results suggest that HMG proteins may play an important role in the regulation of the rRNA gene expression.

Eucaryotic rRNA genes (rDNA) are distinct from genes transcribed by RNA polymerase II or III in that: (i) rRNA genes are transcribed by an α-amanitin-resistant polymerase (class I), (ii) rRNA genes code for the most abundant RNA species in the cell, (iii) rRNA genes are the only transcription unit localized in the nucleolus, and (iv) their transcription is, for the most part, species-specific transcription (13, 17).

rRNA genes are arranged as tandemly repeating units which contain a transcribed (or coding) region and a nontranscribed spacer (NTS). The transcribed domain consists of external transcribed spacers (ETS), the internal transcribed spacers, and the regions coding for the 18S, 5.8S, and 28S RNAs, respectively (30). The functions of the NTS, which is the major portion of the repeat, are just being elucidated.

There are several important transcriptional elements within the NTS. The core promoter element, located between nucleotides −39 and +6 (+1 being the initiation site of the precursor rRNA), is required for the correct initiation of transcription (14, 25, 29, 33, 60). The upstream promoter (or control) element, located between nucleotides −150 and −110, has been shown to be able to effectively stabilize the preinitiation complex (3, 20, 38, 52, 55). Transcription terminators have been identified both downstream and upstream of the initiation site (15, 16, 36). These elements either direct the termination of rRNA synthesis or are sites of rapid RNA processing. In addition, the terminator that lies ca. 200 base pairs (bp) upstream of the initiation site (referred to as T_0 in the mouse or as T_3 in the *Xenopus* rDNA repeat) may act as an upstream element of the adjacent promoter (15, 22, 36). The spacer or NTS promoter (5, 21, 40, 46), which is much weaker than the 45S (or 40S in *Xenopus* species) gene promoter, may also play a role in regulating rRNA synthesis. In addition, enhancer elements (4, 7, 8, 27) have been identified at various distances upstream of the transcription initiation sites of the respective genes.

The development of accurate in vitro transcription systems (2, 13, 19, 34, 56) has led to the identification of factors required for the synthesis of the rRNA (37, 39). Studies on the rDNA terminator and enhancer elements suggest that *trans*-acting factors are involved in these two functions (4, 16, 44).

We have identified the upstream activator sequence (UAS) (or enhancer element) of the rat rRNA gene that is located between nucleotides −1018 (*Kpn*I site) and −286 (*Bam*HI site) (4) and shown that a *Hin*fI fragment (−1106 to −417) encompassing and overlapping the putative UAS element was recognized by a cellular factor (4). We now report the purification of the DNA-binding protein (referred to as p16) to near homogeneity (>95%) by using a gel electrophoresis DNA-binding assay. p16 is very similar to the high-mobility-group (HMG), nonhistone chromosomal protein (24), especially HMG-I (32), on the basis of its size and other properties. The similarity is of interest because some of the HMG proteins have been shown to be preferentially associated with the chromatin and nucleosomes of active genes (23, 47, 54, 57, 58) and to have an effect on transcription in vitro (53).

We found that p16 bound to two regions of the rDNA, one site within the NTS (nucleotides −620 to −417) and the other within the ETS (nucleotides +352 to +545). More importantly, we found that the addition of p16 to in vitro transcription reactions stimulated rRNA synthesis up to ca. fourfold, as long as either one or both of the p16-binding sites remained in *cis* with the promoter. Since one of the p16-binding sites was localized to the region that contained the upstream activator or enhancer sequence (4), as well as the NTS promoter (5), p16 may play a significant role in the regulation of rDNA expression.

MATERIALS AND METHODS

Plasmid DNA. p3.4 is a plasmid containing 3.4 kilobases of the rat rDNA 5′ NTS (nucleotides −286 to −3687, +1 being the initiation site of rDNA transcription) cloned into pBR322 as described previously (41, 61). Plasmid pUCHF3 was generated by subcloning a 688-bp *Hin*fI fragment (4) (nucleotides −1106 to −417) from p3.4 into the *Sma*I site of pUC12

* Corresponding author.

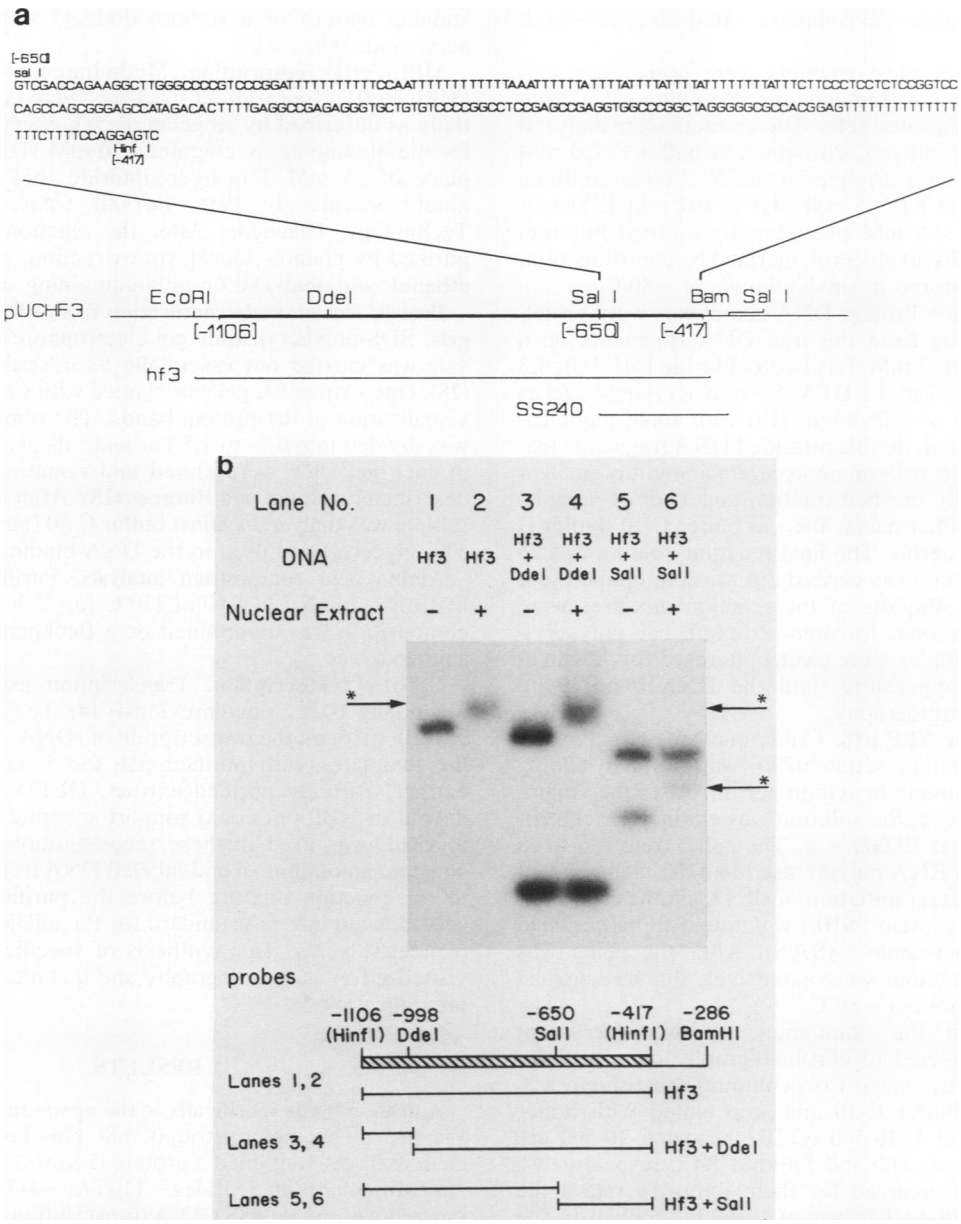

FIG. 1. (a) Schematic diagram of plasmid pUCHF3. pUCHF3 was constructed by subcloning a 688-bp *Hin*fI fragment (4) (nucleotides −1106 to −417) from plasmid p3.4 into the *Sma*I site of pUC12. The nucleotide sequence between −650 (*Sal*I) and −417 (original *Hin*fI site) is shown above the map. DNA fragments hf3 and SS240, which are used in the DNA-binding assay, are also indicated. ▬, Sequences from the insert; ——, vector (pUC12). (b) Mapping of the NBP (p16)-binding site. End-labeled hf3 fragment (lanes 1 and 2) was digested with either *Dde*I (lanes 3 and 4) or *Sal*I (lanes 5 and 6). These fragments were analyzed in the standard DNA-binding assay by using 1.84 μg of nuclear extract, as described in Materials and Methods. Symbols: $\overset{*}{\rightarrow}$, bound form of an individual fragment; +, −, presence and absence, respectively, of that component in the assay. Lower panel: Schematic representation of the *Sal*I and *Dde*I sites.

(Fig. 1a). Other plasmids (or DNA templates) containing portions of the rDNA gene used in this study are p5.1E/X, p5.1ΔH, p1018, and p1742d (see Fig. 7a) (B. G. Cassidy, Ph.D. thesis, Baylor College of Medicine, Houston, Tex., 1987). p5.1E/X (−286 to +638) was cloned into the *Eco*RI and *Bam*HI sites of pBR322. p5.1ΔH (−286 to +124) was constructed by inserting the *Bam*HI-*Hin*dIII fragment from p5.1E/X into the *Bam*HI-*Hin*dIII sites of pBR322. p1018 (−1018 to +638) was constructed by isolating the *Kpn*I (converted to an *Sph*I site)-*Bam*HI fragment from p3.4 and cloning it into the *Sph*I-*Bam*HI sites of p5.1E/X. p1742 (−1742 to +638) was constructed by isolating the *Xba*I (converted to an *Sph*I site)-*Bam*HI fragment from p3.4 and inserting it into the *Sph*I-*Bam*HI sites of p5.1E/X. p1742d, a template used in the transcription assays, was derived by double digestion of p1742 with *Hin*dIII (+124) and *Eco*RI (+638).

Enzymes. Restriction enzymes were purchased from New England BioLabs, Inc., Bethesda Research Laboratories, Inc., or Boehringer Mannheim Biochemicals and were used as recommended by the supplier.

DNA labeling. DNA fragments were 5′ end labeled with [γ-^{32}P]ATP (Amersham Corp.) by using T4 polynucleotide kinase (Bethesda Research Laboratories) after digestion

with bacterial alkaline phosphatase (Bethesda Research Laboratories).

Nuclear extract. Nuclear extracts were prepared essentially as described by Haglund and Rothblum (19) from Novikoff hepatoma ascites cells. The extracts were dialyzed extensively against buffer C-20, which is buffer C (20 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.9], 0.1 M KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 1 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride) containing 20% (vol/vol) glycerol, clarified by centrifugation, quick frozen, and stored in small aliquots at −80°C.

DNA-binding assay. Protein-DNA complexes were visualized after separation from the free DNA fragments on a low-ionic-strength (6.7 mM Tris hydrochloride [pH 7.5], 3.3 mM sodium acetate, 1 mM EDTA) 5% polyacrylamide gel as described by Fried and Crothers (10) with some modifications. The end-labeled, double-stranded DNA fragments (ca. 2 ng) were incubated with an appropriate amount of nuclear extract or a partially purified fraction and 1 μg of poly(dI-dC) · poly(dI-dC) (Pharmacia, Inc.) in buffer C-10 (buffer C containing 10% glycerol). The final reaction volume was 25 μl, and the incubation was carried out at room temperature for 10 to 20 min. Aliquots of the reaction mixtures were loaded onto 16-cm-long, low-ionic-strength 5% polyacrylamide gels. The samples were electrophoresed for 90 min at 160 V at room temperature, and the DNA bands were visualized by autoradiography.

Purification of the NBP p16. Crude nuclear extracts were adjusted with ammonium sulfate to 70% saturation by adding 0.472 g of solid ammonium sulfate per ml. After the ammonium sulfate dissolved, the solution was clarified by centrifugation for 20 min at 39,000 × *g*. The pellet (referred to as Dig X/C) contained RNA polymerase I and the factor(s) that could direct the correct initiation of rRNA synthesis in vitro (19). NTS-binding protein (NBP) was found to be enriched 10-fold in the supernatant (AS70S). After the pellet was suspended, both fractions were extensively dialyzed against buffer C-20 and stored at −80°C.

The proteins in the ammonium sulfate supernatant (AS70S) were subjected to chromatography on a heparin-Sepharose (Sigma Chemical Co.) column. The column was equilibrated with buffer C-10 and step eluted with buffer C-10, C-10-300, and C-10-500 (C-10-300 and C-10-500 are buffer C-10 containing 300 and 500 mM KCl, respectively). The fractions were assayed for their ability to retard the mobility of the *Sal*I-*Sal*I fragment (Fig. 1a, SS240) in the band shift assay. The DNA-binding activity was enriched in the fractions eluted with buffer C-10-500, termed HS500. This fraction was concentrated by ultrafiltration with an Amicon apparatus equipped with a YM-10 membrane and loaded directly onto an Ultrogel ACA54 (LKB Instruments, Inc.) gel filtration column (1.6 by 95 cm). The column was equilibrated with buffer C-10-500, and the flow rate was 9 ml/h. The DNA-binding activity eluted as a ca. 30-kilodalton (kDa) protein. The peak binding fractions, referred to as GF30, were pooled and dialyzed against buffer C-20.

GF30 was then fractionated on a CM-Sephadex (Pharmacia) column. The column was equilibrated with buffer C-10 and step eluted with 100, 200, and 500 mM KCl. Sometimes the protein was eluted with a 200 to 600 mM KCl gradient elution after the 200 mM KCl wash step. Peak binding fractions, which eluted with 500 mM KCl (step elution) or 280 to 350 mM KCl (gradient elution), termed CM500, were pooled and dialyzed against buffer C-20. This fraction contained a single protein, which was >95% pure for NBP as judged from either the Coomassie blue or silver staining pattern of a sodium dodecyl sulfate (SDS)-polyacrylamide gel.

MPE-Fe(II) footprinting. Methidiumpropyl-EDTA Fe(II) [MPE-Fe(II)] footprinting reactions were carried out essentially as described by Angel et al. (1), except that the buffer for the binding assay contained 20 mM HEPES (pH 7.9) in place of 25 mM Tris hydrochloride (pH 7.9). MPE was kindly provided by Peter Dervan, California Institute of Technology, Pasadena. After the reaction, the DNA was purified by phenol-chloroform extraction, precipitated with ethanol, and analyzed on urea-containing sequencing gels.

Protein elution and renaturation from SDS-polyacrylamide gels. SDS-polyacrylamide gel electrophoresis (PAGE) analysis was carried out essentially as described by Laemmli (28). One strip of the gel was stained with Coomassie blue for visualization of the protein bands. The remainder of the gel was divided into 0.5- to 1.5-cm-wide slices, and the proteins in each gel slice were eluted and renatured essentially as described by Hager and Burgess (18). After renaturation, the sample was dialyzed against buffer C-50 (buffer C containing 50% glycerol) and used in the DNA-binding assay.

Amino acid composition analysis. Purified protein was hydrolyzed in 5.7 N HCl at 110°C for 22 h. The amino acid composition was determined on a Beckman 121 MB amino acid analyzer.

***In vitro* transcription.** Transcription assays were done essentially as described previously (4). To examine the effect of NBP (p16) on the transcription of rDNA, we preincubated the templates with purified p16 for 5 min before adding extract. Partially purified extract, DE175 (19), which was devoid of NBP but could support accurate rRNA synthesis in vitro, was used in these reconstitution experiments. A constant amount of an end-labeled DNA fragment was added to the reaction mixture before the purification steps and served as an internal standard for the efficiency of recovery of nucleic acids. The synthesis of specific transcripts was visualized by autoradiography and quantitated as described previously (4, 5).

RESULTS

A protein binds specifically to the upstream region of rDNA. We previously demonstrated that Novikoff hepatoma nuclear extracts contained a protein factor(s) that recognized a *Hin*fI fragment (nucleotides −1106 to −417) from the region upstream of the rat 45S rRNA transcription initiation site (4). This region of the gene overlaps the putative upstream activator sequence in vitro (4). Treatment of the nuclear extract with proteinase K (1 mg/ml) at 0°C for 30 min eliminated the interaction between the nuclear extract and this fragment (results not shown), indicating that the DNA-binding factor was a protein or at least contained a proteinaceous component required for the recognition of the specific rDNA fragment.

To define the region within the *Hin*fI fragment recognized by the protein(s) in the nuclear extract, ^{32}P-labeled *Hin*fI fragment (hf3) was isolated and digested with either *Dde*I or *Sal*I, generating smaller, labeled fragments (Fig. 1b). The hf3 fragment (Fig. 1b, lane 2), the larger *Dde*I fragment (nucleotides −998 to −417) (lane 4), and the smaller *Sal*I fragment (nucleotides −650 to −417) (lane 6) demonstrated mobility shifts when incubated with nuclear extracts, whereas the smaller *Dde*I fragment (ca. 110 bp) and the larger *Sal*I fragment (ca. 460 bp) were not shifted. These two fragments thus served as a negative control for the mobility shift assay. These results suggest that the region between the *Sal*I

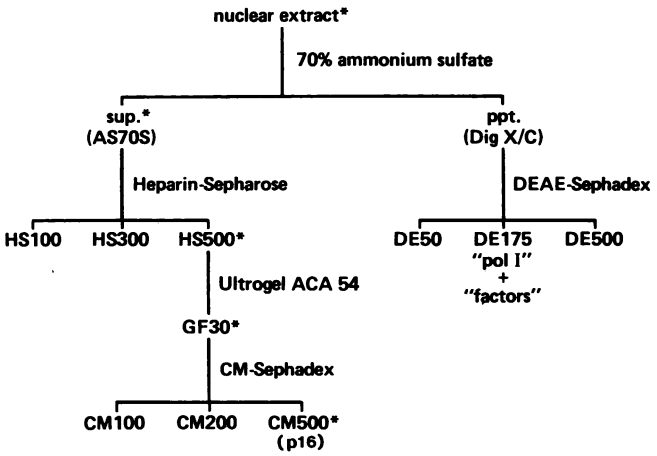

FIG. 2. Scheme for purification of NBP (p16) from nuclear extracts of Novikoff hepatoma cells. Symbol: *, fractions that contain NBP (p16) activity.

(nucleotide −650) and *Hin*fI (nucleotide −417) sites has a sequence that binds specifically to a protein(s) present in the nuclear extract. Long stretches of AT sequences present in this region of the DNA (Fig. 1a) may be important for this specificity. The *Sal*I-*Sal*I fragment (Fig. 1a, SS240) was then used in the mobility shift assay during the purification.

Purification of the NBP p16. The general protocol used to purify the NBP is shown in Fig. 2. When crude nuclear extract was adjusted with ammonium sulfate to 70% saturation, NBP was enriched in the supernatant fraction (AS70S). This fraction represented about 10% of the total proteins present in the crude nuclear extract, but contained ca. 90% of the NBP activity. NBP was then further purified through fractionation on heparin-Sepharose, Ultrogel ACA54, and carboxymethyl-Sephadex columns. Chromatography over heparin-Sepharose was a critical step in which the NBP was enriched at least 100-fold from the AS70S fraction. The binding activity was in the fractions that eluted with 500 mM KCl (HS500) (data not shown). The HS500 fraction was subjected to chromatography on a calibrated Ultrogel ACA54 gel filtration column (Fig. 3). The peak of the binding activity eluted in fractions 41 to 43, equivalent to the elution volume for a 30-kDa globular protein (see inset in Fig. 3). (As the protein was purified, we noted the appearance of multiple bands in the gel retardation assay. The nature of these extra complexes is unknown.) The three fractions containing the peak of the DNA-binding activity were pooled (GF30). This pooled fraction represented 16 to 20% of the proteins applied to the ACA54 column. GF30 was further purified by carboxymethyl-Sephadex column chromatography. NBP was enriched in the fractions that eluted at 500 mM KCl (data not shown) when the column was step eluted with buffers containing 100, 200, and 500 mM KCl. When the carboxymethyl-Sephadex column was eluted with a 200 to 600 mM KCl gradient, the peak of the DNA-binding activity eluted at 280 to 350 mM KCl (data not shown). The peak fractions from the carboxymethyl-Sephadex column (step or gradient elution) (CM500) represent ca. 20% of the total protein applied. We estimate that the overall recovery of the NBP throughout the purification steps was 50%; thus, the NBP has been purified about 12,000-fold from the nuclear extract. SDS-PAGE analysis of the protein components of CM500 demonstrated a single polypeptide with an apparent

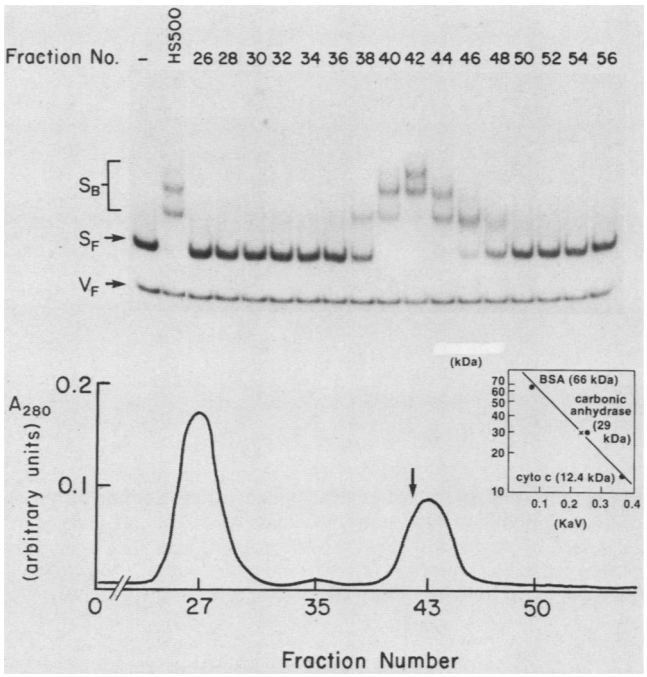

FIG. 3. Gel filtration (Ultrogel ACA54) column chromatography of NBP (p16). Standard DNA-binding assays were performed by the DNA fragment SS240 (Fig. 1a) and the protein fractions eluted from the column. The upper panel shows the results of the DNA-binding assay. The protein-DNA complex is indicated as S_B. S_F denotes the free form of labeled SS240. V_F represents a nonspecific fragment, derived from the vector (pUC12), which was added to the assays as an internal control for specificity. The lower panel shows the protein elution profile of the ACA54 column. The arrow indicates the position of the peak of the DNA-binding activity. The inset in the lower panel shows the calibration of this ACA54 column (1.6 by 95 cm) with protein standards (bovine serum albumin [BSA], carbonic anhydrase, and cytochrome *c*); ×, elution volume of p16.

molecular weight of 16,000 (Fig. 4). Therefore, we refer to this protein as p16.

Amino acid composition analysis of p16. The amino acid composition of p16 (Table 1) is similar to that of members of the HMG family of proteins, especially HMG-I (12, 32) (Table 1). Preliminary protein sequencing data indicated that the N-terminal 13 amino acids of p16 were very similar, although not identical, to the putative N-terminal sequences of HMG-I (31) (data not shown). These results suggest either that p16 is a new member of the HMG-I family of proteins or that it belongs to a new class of proteins that are closely related to HMG-I.

The purified NBP is not a degradation product. To determine whether p16 was the protein responsible for the original DNA-binding activity, and not the degradation product of a higher-molecular-weight precursor, we carried out the following experiment. The crude AS70S fraction was separated by SDS-PAGE, the gel was divided into regions, and the proteins from each slice were eluted, renatured (18), and tested for their specific DNA-binding activity (Fig. 5). Only the proteins from gel slice H (Fig. 5) contained the activity. Since gel slice H contained proteins within the 14- to 20-kDa range, these results suggested that p16 was not a degradation product of a higher-molecular-weight polypeptide.

Since p16 behaved as a 30-kDa polypeptide during gel filtration column chromatography and was a 16-kDa polypeptide on SDS-PAGE, the possibility exists that native p16

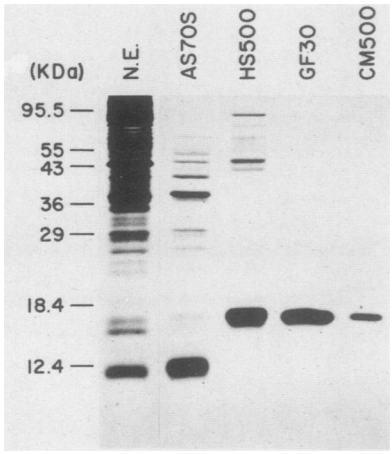

FIG. 4. SDS-PAGE analysis of the proteins obtained at each purification step. Protein samples from each step were analyzed on an SDS-containing 15% polyacrylamide gel. The proteins were visualized by staining with Coomassie brilliant blue. The amount of protein loaded in each lane was 74 μg (nuclear extract, N.E.), 34 μg (AS70S), 10 μg (HS500), 3 μg (GF30), and 0.34 μg (CM500).

TABLE 1. Amino acid composition

Amino acid	% Amino acid found in:		
	HMG-I (rat thymus)[a]	HMG-I (HeLa)[b]	p16[c]
Asx	4.9	3.0	3.0 ± 0.4
Thr	7.1	7.2	8.3 ± 0.6
Ser	11.9	12.2	13.2 ± 0.7
Glx	17.8	18.6	12.6 ± 2.0
Pro	9.0	10.5	12.0 ± 0.6
Gly	11.1	10.4	11.7 ± 0.7
Ala	5.2	5.3	4.1 ± 1.1
Cys	–	–	–
Val	4.7	3.2	3.9 ± 0.3
Met	–	–	–*
Ile	1.0	1.2	1.1 ± 0.1
Leu	1.6	3.3	3.1 ± 0.2
Tyr	–	–	0.3 ± 0.0
Phe	–	–	0.2 ± 0.0
His	0.6	0.4	–*
Lys	16.5	15.0	18.0 ± 1.5
Arg	9.2	9.0	8.8 ± 0.6

[a] Taken from reference 12.
[b] Taken from reference 32.
[c] Average of three independent preparations. –*, Trace amount of this amino acid was detected in only one of the three preparations.

is a homodimer of the 16-kDa polypeptide. Preliminary cross-linking experiments with the purified protein indicated that only a small fraction, less than 5%, of the p16 present in the purified material could be cross-linked with dimethylsuberimidate in vitro (results not shown), indicating that it is unlikely that the protein is a dimer. It is more likely to be an elliptical or rodlike protein which would behave as a protein of higher molecular weight on the molecular sieve column during chromatography.

Recognition sequence of p16. We have used MPE-Fe(II) to footprint the DNA-protein complex. Purified p16 protected several regions between nucleotides −620 and −417 (Fig. 6a and b). This region of the DNA is characterized as having several long stretches of poly(dA-dT) sequences (Fig. 1a). p16 seems to have a higher affinity for sequences with eight or more consecutive thymidines (or adenines), e.g., nucleotides −620 to −610, −605 to −594, and −442 to −429, than for sequences with other combinations of A and T nucleotides, since only the former sequences were completely protected when smaller amounts of p16 were added to the assay mixture (Fig. 6a, lane 3). The binding of p16 to this A+T-rich region is also very specific, since the sequence AAAAGAAA, adjacent to the −429 position, was not recognized by p16 (Fig. 6b, right panel). We previously reported that the ETS of the rat rDNA gene contains a region which is enriched in A+T-rich sequences (+352 to +545) (see Fig. 3 in reference 45). Using the band shift assay, we determined that p16 also recognized those sequences (data not shown). Footprinting experiments confirmed that p16 indeed recognized the A+T-rich sequences within the ETS segment: sequences from +352 to +406 and from +501 to +545 (Fig. 6c). These results indicate that there are at least two p16-binding regions within the rDNA repeat. One is localized in the NTS (or upstream) region (sequences −620 to −417), and the other is within the ETS (or downstream) region (sequences +352 to +545).

Transcriptional role of p16. The original NTS binding activity demonstrated by p16 was of interest since this protein bound to an rDNA region responsible in *cis* for stimulating rRNA synthesis in vitro (4). Thus, we examined the possibility that the addition of p16 to an in vitro transcription system would result in an elevated level of transcription.

The addition of p16 to a transcription reaction mixture that consisted of a template with the NTS binding region in *cis* with the promoter (p1742d [Fig. 7a]) resulted in a three- to fourfold stimulation of transcription (Fig. 7b, lanes 1 and 2). The transcription of a control template, which did not contain this binding site (p5.1ΔH [Fig. 7a]), was unaffected by the addition to p16 (Fig. 7b, lanes 3 and 4). Since p16 also recognized sequences downstream of the rDNA initiation site, we examined the effect of p16 on the transcription of such a template (p5.1E/X [Fig. 7a]). We found that p16 could

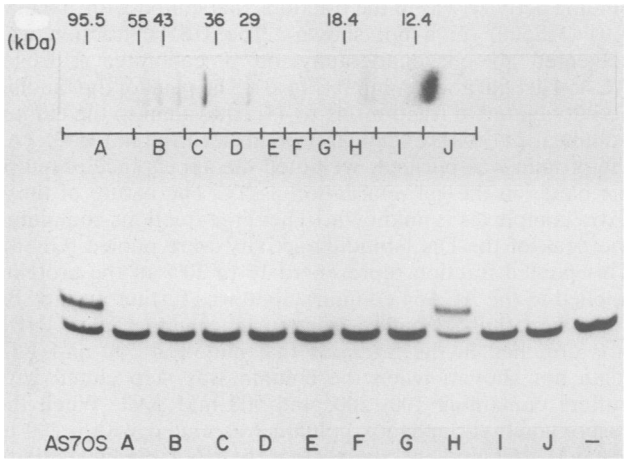

FIG. 5. Identification of the proteins in the crude extract that contains NTS-binding activity. Proteins from the AS70S fraction were fractionated on an SDS-containing 15% polyacrylamide gel. The gel was then sliced into several regions, designated A through J, as shown at the top of the figure. Proteins from each gel slice were eluted, renatured as described by Hager and Burgess (18), and analyzed for their specific DNA-binding activity (bottom panel). Only proteins that eluted from gel slice H contained the same activity as p16.

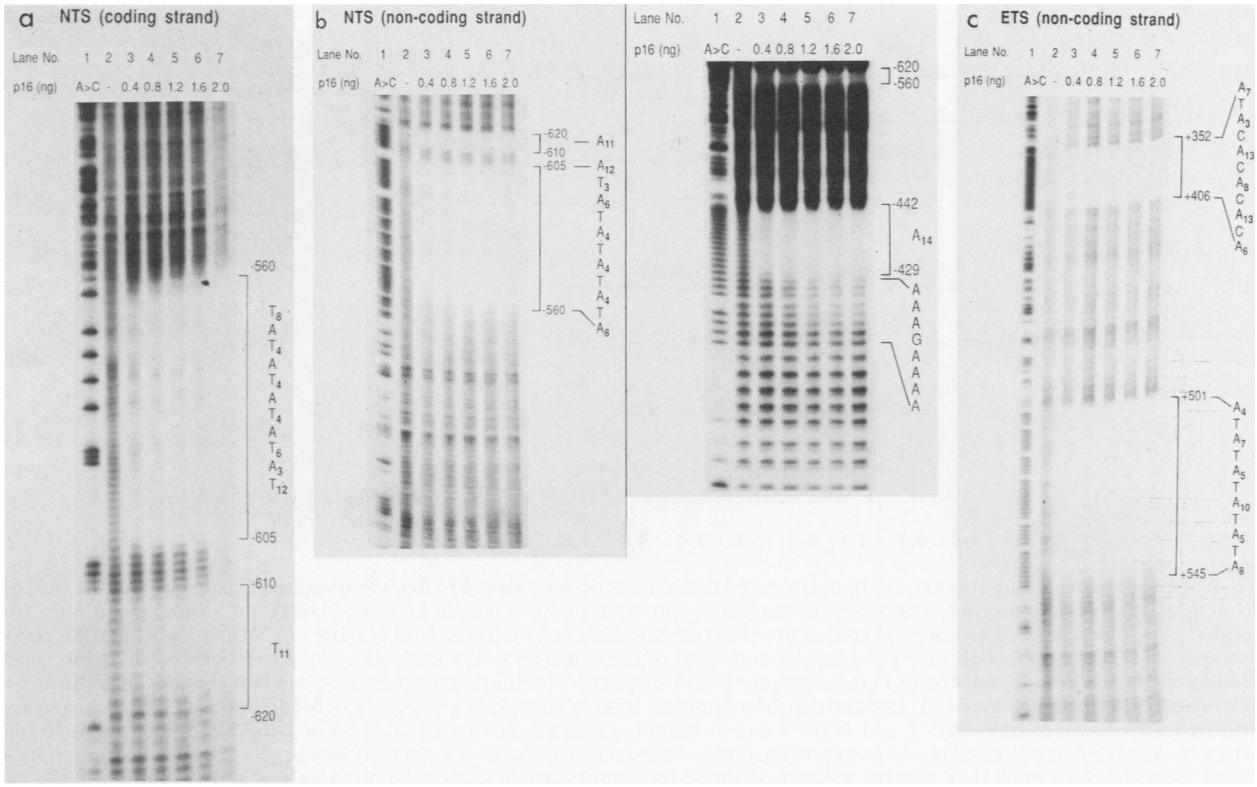

FIG. 6. Binding of purified p16 to the NTS and the ETS regions of the rat rDNA gene. (a) The coding-strand probe, the *Sal*I (−650)-*Bgl*I (−468) fragment, was prepared by 5′ end labeling the *Sal*I site of pUCHF3. (b) Two noncoding-strand probes, fragments *Bam*HI (in the vector next to −417)-*Sal*I (−650) (Fig. 1a) (b) and *Eco*RI (+638)-*Hin*dIII (+124) (Fig. 7a) (c), were 5′ end labeled at the *Bam*HI and *Eco*RI sites, respectivley. Approximately 4 to 8 ng of labeled DNA was incubated with increasing amounts of purified p16 (0.4 to 2.0 ng; lanes 3 to 7 in panels a through c), subjected to partial digestion by MPE-Fe(II), and analyzed on urea-containing sequencing gels. Footprinted regions are indicated by brackets. The right side of panel b is an analysis of the same reaction products as in the left side (6% gel), but displayed on a 20% gel to show the binding site that lies between positions −429 and −442. Control reaction mixtures (lane 2) contained 2 ng of bovine serum albumin. The A>C sequencing ladders (35) (lane 1) of the same probe were run adjacent to each footprinting reactions to identify the sequences of the binding sites.

stimulate the transcription of p5.1E/X up to ca. fourfold (Fig. 7c, lanes 1 to 4). We also found that the transcription of templates with both binding regions (p1018 [Fig. 7a]) was stimulated to approximately the same extent (Fig. 7c, lanes 5 to 8) as the transcription of templates with either the upstream or the downstream binding region. These experiments indicated that the maximal effect of p16 on transcription required only one of its binding domains to be in *cis* to the promoter and that the binding region could be located either upstream or downstream of the promoter. In this regard it should be noted that p1742d contained a second RNA polymerase I promoter (5) upstream of the p16-binding site in the NTS region and that transcription from this promoter was also stimulated (Fig. 7b, arrowhead in lanes 1 and 2). The lack of stimulation of templates that did not contain p16-binding sites was significant, since it indicated that the stimulation of transcription was not the result of a nonspecific anti-inhibitor, e.g., an inhibitor of RNase, or of nonspecific DNA binding.

The effect of p16 was both proportional to the amount of p16 added to the assay and saturable. The addition of increasing amounts of p16 to a transcription reaction mixture resulted in elevated levels of transcription (Fig. 7c). However, in this case the addition of more than 40 ng of p16 resulted in a slightly decreased level of transcription (results not shown). This was probably due to nonspecific DNA binding at that ratio of p16 to template. The stimulation was also specific for p16, since a fraction devoid of p16, HS300, generated during the purification of p16 (see Materials and Methods), had no effect on transcription (data not shown).

DISCUSSION

We have purified an rDNA-binding protein, p16, from Novikoff hepatoma cells to near homogeneity. Protein p16 recognized two regions of the rat rDNA gene in vitro. One of these regions, which lies upstream of the transcription initiation site (between nucleotides −620 and −417), is part of the UAS (4). The second region is within the ETS (between +352 and +545) (45). Both of these binding domains contained long, pure dA-dT sequences (Fig. 1a and 3 in reference 45). A search of the GenBank (MicroGenie, Beckman Instruments, Inc.) revealed that sequences containing 10 to 20 consecutive thymidines (or adenines) are quite common. However, sequences containing ≥45 base pairs of pure A and T, as those sequences present in the rDNA region, are relatively rare. Footprinting experiments showed that these dA+dT-rich sequences were the recognition sites of p16.

We have also shown that when either the upstream (NTS) or the downstream (ETS) binding region (or both) was in *cis* with an rDNA promoter, the addition of p16 to a transcrip-

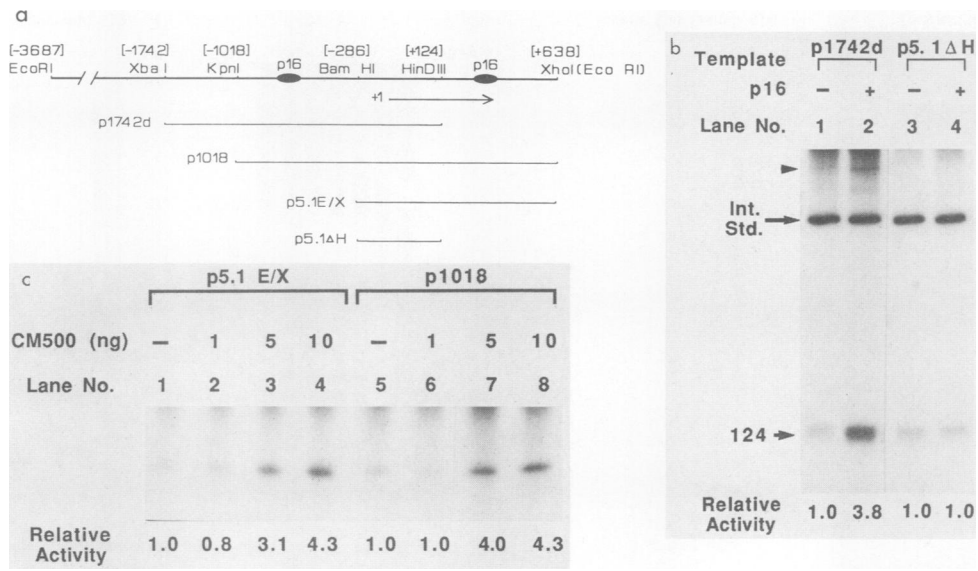

FIG. 7. Effect of p16 on transcription. (a) Templates used in the transcription assay. p5.1ΔH was truncated at a unique *Hin*dIII site, which results in a 124-nucleotide transcript; otherwise, the templates were truncated at a unique *Xho*I (or *Eco*RI) site, resulting in 638-nucleotide transcripts. p1742d was derived by double digestion of p1742 (see Materials and Methods) with *Hin*dIII (+124) and *Eco*RI (+638), releasing the downstream (ETS) p16-binding site. The resulting transcript of this template is 124 nucleotides in length (marked as 124 in panel b). p16-binding sites are shown as solid ovals in the diagram. (b) p16 stimulates the transcription of a template that contains only the upstream binding region, p1742d (lanes 1 and 2), but not that of a template lacking any binding region, p5.1ΔH (lanes 3 and 4). Templates were preincubated with 10 ng of p16 (lanes 2 and 4) for 5 min at room temperature before the addition of DE175 and nucleotides to initiate transcription. The RNA synthesized in vitro was purified and analyzed on an 8 M urea–4% polyacrylamide gel. The specific transcripts were visualized by autoradiography. The relative amount of transcript synthesized in each reaction is shown at the bottom of the figure. Abbreviations: 124, specific transcript 124 nucleotides in length; Int. Std., internal standard (see Materials and Methods). An arrowhead points to the transcripts initiated from the NTS promoter (see text). (c) p16 stimulates the transcription of templates containing the ETS binding region, p5.1E/X (lanes 1 to 4) and that of a template with both binding regions in *cis* with the promoter, p1018 (lanes 5 to 8). The amount of p16 (labeled as CM500) added to each assay was as indicated (1 to 10 ng). The relative amount of transcription of each reaction is indicated at the bottom of the figure. —, No p16 (fraction CM500) was added.

tion assay resulted in a three- to fourfold stimulation. Since the protected sites of both the upstream and the downstream binding regions are quite large (more than 70 nucleotides in each case), it is likely that more than one molecule of p16 binds to each region.

p16 was soluble in 5% perchloric acid (data not shown), and its amino acid composition was similar to the reported amino acid compositions of the HMG-I (12, 32) proteins. The DNA-binding sites of the α-protein (50) and other members of the HMG-I family (9) have been reported to be A+T-rich regions of ≥6 bp. p16 also recognizes A+T-rich sequences; however, it has a higher affinity for sequences of eight or more consecutive thymidines (or adenines), suggesting that if p16 is a member of the highly conserved HMG family or proteins, it may belong to a subclass of the HMG-I family. In this case, we have demonstrated a new, nondenaturing method for purifying this member of HMG-I proteins. In studies of yeast genes, poly(dA-dT) sequences have been shown to activate transcription bidirectionally, and the degree of activation appears to be proportional to the length of the poly(dA-dT) sequences (51). These A+T-rich DNA regions were reported to be free of nucleosome structure (26, 42, 49), and they caused the DNA to kink (59). It was suggested (51) that the activation of transcription by the poly(dA-dT) DNA was not due to a *trans*-acting factor, but to the unusual structure of the sequences themselves. However, in our studies on the rDNA gene, we have not observed any effect of the poly(dA-dT) sequences alone. We found that the AT elements played a role in transcription only when the assay mixtures were supplemented with p16. These results would suggest that the A+T-rich sequences in the rat rDNA might be functionally different from that in the yeast genome in that they would require the binding of a cellular protein which might, then, change the conformation of the DNA and activate transcription. Preliminary DNase I footprinting experiments (results not shown) supported this model. We have observed an alteration in the structure of the surrounding DNA, as reflected by the occurrence of DNase I-hypersensitive sites, when p16 was bound to the DNA.

The identification of an ETS binding site which could mediate the stimulatory effect of p16 was unexpected. Except for the internal promoter elements of the class III genes, very few intragenic transcription activators have been reported (11, 43). One example of such an element is the tissue-specific enhancer found in the intron of the immunoglobulin heavy-chain gene (11). Although we have not tested the ETS p16-binding site for properties such as orientation and distance dependence (properties common to other enhancers), the ETS binding site demonstrated similarities to the immunoglobulin heavy-chain gene enhancer. For example, these two elements are transcribed as part of their genes, recognized by specific cellular factors (48), and can activate their own transcription. We have previously identified an additional RNA polymerase I promoter in the NTS region (5). This NTS transcript initiates at nucleotide −713 with respect to the initiation site of 45S RNA synthesis. The NTS binding site of p16 lies downstream of this NTS promoter (nucleotides −620 to −417), and thus will be transcribed as part of the transcripts initiate at the NTS

promoter. Our data showed that transcription from this NTS promoter can also be stimulated by p16 as long as the NTS binding site is in *cis* with this promoter.

De Winter and Moss (6) reported that at least one spacer promoter is necessary for efficient enhancement of *Xenopus laevis* rDNA transcription and suggested that these spacer promoters and the 60/81-bp enhancer repeats probably act together to enhance 40S pre-RNA synthesis. The NTS promoter that we identified in the rat rDNA gene lies within, or close to, the UAS. It is possible that the NTS promoter is one of the functional parts of the UAS and that it plays a role very similar to that of the spacer promoter in the *Xenopus* ribosomal repeat (6). Since the transcription of the DNA template without the UAS can also be stimulated by p16, as long as the downstream binding domain is *cis*-linked to the promoter, this would suggest that p16 is not the factor (or the only factor) that is responsible for the UAS effect observed previously (4). However, the proximity of these three functional elements, i.e., the p16-binding site, the NTS promoter, and the UAS, may suggest the existence of a presently undefined mechanism that regulates rRNA synthesis in vivo.

The mechanism by which p16 stimulates rRNA synthesis in vitro is unknown. It is possible that the binding of p16 alters the conformation of the DNA and makes the template more accessible to the transcription machinery. While this study was in progress, Tremethick and Molloy (53) reported that HMG 1 and HMG 2 stimulated the transcription of genes transcribed by RNA polymerase II and RNA polymerase III in vitro. Although it was not determined whether this stimulatory effect required specific sequences, as we have shown for p16, their results suggested that it was due to DNA-protein interactions. It has also been reported that HMG 14 and HMG 17 are associated with actively transcribed genes (57, 58). These results might suggest that most, if not all, HMG proteins can activate gene transcription, although their mechanisms of action may not be identical. The effect of p16 on transcription of the class II and III genes is under investigation.

The role of the HMG proteins in the regulation of gene expression has been the source of some speculation (23, 24, 47, 54, 57, 58). We have demonstrated that a possible member of this group of proteins, p16, has a positive effect on transcription. Since the binding of p16 to the transcribed region of rDNA gene appears to be sufficient for the stimulatory effect, this might represent a novel way of regulating gene expression. The synthesis of the HMG-I proteins has been shown to be dependent on the growth rate of the cells (12), as is the synthesis of rRNA. It will be of interest to determine whether the expression or posttranslational modification of p16 is one of the key steps involved in the growth-dependent expression of the rRNA gene family.

ACKNOWLEDGMENTS

We thank S. Chirala, R. Reddy, R. Schwartz, M.-J. Tsai, and L. Yeoman for critical reading of this manuscript. We are especially indebted to L. Yeoman for his help in the amino acid composition analysis and to B. Cassidy for his initial input into this project.

This research was supported by Public Health Service grants GM-37252 from the National Institute of General Medical Sciences and CA-10893, P4 from the National Cancer Institute.

LITERATURE CITED

1. **Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin.** 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. Cell **49**:729–739.
2. **Birkenmeier, E. H., D. D. Brown, and E. Jordan.** 1978. A nuclear extract of *Xenopus laevis* oocytes that accurately transcribes 5S RNA genes. Cell **15**:1077–1086.
3. **Cassidy, B., R. Haglund, and L. I. Rothblum.** 1987. Regions upstream from the core promoter of the rat ribosomal gene are required for the formation of a stable transcription initiation complex by RNA polymerase I *in vitro*. Biochim. Biophys. Acta **909**:133–144.
4. **Cassidy, B. G., H.-F. Yang-Yen, and L. I. Rothblum.** 1986. Transcriptional role for the nontranscribed spacer of rat ribosomal DNA. Mol. Cell. Biol. **6**:2766–2773.
5. **Cassidy, B. G., H.-F. Yang-Yen, and L. I. Rothblum.** 1987. Additional RNA polymerase I initiation site within the nontranscribed region of the rat rRNA gene. Mol. Cell. Biol. **7**:2388–2396.
6. **De Winter, R. F. J., and T. Moss.** 1986. Spacer promoters are essential for efficient enhancement of *X. laevis* ribosomal transcription. Cell **44**:313–318.
7. **Dixit, A., L. C. Garg, W. Chao, and S. T. Jacob.** 1987. An enhancer element in the far upstream spacer region of rat ribosomal RNA gene. J. Biol. Chem. **262**:11616–11622.
8. **Elion, E., and J. Warner.** 1986. An RNA polymerase I enhancer in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6**:2089–2097.
9. **Elton, T. S., M. S. Nissen, and R. Reeves.** 1987. Specific AT DNA sequence binding of RP-HPLC purified HMG-I. Biochem. Biophys. Res. Commun. **143**:260–265.
10. **Fried, A., and D. M. Crothers.** 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
11. **Gillies, S. D., S. L. Morrison, V. T. Oi, and S. Tonegawa.** 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell **33**:717–728.
12. **Goodwin, G. H., P. N. Cockerill, W. Kellam, and C. A. Wright.** 1985. Fractionation by high-performance liquid chromatography of the low-molecular-mass high-mobility-group (HMG) chromosomal proteins present in proliferating rat cells and an investigation of the HMG proteins present in virus transformed cells. Eur. J. Biochem. **149**:47–51.
13. **Grummt, I.** 1981. Specific transcription of mouse ribosomal DNA in a cell free system that mimics control *in vivo*. Proc. Natl. Acad. Sci. USA **78**:737–731.
14. **Grummt, I.** 1982. Nucleotide sequence requirements for specific initiation of transcription by RNA polymerase I. Proc. Natl. Acad. Sci. USA **79**:6908–6911.
15. **Grummt, I., A. Kuhn, I. Bartsch, and H. Rosenbauer.** 1986. A transcription terminator located upstream of the mouse rDNA initiation site affects rRNA synthesis. Cell **47**:901–911.
16. **Grummt, I., H. Rosenbauer, I. Niedermeyer, U. Maier, and A. Öhrlein.** 1986. A repeated 18 bp sequence motif in the mouse rDNA spacer mediates binding of a nuclear factor and transcription termination. Cell **45**:837–846.
17. **Grummt, I., E. Roth, and M. R. Paule.** 1982. Ribosomal RNA transcription *in vitro* is species specific. Nature (London) **296**:173–174.
18. **Hager, D. A., and R. R. Burgess.** 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with Sigma subunit of *Escherichia coli* RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. Anal. Biochem. **109**:76–82.
19. **Haglund, R. E., and L. I. Rothblum.** 1987. Isolation, fractionation and reconstitution of a nuclear extract capable of transcribing ribosomal DNA. Mol. Cell. Biochem. **73**:11–20.
20. **Haltiner, M., S. Smale, and R. Tjian.** 1986. Two distinct promoter elements in the human rRNA gene identified by linker-scanning mutagenesis. Mol. Cell. Biol. **6**:227–235.
21. **Harrington, C. A., and D. M. Chikaraishi.** 1987. Transcription of spacer sequences flanking the 45S rDNA gene of rat. Mol. Cell. Biol. **7**:314–325.
22. **Henderson, S., and B. Sollner-Webb.** 1986. A transcriptional terminator is a novel element of the promoter of the mouse

ribosomal RNA gene. Cell **47**:891–900.
23. **Jackson, J. B., J. M. Pollock, Jr., and L. R. Randolph.** 1979. Chromatin fractionation procedure that yields nucleosomes containing near-stoichiometric amounts of high mobility group nonhistone chromosomal proteins. Biochemistry **18**:3739–3748.
24. **Johns, E. W. (ed.).** 1982. The HMG chromosomal proteins. Academic Press, Inc., New York.
25. **Kohorn, B., and P. Rae.** 1983. Localization of DNA sequences promoting RNA polymerase I activity in Drosophila. Proc. Natl. Acad. Sci. USA **80**:3265–3268.
26. **Kunkel, G. R., and H. G. Martinson.** 1981. Nucleosomes will not form on double-stranded RNA or over poly(dA) · poly(dT) tracts in recombinant DNA. Nucleic Acids Res. **9**:6869–6888.
27. **Labhart, P., and R. H. Reeder.** 1984. Enhancer-like properties of the 60/81 bp elements in the ribosomal gene spacer of *Xenopus laevis*. Cell **37**:285–289.
28. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
29. **Learned, R., S. Smale, M. Haltiner, and R. Tjian.** 1983. Regulation of human ribosomal RNA transcription. Proc. Natl. Acad. Sci. USA **80**:3558–3562.
30. **Long, E., and I. Dawid.** 1980. Repeated genes in eukaryotes. Annu. Rev. Biochem. **49**:727–764.
31. **Lund, T., K. H. Dahl, E. Mørk, J. Holtlund, and S. G. Laland.** 1987. The human chromosomal protein HMG-I contains two identical palindrome amino acid sequences. Biochem. Biophys. Res. Commun. **146**:725–730.
32. **Lund, T., J. Holtlund, M. Fredriksen, and S. G. Laland.** 1983. On the presence of two new high mobility group-like proteins in HeLa S3 cells. FEBS Lett. **152**:163–167.
33. **Mandal, R.** 1984. The organization and transcription of eucaryotic ribosomal RNA genes. Prog. Nucleic Acid Res. Mol. Biol. **31**:115–160.
34. **Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gefter.** 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc. Natl. Acad. Sci. USA **77**:3855–3859.
35. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
36. **McStay, B., and R. H. Reeder.** 1986. A termination site for *Xenopus* RNA polymerase I also acts as an element of an adjacent promoter. Cell **47**:913–920.
37. **Miesfeld, R., and N. Arnheim.** 1984. Species-specific rDNA transcription is due to promoter-specific binding factor. Mol. Cell. Biol. **4**:221–227.
38. **Miller, K., J. Tower, and B. Sollner-Webb.** 1985. A complex control region of the mouse rRNA gene directs accurate initiation by RNA polymerase I. Mol. Cell. Biol. **5**:554–562.
39. **Mishima, Y., L. Financsek, R. Kominami, and M. Muramatsu.** 1982. Fractionation and reconstitution of factors required for accurate transcription of mammalian ribosomal RNA genes: identification of a species-specific factor. Nucleic Acids Res. **10**: 6659–6670.
40. **Moss, T.** 1983. A transcriptional function for the repetitive ribosomal spacer in *Xenopus laevis*. Nature (London) **302**:223–228.
41. **Mroczka, D., B. Cassidy, H. Busch, and L. Rothblum.** 1984. Characterization of rat ribosomal DNA: the highly repetitive sequences that flank the ribosomal transcription unit are homologous and contain RNA polymerase III transcription initiation sites. J. Mol. Biol. **174**:141–162.
42. **Prunell, A.** 1982. Nucleosomal reconstitution on plasmid-inserted poly(dA) poly(dT). EMBO J. **1**:173–179.
43. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream elements. Cell **33**:741–748.
44. **Reeder, R. H., J. G. Roan, and M. Dunaway.** 1983. Spacer regulation of *Xenopus* ribosomal gene transcription: competition in oocytes. Cell **35**:449–456.
45. **Rothblum, L. I., R. Reddy, and B. Cassidy.** 1982. Transcription initiation site of rat ribosomal DNA. Nucleic Acids Res. **10**: 7345–7362.
46. **Rungger, D., M. Crippa, M. F. Trendelenburg, U. Scheer, and W. W. Franke.** 1978. Visualization of rDNA spacer transcription in *Xenopus* oocytes treated with fluorouridine. Exp. Cell Res. **116**:481–486.
47. **Sandeen, G., W. I. Wood, and G. Felsenfeld.** 1980. The interaction of high mobility proteins HMG 14 and 17 with nucleosomes. Nucleic Acids Res. **8**:3757–3778.
48. **Schöler, H. R., and P. Gruss.** 1985. Cell type-specific transcriptional enhancement *in vitro* requires the presence of trans-acting factors. EMBO J. **4**:3005–3013.
49. **Simpson, R. R., and P. Kunzler.** 1979. Chromatin and core particles formed from the inner histones and synthetic polydeoxyribonucleotides of defined sequence. Nucleic Acids Res. **6**:1387–1415.
50. **Solomon, M. J., F. Strauss, and A. Varshavsky.** 1986. A mammalian high mobility group protein recognizes any stretch of six AT base pairs in duplex DNA. Proc. Natl. Acad. Sci. USA **83**: 1276–1280.
51. **Struhl, K.** 1985. Naturally occurring poly (dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. Proc. Natl. Acad. Sci. USA **82**:8419–8423.
52. **Tower, J., V. C. Culotta, and B. Sollner-Webb.** 1986. Factors and nucleotide sequences that direct ribosomal DNA transcription and their relationship to the stable transcription complex. Mol. Cell. Biol. **6**:3451–3462.
53. **Tremethick, D. J., and P. L. Molloy.** 1986. High mobility group proteins 1 and 2 stimulate transcription *in vitro* by RNA polymerases II and III. J. Biol. Chem. **261**:6986–6992.
54. **Vidali, G., L. C. Boffa, and V. G. Allfrey.** 1977. Selective release of chromosomal proteins during limited DNase I digestion of avian erythrocyte chromatin. Cell **12**:409–415.
55. **Wandelt, C., and I. Grummt.** 1983. Formation of stable preinitiation complexes is a prerequisite for ribosomal DNA transcription *in vitro*. Nucleic Acids Res. **11**:3795–3809.
56. **Weil, P. A., J. Segall, B. Harris, S.-Y, Ng, and R. G. Roeder.** 1979. Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. J. Biol. Chem. **254**:6163–6173.
57. **Weisbrod, S., M. Groudine, and H. Weintraub.** 1980. Interaction of HMG 14 and 17 with actively transcribed genes. Cell **19**:289–301.
58. **Weisbrod, S., and H. Weintraub.** 1981. Isolation of actively transcribed nucleosomes using immobilized HMG 14 and 17 and analysis of α-globin chromatin. Cell **23**:391–400.
59. **Wu, N. M., and D. M. Crothers.** 1984. The locus of sequence-directed and protein-induced DNA bending. Nature (London) **308**:509–513.
60. **Yamamoto, O., N. Takakusa, Y. Mishima, R. Kominami, and M. Muramatsu.** 1984. Determination of the promoter region of mouse ribosomal RNA gene by an *in vitro* transcription system. Proc. Natl. Acad. Sci. USA **81**:299–303.
61. **Yang-Yen, H. F., and L. I. Rothblum.** 1986. Partial nucleotide sequence of a 3.4 kb fragment from the rat ribosomal DNA nontranscribed spacer. Nucleic Acids Res. **14**:5557.