The EMBO Journal vol.4 no.13A pp.3539–3547, 1985

# A general upstream binding factor for genes of the yeast translational apparatus

**Janine Huet, Patrick Cottrelle, Marc Cool, Marie-Luce Vignais, Danielle Thiele, Christian Marck, Jean-Marie Buhler, André Sentenac and Pierre Fromageot**

Département de Biologie, Service de Biochimie, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette Cedex, France

Communicated by A.Sentenac

**Fractionation of yeast extracts on heparin-agarose revealed the presence of a DNA footprinting activity that interacted specifically with the 5'-upstream region of *TEF1* and *TEF2* genes coding for the protein synthesis elongation factor EF-1α, and of the ribosomal protein gene *RP51A*. The protected regions encompassed the conserved sequences 'HOMOL1' (AACATC$^{TA}_{CG}$T$^{A}_{G}$CA) or RPG-box (ACCCATACATT$^{T}_{A}$) previously detected 200–400 bp upstream of most of the yeast ribosomal protein genes examined. Two types of protein-DNA complexes were separated by a gel electrophoresis retardation assay. Complex 1, formed on *TEF1*, *TEF2* and *RP51A* 5'-flanking region, was correlated with the protection of a 25-bp sequence. Complex 2, formed on *TEF2* or *RP51A* probes at higher protein concentrations, corresponded to an extended footprint of 35–40 bp. The migration characteristics of the protein-DNA complexes and competition experiments indicated that the same component(s) interacted with the three different promoters. It is suggested that this DNA factor(s) is required for activation and coordinated regulation of the whole family of genes coding for the translational apparatus.**

*Key words:* footprinting factor/consensus sequences/promoter/*TEF1*, *TEF2*, *RP51A* genes/yeast

## Introduction

The eukaryotic promoter for protein-coding genes comprises several control regions located upstream from the mRNA start site (*cis*-acting elements) required for accurate and efficient initiation of transcription. Recognition of these upstream sequences is not done directly by RNA polymerase B(II) but involves specific DNA-binding proteins (*trans*-acting factors). Template-competition experiments, using *in vitro* transcription systems clearly demonstrated the existence of distinct transcription factors forming stable pre-initiation complexes with the 'TATA' box region (Davison *et al.*, 1983) or other promoter elements (Miyamoto *et al.*, 1984; Wildeman *et al.*, 1984). Dynan and Tjian (1983a, 1983b) isolated the first promoter-specific transcription factor (Sp1) from human cells and demonstrated its interaction with an upstream control element of the SV40 promoter. Similarly, Parker and Topol (1984a) partially purified from insect cells another promoter-specific factor (HSTF) which, in footprinting experiments, protected a region immediately upstream from the 'TATA' box of the *Drosophila* heat-shock gene *hsp*70. The same authors (Parker and Topol, 1984b) isolated a general factor which protected the 'TATA' box region and start site of several *Drosophila* genes. Both the HSTF and the 'TATA' box factor could bind to the *hsp*70 gene simultaneously. However, little is known about these proteins and their role in promoter activation and recognition by RNA polymerase B.

The promoter region of several yeast genes studied by *in vitro* mutagenesis contains at least two separate, *cis*-acting elements: a 'TATA' box region whose deletion severely reduces the level of transcription *in vivo*, and upstream activator sites (UAS) similar in many respects to enhancer sites found in mammalian systems (reviewed by Guarente, 1984; Guarente and Hoar, 1984; Struhl, 1983, 1984). Genetic studies have revealed several regulatory proteins, the best example of which is the *GAL4* gene product which regulates positively several genes involved in galactose utilization. This protein is thought to interact with the UAS element of these genes (Guarente *et al.*, 1982; Johnston and Davis, 1984; West *et al.*, 1984; Yocum *et al.*, 1984). Recently, Bram and Kornberg (1985) partially purified a DNA-binding factor(s) (very likely the *GAL4* gene product) which, in footprint experiments, protected two homologous 30-bp regions contained in the $UAS_G$ element of the *GAL1-GAL10* promoter. These results were well correlated with *in vivo* protection experiments (Giniger *et al.*, 1985).

These findings, as well as others concerning the recognition of tRNA gene promoters (Camier *et al.*, 1985) suggest the existence in yeast, as in animal cells, of a class of transcription/regulatory factors which interact with specific parts of promoters in the absence of other transcriptional components. The isolation of these factors, based on their specific DNA-binding properties, would be a first step in the reconstitution of a yeast *in vitro* transcription system for class B(II) genes, which has yet to be developed, and would contribute to the understanding of promoter activation. With this goal in mind, we have undertaken the isolation of promoter-binding proteins using the promoters of *TEF1* and *TEF2* genes which code for the elongation factor EF-1α in *Saccharomyces cerevisiae*. These two genes are functional, actively expressed, code for exactly the same protein, but have highly divergent promoter sequences (Cottrelle *et al.*, 1985a, 1985b; Schirmaier and Philippsen, 1984).

In the present work we have identified, by DNA footprinting, a DNA-binding activity which interacts with specific regions of *TEF1* and *TEF2* promoters. Factor-promoter complexes were also isolated by polyacrylamide gel electrophoresis. The binding sites were related to two conserved sequences (HOMOL1 and RPG boxes) recently detected by computer analysis upstream of most of the ribosomal protein genes analyzed (Teem *et al.*, 1984; Leer *et al.*, 1985). We show by footprint analysis and by competition experiments that the same component(s) interact with the HOMOL1 region of the ribosomal protein gene *RP51A*. We suggest that this DNA binding component could be a general factor for transcription activation of a large family of genes coding for the translational machinery.

## Results

### *Specific protein binding to TEF1 and TEF2 promoters*

Two unlinked genes *TEF1* and *TEF2* code for the elongation factor EF-1α in *S. cerevisiae*. The sequence of both genes and flank-

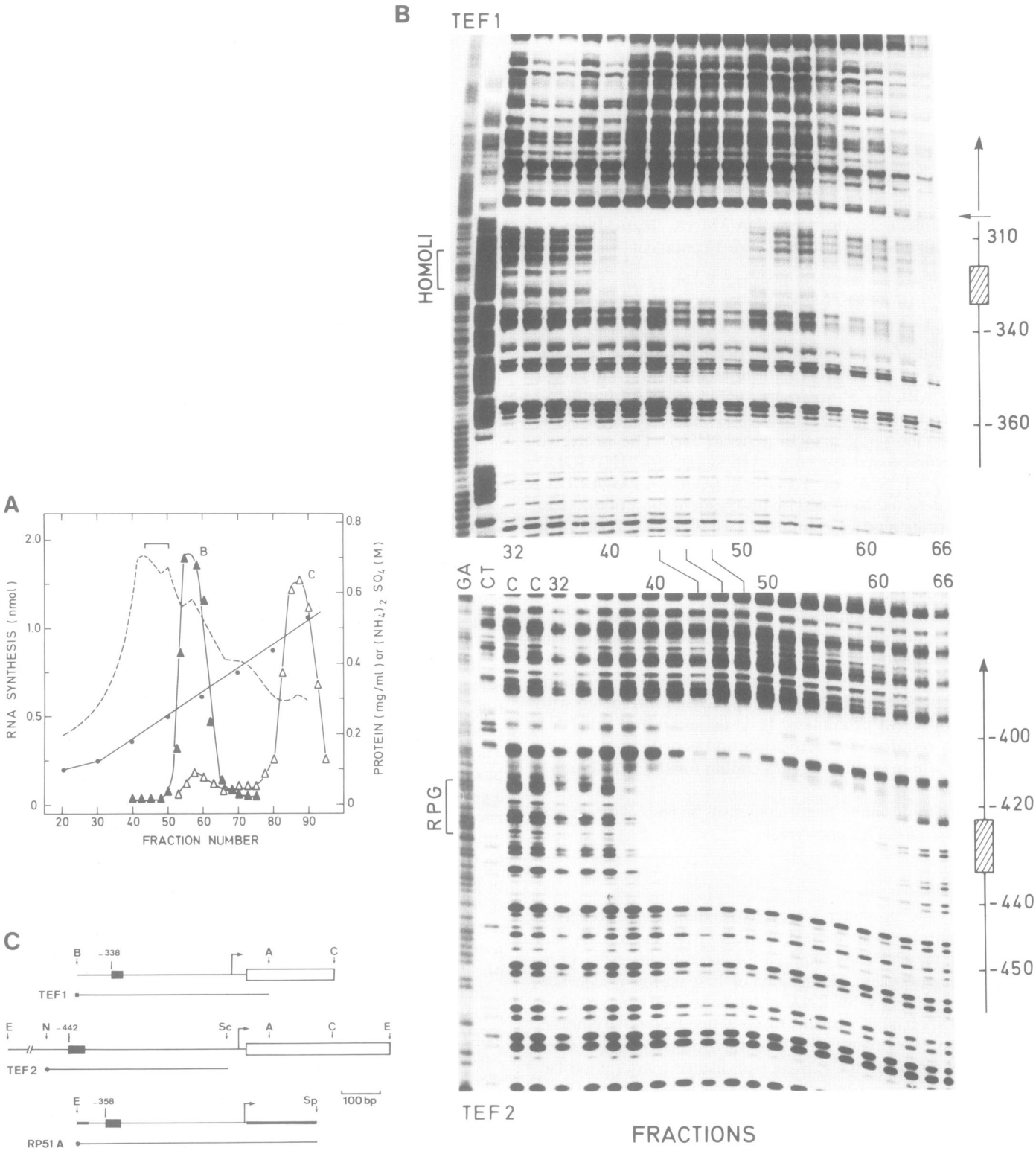


**Fig. 1.** A DNA-binding activity interacts with *TEF1* and *TEF2* 5′-flanking regions. **(A)** Heparin-agarose chromatography. Protein fractions eluted by the salt gradient (●) were assayed for proteins (---), and for RNA polymerase B(▲) or C (△) activity on 10 μl, and every second fraction (14 μl) was used in the footprint assay. **(B)** For footprint analysis the protein fractions from the heparin-agarose column were incubated with the *TEF1* probe 3′ end-labeled at the *Bam*HI site, or with the *TEF2* probe, 3′ end-labeled at the *Nar*I site, the complexes were digested lightly with DNase I, and the DNA fragments were analyzed on an 8% polyacrylamide gel as described in Materials and methods, together with the G+A, and C+T cleavage products of the same probes. Fraction numbers are indicted, as well as control samples of DNA **(lane C)** digested with DNase I in absence of proteins. The location of the RPG or HOMOL1 boxes are indicated as hatched boxes alongside the footprints, with distances to the ATG initiation codon. Sites hypersensitive to DNase I are indicated by the arrows. **(C)** The diagram shows the 5′ upstream region of *TEF1*, *TEF2* and *RP51A* genes and the probes used in the footprint and gel retardation assays. Restriction enzyme sites are abbreviated as follows: A, *Acc*I; B, *Bam*HI; C, *Cla*I; E, *Eco*RI; N, *Nar*I; Sc, *Sca*I; Sp, *Sph*I. Open bars show the coding regions; the bold line indicates pBR322 sequences; arrows show the transcription start sites; lines with the stars are the $^{32}$P-labeled probes. Filled boxes are the DNase I-protected regions; numbering is from the ATG initiation codon.

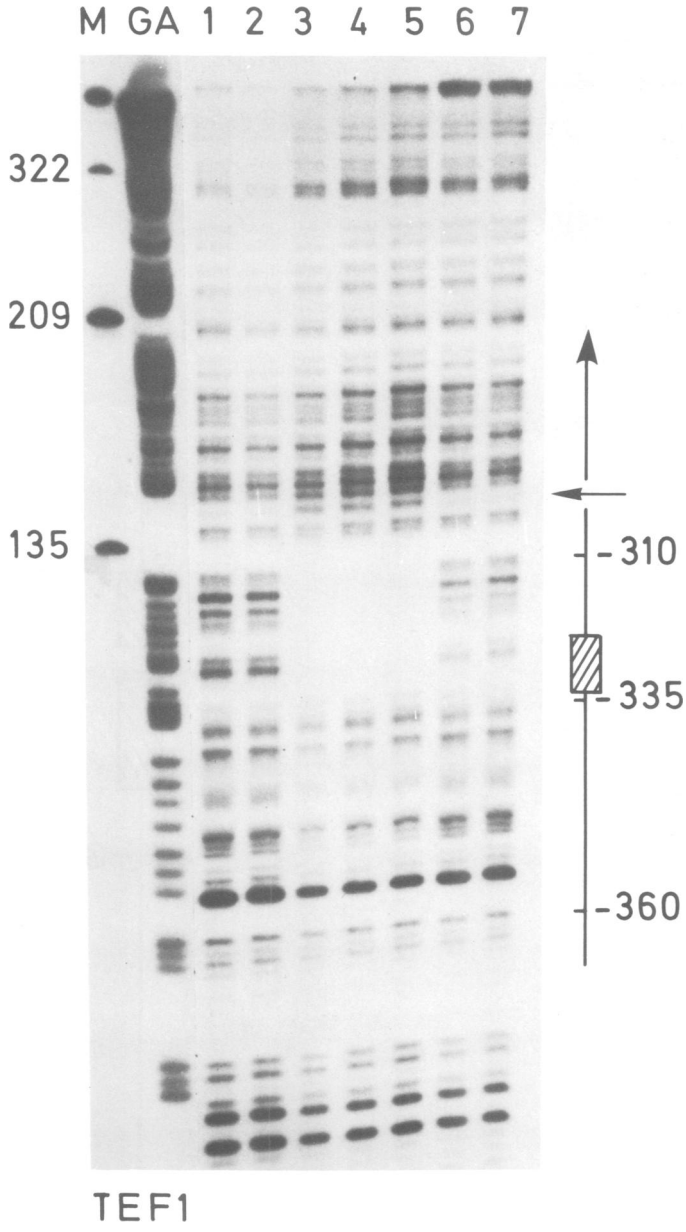


**Fig. 2.** Footprint analysis on the opposite strand of the *TEF1* probe. Protein fractions from the heparin-agarose column shown in Figure 1A were incubated with the *TEF1* probe $^{32}$P end-labeled on the 5′ strand at the *Bam*HI site. DNase I digestion and analysis of the DNA products were as described in Materials and methods. **Lane M** corresponds to size markers; **lane GA** shows the G+A degradation product of the probe; **lanes 1–7** correspond to protein fractions 30, 32, 42, 44, 46, 52 and 54, respectively. Distances to the ATG codon, the location of the HOMOL1 box and an enhanced cleavage site are indicated on the diagram alongside the footprint.

ing regions is known (Cottrelle *et al.*, 1985a; Schirmaier and Philippsen, 1984) as well as their transcription start sites (see Figure 1C). Both genes are functional:gene disruption experiments showed that either gene is required and sufficient for cell viability (Cottrelle *et al.*, 1985b). Dot matrix analysis disclosed no noticeable sequence homology of their 5′-flanking region except for dT clusters often found in promoters of yeast genes (not shown).

The footprint assay was used for detection of promoter binding activities. The DNA fragments used as labeled probes are presented in Figure 1C. As a preliminary affinity chromatography step, a yeast protein extract was fractionated on a heparin-agarose column. Proteins were eluted by a linear salt gradient between 0.1 M and 0.75 M ammonium sulfate, and a series of footprint analyses was performed using every second fraction of the column. A 40-fold excess of carrier DNA was present to trap non-specific DNA-binding proteins which could compete for the labeled probe. As shown in Figure 1B, one region of the *TEF1* probe (3′-labeled at position −425) was protected from DNase I by protein fractions 40–50, eluted at ~0.2 M salt before RNA polymerase B activity. The protected region was ~25 bp long, and was located at positions −338 to −314 from the translation initiation codon. These are minimum values; the exact extent of protection on the 3′ side could not be determined precisely as not all phosphodiester bonds were equally sensitive to DNase I. A strong hypersensitive site was created downstream of the binding region, at position ~−302, within a purine cluster. Another enhanced cleavage site was found upstream at −339. The fractions from the same column were assayed for promoter binding activity using the *TEF2* probe labeled at its 3′ end at position −496. As shown in Figure 1B, the peak fractions that contained maximal *TEF1* binding activity also protected maximally a region of the *TEF2* probe. Curiously, the peak fractions 44–48 protected a larger region of DNA (~35 bp) than side fractions (like 40 or 52) which protected a region of ~25 bp. We have found the same dual footprint pattern by simply lowering the concentration of the two most active fractions 44 and 46: the smaller footprint of ~25 bp was obtained by a 4-fold reduction in protein concentration. This was suggestive of the binding of two components (possibly a protein dimer) on the *TEF2* probe.

A similar footprint experiment was performed with the same column fractions using the *Bam*HI-*Acc*I *TEF1* DNA fragment 5′ end-labeled on the opposite strand at the *Bam*HI site. A footprint was also found on this strand (Figure 2). Comparison with the G+A and C+T sequence ladders showed that the size of the footprint was the same as that seen previously (~25 bp) at approximately the same place, with a slight shift of ~5 bp towards the gene. The same hypersensitive site was found downstream of the protected region at position ~−302.

From these initial results it was not yet clear whether the same component was binding the *TEF1* and *TEF2* probes. Indirect evidence would suggest the presence of two distinct factors: the *TEF2* binding component was present in more fractions than the *TEF1* binding activity, gave two different size footprints and hypersensitive sites were only created in the *TEF1* promoter. However, the region protected contained the same sequence CACCCA--CA at the same distance (~7 bp) from the 5′ border of the complex (see Figure 4).

### *Relationship of DNA binding sites to consensus sequences found upstream of the ribosomal protein genes family*

The sequences protected by the footprint factor(s) showed a remarkable analogy with two dodecanucleotide sequences highly conserved and usually closely spaced, upstream of most of the ribosomal protein genes analyzed so far (Leer *et al.*, 1985). The *TEF1* binding site contained a sequence related to the consensus sequence HOMOL1 (AACATC$^{TA}_{CG}$T$^{A}_{G}$CA), while the *TEF2*-protected site included the RPG-box consensus sequence (ACCCATACATT$^{T}_{A}$) as well as a potential HOMOL1 sequence (see Figure 4 for sequence data). This observation prompted us to investigate whether the footprint component(s) would similarly interact with these 5′ upstream sequences from a ribosomal protein gene. Fractions from heparin-agarose columns showing the footprinting activity on *TEF1* and *TEF2* probes were incubated with a $^{32}$P-labeled

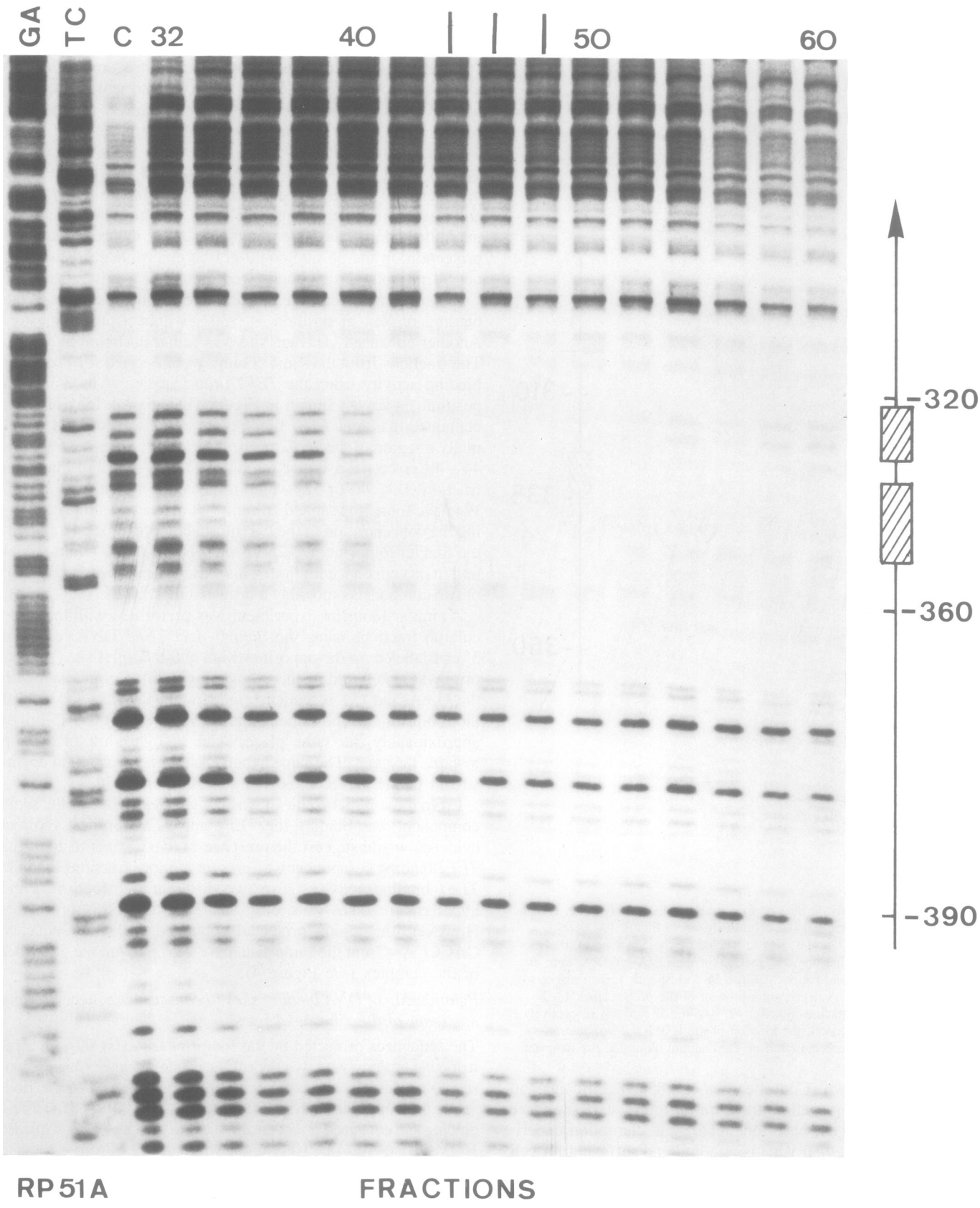


**Fig. 3.** Interaction of the footprint factor with the 5′-flanking region of *RP51A* gene. Protein fractions (14 μl) from the heparin-agarose column shown in Figure 1A were incubated with the *RP51A* probe, the 602-bp *Eco*RI-*Sph*I DNA fragment 3′ end-labeled at the *Eco*RI site. DNase I digestion and analysis of cleavage products were as described in Materials and methods. Every second column fraction was assayed **(lanes 32–60)**; **lane C**, control DNase I digestion; GA and CT, cleavage products of sequencing reactions of the same probe; the bars show the peak of footprinting activity; the location of the two HOMOL1 repeats is indicated in the diagram with positions by reference to the ATG initiation codon.

DNA fragment containing the promoter region of the ribosomal gene *RP51A* which included the canonical HOMOL1 box and a slight variant of that sequence in close proximity. DNA binding activity was revealed by footprint analysis as shown in Figure 3. Indeed, one region of the *RP51A* promoter, ~35–40 bp long, was clearly protected from DNase I digestion. The footprint encompassed the two HOMOL1 boxes (Figure 3 and the sequence data in Figure 4). Footprinting activity on the *RP51A* HOMOL1 region was found to co-elute with the *TEF1* and *TEF2* binding factor from the same heparin-agarose column (compare with Figure 1B). No strong hypersensitive site was found at the boundaries of the footprint as in *TEF2* but in contrast to *TEF1*.

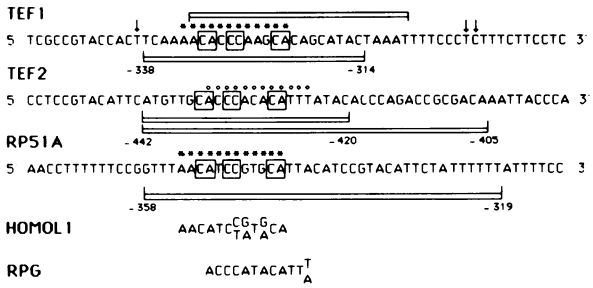


**Fig. 4.** Summary and comparison of protected DNA sequences. The DNA sequence of the relevant 5′ upstream region of *TEF1*, *TEF2* and *RP51A* gene are shown (5′ strand). The sequence upstream of the *TEF2* gene (Schirmaier and Philippsen, 1984) has been extended to the *Nar*I site (position −496) by M. Cool. Nucleotides are numbered from the ATG initiation codon. Protected regions (3′ strand) are underlined. Protection of the 5′ strand of the *TEF1* probe is indicated by the overline and enhanced cleavage sites by arrows. The HOMOL1 and RPG consensus sequences (Teem *et al.*, 1984; Leer *et al.*, 1985) are indicated. Nucleotides residues conserved at a fixed distance from the 5′ boundary of the protected region are boxed.

Figure 4 gives the maps of the factor(s) binding sites on *TEF1*, *TEF2* and *RP51A*, as derived from the data of Figures 1, 2 and 3 and a number of similar experiments. The size of the protected region on *TEF1* (~25 bp) was the same as the small footprint on the *TEF2* probe. The large footprints on *TEF2* or *RP51A* also had a comparable size (35–40 bp). (The full extent of protection is not known precisely on the gene proximal side because of irregular cleavage by DNase I.) The pattern of protection on *TEF1* was asymmetric, with regard to the two DNA strands, by ~5 bp. The consensus sequence HOMOL1 appeared at the same distance (5 bp) from the gene-distal 5′ boundary of the footprint on *RP51A* and *TEF1* genes. The RPG-box ACCCA(C)ACATTT was found in the *TEF2* footprint at 7 bp from its 5′ boundary. A potential HOMOL1 sequence (with two mismatches) started 13 bp from the 5′ border and overlapped this RPG-box.

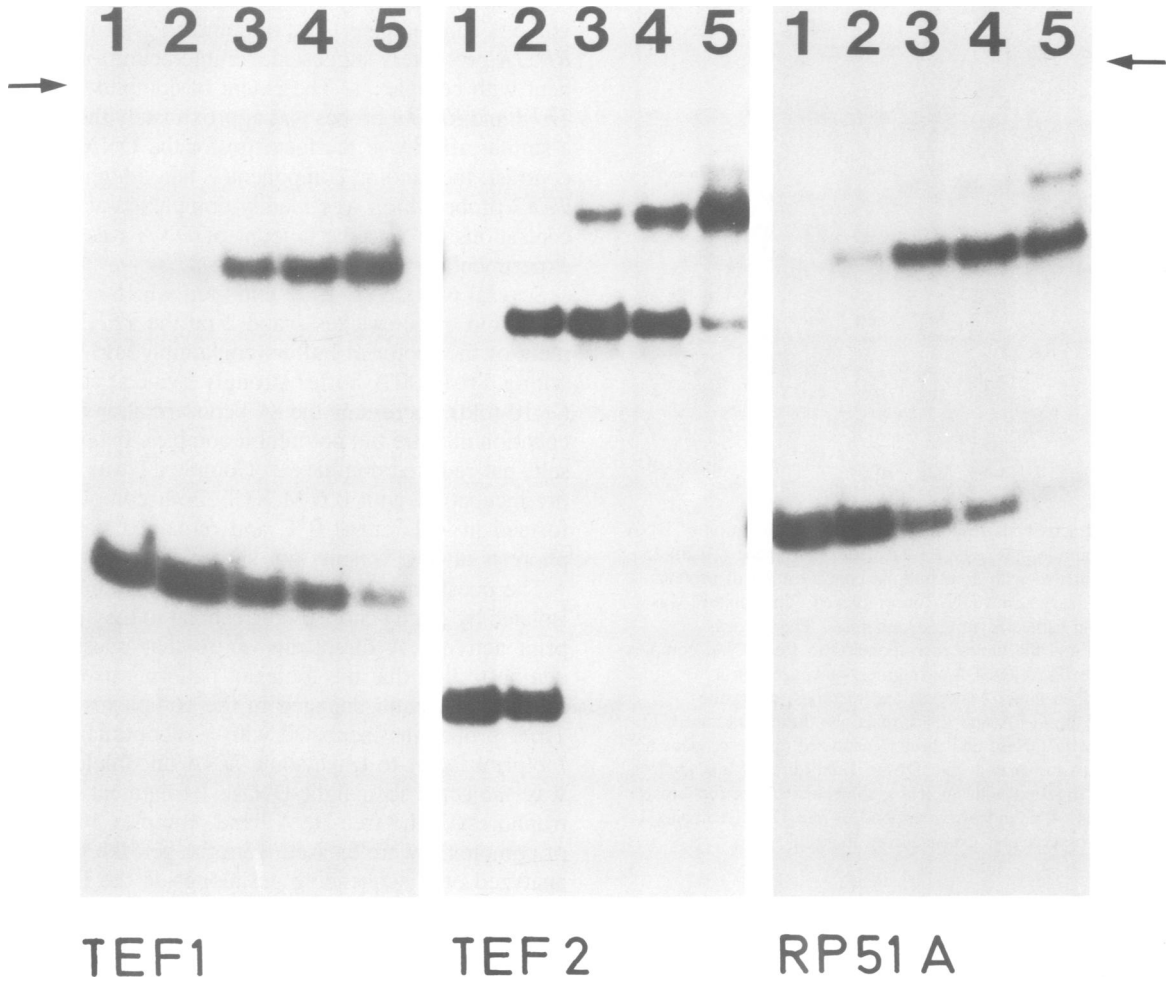


**Fig. 5.** Analysis of protein-DNA complexes by gel electrophoresis. Protein-DNA complexes were formed with three different labeled probes, *TEF1* (486 bp; 5.4 fmol), *TEF2* (452 bp; 5.8 fmol) or *RP51A* probe (602 bp; 6 fmol) in the presence of 0.54 μg of pBR322 DNA and varying amounts of proteins (fraction 45 from Figure 1A). After incubation, the samples were analyzed by electrophoresis as described under Materials and methods. The figure shows the autoradiogram of the slab gel. In **lanes 1–5** the amount of proteins was 0, 0.07, 0.175, 0.35 and 0.7 μg, respectively. The arrows indicate the origin.

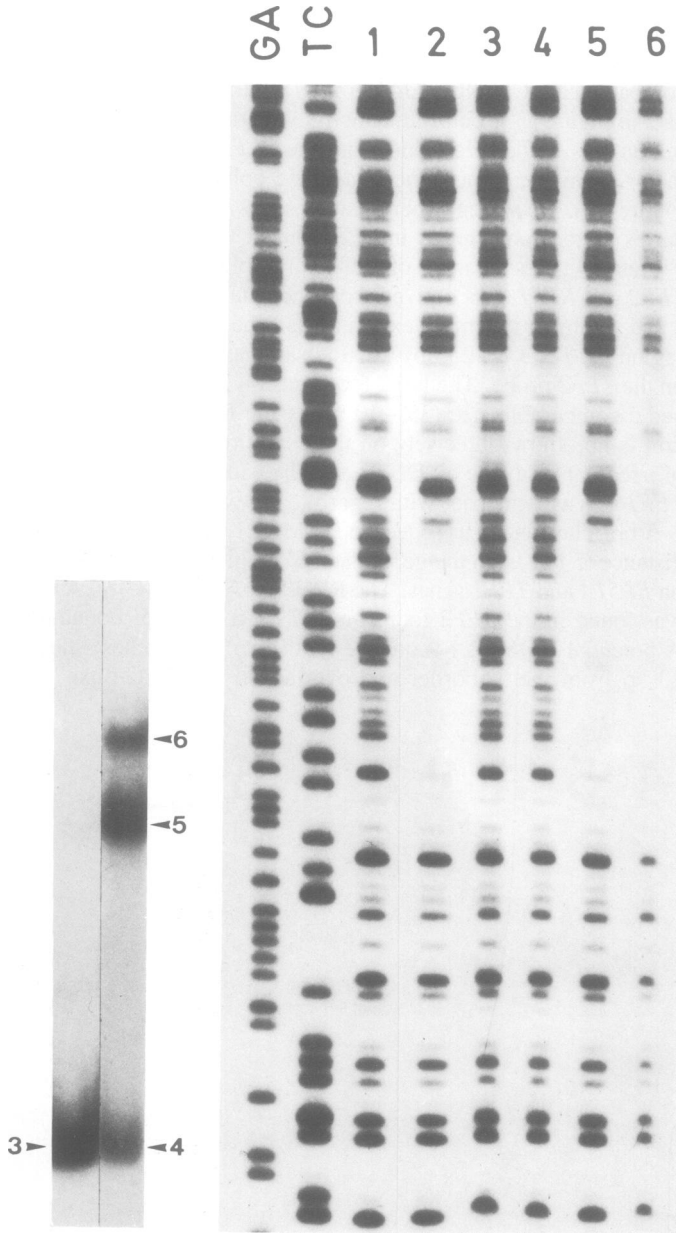


**Fig. 6.** The footprinting activity is responsible for altered migration of DNA complexes on polyacrylamide gels. The $^{32}P$ labeled *TEF2* probe (10 fmol; 24 000 c.p.m.) was incubated with a suboptimal concentration of proteins (0.7 μg; fraction 45) to leave some DNA uncomplexed. The mixture was treated with DNase I and subjected to electrophoresis. The gel was autoradiographed (inset) and the bands corresponding to free DNA, complex 1 and complex 2 were excised and DNA fragments analyzed on a sequencing gel as described under Materials and methods, alongside degradation products of the G+A and C+T reactions (**lanes GA** and **CT**). **1**, DNA probe treated with DNase and directly analyzed on the sequencing gel; **2**, Protein-DNA mixture treated with DNase I and directly analyzed on the gel; **3**, DNA probe digested with nuclease, subjected to electrophoresis and excised from the gel; **4, 5** and **6**, correspond to free DNA, complex 1 and complex 2 bands, respectively, excised from the gel.

*Analysis of DNA-protein complexes by gel electrophoresis*

DNA-binding proteins were found to alter markedly the migration rate of specific, small DNA fragments upon electrophoresis in polyacrylamide gels under non-denaturing conditions. This gel retardation assay was initially described by Fried and Crothers (1981) and by Garner and Revzin (1981). The protein-DNA complexes can be vizualized in partially purified fractions provided enough carrier DNA is added to prevent non-specific binding of unrelated proteins (Strauss and Varshavsky, 1984). We used this technique to isolate and characterize the complexes formed on the different promoter fragments. Should the same factor interact with *TEF1*, *TEF2* and *RP51A* promoters, the resulting effect on the electrophoretic migration rate of the DNA fragments should be the same (i.e., the complex should have the same $R_f$). In the experiment shown in Figure 5, the 486-bp *TEF1* probe was titrated with increasing amounts of the footprinting fraction and the mixture subjected to electrophoresis. The appearance of a sharp, slowly migrating band attested to the formation of a specific protein-DNA complex that had remained stable during electrophoresis. The complex migrated with about one-third the velocity of the free DNA band (the measured $R_f$ was 0.37). When the same titration was performed on the same slab gel with the *TEF2* probe (452 bp), at low protein concentration a complex was formed with the same relative migration ($R_f$ = 0.38); at increasing protein concentrations a second complex appeared which became predominant ($R_f$ = 0.2). The concomitant decrease in complex 1 suggested that it was an intermediate in the formation of complex 2. Similarly, the titration of the 602-bp *RP51A* probe in the same gel resulted in the formation of two complexes, first complex 1, with an $R_f$ = 0.4, then complex 2 with an $R_f$ = 0.22. Since protein binding in complex 1 induced the same alteration in the migration velocity of the three DNA fragments it was likely that the same binding component was involved. The characteristics of complex 2 formation on *TEF2* or *RP51A* promoters suggested the interaction of another component with complex 1. The extent of complex 1 formation with *TEF1* and *RP51A* probes was approximately the same suggesting a similar affinity of the factor(s) for the DNA binding sites. In contrast, the binding component(s) had a higher affinity for the *TEF2* probe which was readily complexed at low protein concentrations. (The same amount of DNA was used in all three experiments.)

Several parameters were explored which could influence the formation of complexes 1 and 2 on the *TEF2* probe. Replacement of the footprint buffer (containing $Mg^{2+}$ and $Ca^{2+}$ ions) with a Tris-EDTA buffer strongly reduced complex formation (~10-fold). Increasing the KCl concentration of the standard incubation mixture did not inhibit complex 1, up to at least 0.3 M salt, but reduced complex 2. Complex 1 was still present after pre-incubation with 0.6 M KCl. Both complex 1 and 2 were formed in <1 min at 0°C and remained stable upon electrophoresis at 4°C (results not shown).

The question remained whether the protein-DNA complexes isolated by electrophoresis correspond to the binding of the footprint activity. A direct answer to that question required the demonstration that the footprint pattern was exclusively found in the DNA band engaged in the complexes. The $^{32}P$-labeled *TEF2* probe was incubated with a suboptimal concentration of footprint factor to leave some DNA uncomplexed, the mixture was subjected to a light DNase I treatment then to gel electrophoresis. The free DNA band, complex 1 and a small band of complex 2 were excised from the gel, the DNA isolated and analyzed on a sequencing gel alongside the DNase I digestion products of the probe alone, similarly isolated from the gel, or of the protein-DNA mixture prior to electrophoresis. As seen in Figure 6 no footprint was found on the free DNA band (compare lane 4 with control lanes 3 and 1). In contrast, the footprint pattern was recovered in the band corresponding to complex 1 (lane 5). This was the small size footprint (23 bp) found at low protein concentration. The band of complex 2 presented the large-

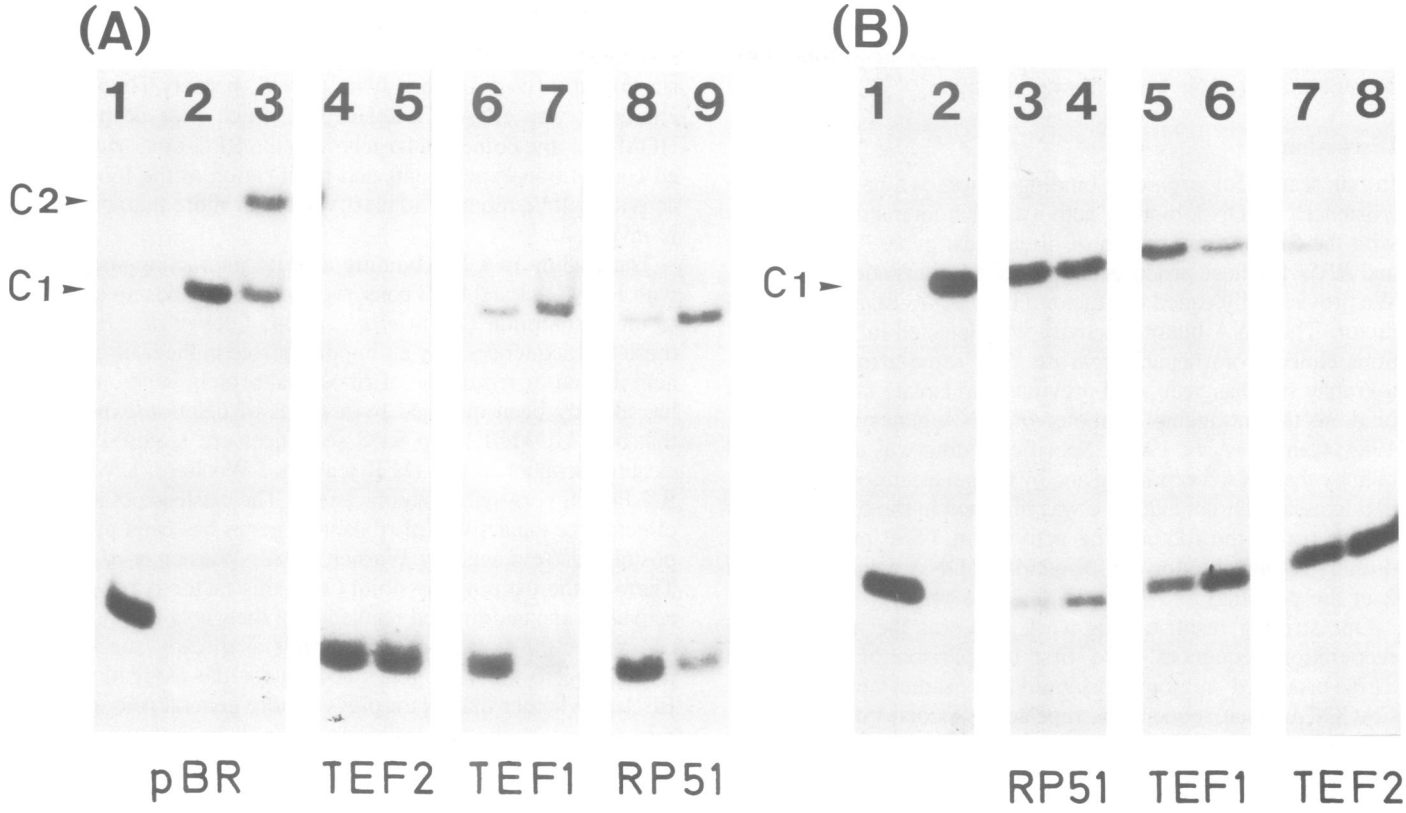


**Fig. 7.** Competition experiments using the gel retardation assay. **(A)** The footprinting protein fraction (0.2 or 0.7 μg of fraction 45) was pre-incubated for 10 min at 25°C with 0.8 μg of different plasmid DNA:pBR322 (0.27 pmol), pYI51A (RP51) (0.14 pmol), pLB25 (*TEF2*) (0.15 pmol), pLB15-1 (*TEF1*) (0.25 pmol) as indicated. The $^{32}$P-labeled *TEF2* probe (5.6 fmol; 6500 c.p.m.) was then added, further incubated for 15 min and the protein-DNA complexes analyzed by electrophoresis (see Materials and methods). Amount of proteins added in **lane 1:** 0 μg; in **lanes 2, 4, 6** and **8:** 0.2 μg; in **lanes 3, 5, 7** and **9:** 0.7 μg. The migration of complex 1 (C1) and complex 2 (C2) is indicated. **(B)** The protein fraction (0.5 μg of fraction 45) was incubated for 15 min with the labeled *RP51A* probe (6.8 fmol; 4800 c.p.m.) in the presence of 0 fmol **(lane 2)** 30 fmol **(lanes 3, 5, 7)** or 100 fmol **(lanes 4, 6, 8)** of a competitor plasmid DNA containing the 5′ upstream region of *RP51A*, *TEF1*, or *TEF2* genes (see text). The amount of vector DNA pBR322 was kept constant at 0.5 μg. The migration of complex 1 is indicated (C1).

size footprint (lane 6). The footprint pattern obtained in protein-DNA mixtures prior to electrophoresis was a combination of these two types of footprint (lane 2). This experiment correlated well the footprint data with the gel electrophoresis assay.

*Competition experiments*

Competition experiments were performed to determine whether the same component(s) interact with the three promoters. Complex formation was followed by gel electrophoresis. Figure 7A shows an experiment where the footprinting activity was pre-incubated with an excess relative to the probe (a 400-fold mass excess or 30- to 50-fold on a molar basis), of different plasmids DNAs, pBR322, pYI51A (containing the *RP51A* gene), pLB15-1 or pLB25 (containing the 5′ upstream region of *TEF1* or *TEF2*). Then the $^{32}$P-labeled *TEF2* probe was added, the protein-DNA mixtures were further incubated and analyzed by electrophoresis. Pre-incubation with excess pBR322 DNA did not prevent the formation of complex 1 or the mixture of complex 1 plus 2 (Figure 7A, lanes 2 and 3). As expected, the DNA binding activity was totally preempted by pre-incubation with the *TEF2* plasmid (lanes 4, 5). Pre-incubation with *TEF1* or *RP51A* DNA also efficiently competed out the formation of complex 1 on the *TEF2* probe as evidenced by the restoration of the free DNA band (lanes 6 and 8). Strikingly, the formation of complex 2 was also totally prevented by both DNAs (lanes 7 and 9) although the *TEF1* promoter did not form a type 2 complex (see Figure 5). The presence of a residual amount of complex 1 further indicated that *TEF1* and *RP51A* DNA were less efficient than *TEF2* in binding the footprint factor(s). These results were confirmed by a similar pre-binding experiment, this time using a *TEF1*-labeled probe and analyzing the *TEF1*-protein complexes by footprinting. Again pre-incubation with plasmid DNAs containing the *TEF1*, *TEF2* or *RP51A* promoters totally prevented the subsequent binding of the factor to the *TEF1* probe. Plasmid pBR322 had no effect on the footprint (results not shown). These experiments indicate that the proteins responsible for complex 1 and complex 2 formation interact with the three competitor DNAs.

The direct competition experiment shown in Figure 7B further supported the conclusion that the same component(s) interacted with the promoters of *TEF1*, *TEF2* and *RP51A* genes. The footprint fraction was equilibrated for 15 min at 25°C with the *RP51A* $^{32}$P-labeled probe in the presence of increasing amounts of *TEF2* (the *Nar*I-*Sca*I fragment of 452 bp), *TEF1* (the *Bam*HI-*Acc*I fragment of 486 bp) or *RP51A* (the 392-bp *Asu*II-*Bgl*II fragment) promoter regions, each cloned into pBR322. The amount of vector DNA was kept constant by supplementing with pBR322 plasmid DNA. Complex formation on *RP51A* was then analyzed by the electrophoresis assay (Figure 7B). The amount of protein added was sufficient to convert most of the DNA probe into complex 1. The best competitor DNA was *TEF2* promoter: a 5-fold molar excess over the *RP51A* probe inhibited complex 1 formation 95%, while a 15-fold molar excess of *TEF1* or

*RP51A* unlabeled competitor DNA was required to inhibit complex 1 formation 90% and 30%, respectively. These results indicate that the same component(s) interact with the three promoters, although with varying affinity.

## Discussion

In our search for promoter-binding factors we have found the existence of a DNA binding activity which interacts specifically with the 5′ upstream region of three yeast genes, *TEF1*, *TEF2* and *RP51A*, whose products are part of the translation apparatus. We provisionally called this factor TUF, for translation upstream factor. The DNA binding activity was detected in protein fractions eluted from heparin-agarose. The same affinity chromatography step has been used previously to isolate factor $\tau$ which binds to the intragenic promoter of tRNA genes (Ruet *et al.*, 1984; Camier *et al.*, 1985). Specific binding was demonstrated both by the DNA footprinting and by the gel electrophoresis retardation assay. Direct evidence was obtained in the experiment of Figure 6 that the footprinting activity on *TEF2* promoter was directly responsible for the formation of DNA complexes 1 and 2 in the gel assay.

One striking result of this work concerns the nature of the recognition sequences. The first comparison of *TEF1*- and *TEF2*-protected regions disclosed the same hexanucleotide CACCCA. That sequence is repeated a second time in the extended footprint region of *TEF2* and was therefore thought to be part of the recognition site. A more general comparison further revealed that the *TEF1*-protected region contained a sequence related (with two mismatches) to the 12-bp consensus sequence 'HOMOL1' discovered recently by Teem *et al.* (1984) some 300 bp upstream from several ribosomal protein genes. Leer *et al.* (1985) extended that finding to 16 out of the 21 ribosomal protein genes examined. This led us to investigate the binding of the footprint factor on *RP51A* promoter. That gene contains the canonical HOMOL1 sequence and, within 2 bp, a direct repeat of that sequence with a slight variation. Indeed the footprint on *RP51A* exactly encompassed these two sequences (see Figure 4). At this point, the simplest interpretation was that the factor interacted with HOMOL1 sequence to give the 25-bp footprint and complex 1. Complex 2 and the large footprint would be formed on *RP51A* by binding a second molecule of factor to the HOMOL1 repeat. It may be noted that on *RP51A* the small-size footprint may not be found experimentally if the two binding sites are independent and have the same affinity for the factor. This remains to be investigated.

A complication arises from the analysis of the *TEF2*-protected sequence. Of the three genes, the *TEF2* promoter shows the highest affinity for the factor, although no convincing HOMOL1 sequence could be found in the 23-bp footprint (complex 1). Instead, a distinct sequence was found, clearly related (with only one mismatch) to an additional 12-bp conserved sequence, the 'RPG-box', recently discovered by Leer *et al.* (1985) in most ribosomal protein genes, often in close proximity and 3′ to the HOMOL1-sequence (Figure 4). This finding raises the question of the presence of two distinct factors in the footprinting fraction. However, several observations seem to argue against this possibility. The similar 23-bp footprint on *TEF1* and *TEF2*, the similar alteration in migration velocity of type I complexes on the three promoters and the co-elution of the peak of footprinting activity from heparin-agarose. Also there is a considerable similarity in the two consensus sequences (see Figure 4). The competition experiments indicate that the same DNA binding activity interacts with the 5′ upstream regions of *TEF1*, *TEF2* and *RP51A* genes but do not exclude the existence of multiple components. Assuming that the same factor interacted with the three promoters we looked for a homology in their protected regions starting at a fixed distance from their 5′ border. The sequence emerging was RCAYCCRHRCAY which is a derivative of HOMOL1 and coincides largely with the RPG-box. More refined competition experiments and purification of the footprinting activity will be required to clarify whether more than one factor is involved.

The finding of DNA binding activity interacting specifically with HOMOL1 and RPG conserved sequences adds much weight to the assumption (Teem *et al.*, 1984; Leer *et al.*, 1985) that these two sequences play an important role in the transcriptional activity and/or regulation of ribosomal protein genes. Evidence has already been obtained by a series of deletion experiments that both HOMOL1 and RPG sequences are required for efficient transcription *in vivo* (M.Rotenberg, J.Woolford, L.Woudt and R.J.Planta, personal communication). The existence of a positive effector for the activity of ribosomal genes has been previously postulated (reviewed by Warner, 1982; Warner *et al.*, 1985). There is the exciting possibility that this factor is required for activation and coordinated regulation of the whole family of genes coding for the translational apparatus as already suggested by its binding to the *TEF* genes coding for the elongation factor EF-1$\alpha$. Whether that factor plays a more general role in growth rate control is also a possibility that will be investigated.

## Materials and methods

*Plasmid and DNA probes*

All plasmids were cloned in *Escherichia coli* RRI strain. Plasmid pLB15 is a derivative of pLB1 that carries the *TEF1* gene (Cottrelle *et al.*, 1985a). It contains 425 bp of the upstream sequence and part of the *TEF1* coding region in a *Sau*3A-*Cla*I fragment inserted in the *Bam*HI/*Cla*I site of pBR322. The *TEF1* probe was the 486-bp *Bam*HI-*Acc*I fragment labeled at the *Bam*HI site with [$\alpha$-$^{32}$P]dGTP by DNA polymerase I filling (Klenow fragment) or by posphorylation with [$\gamma$-$^{32}$P]ATP and T4 polynucleotide kinase. Plasmid pLB25 is a derivative of pLB2 which carries the complete *TEF2* gene (M. Cool and J.M. Buhler, unpublished). Plasmid pLB25 contains two copies of a 1.9-kb *Eco*RI-*Eco*RI fragment encompassing 363 bp of the *TEF2* coding sequence and 1.6 kb of the 5′ upstream region. The *TEF2* probe was a 452-bp fragment extending from the *Nar*I site (position −496) to the *Sca*I site, $^{32}$P-labeled by DNA polymerase I filling at the *Nar*I site. The same fragment was subcloned into the *Cla*I/*Bam*HI sites of pBR322 to give plasmid pLB25-1. Plasmid pYI51A, a gift from M. Rosbash, carries the gene coding for the ribosomal protein rp51A on a 3.2-kb *Eco*RI-*Hind*III yeast DNA fragment (Abovich and Rosbash, 1984). A 392-bp *Asu*II-*Bgl*II fragment extending from position −416 to −24 was subcloned in the *Cla*I/*Bam*HI sites of pBR322, resulting in plasmid p51A-P. The *RP51A* probe is the 602-bp *Eco*RI-*Sp*HI fragment prepared from p51A-P and 3′ end-labeled at the *Eco*RI site. Bacterial transformation, purification of plasmids, restriction enzyme digestions, DNA modification, gel electrophoresis of DNA and purification of the labeled DNA probes were performed by standard methods (Maniatis *et al.*, 1982).

*Isolation of the footprinting activity*

Yeast cells *S. cerevisiae* 20B-12 (pep 4-3) collected in the exponential phase of growth (150 g) were broken in a Manton Gaulin homogenizer in a high salt buffer, centrifuged at high speed and the supernatant applied to a heparin-agarose column, as previously described (Ruet *et al.*, 1984). Protein fractions (10 ml) eluted by a salt gradient from 0.1 to 0.75 M ammonium sulfate were assayed for RNA polymerase B and C activity, for proteins, dialyzed against a buffer containing 20 mM Tris-HCl pH 8, 50 mM KCl, 1 mM dithiothreitol (DTT), 0.2 mM EDTA, 10% glycerol (v/v) and stored at −70°C in small aliquots.

*DNase I protection experiment*

The footprint assay mixture (20 $\mu$l final volume) contained 20 mM Tris-HCl, pH 8, 70 mM KCl, 5 mM $MgCl_2$, 0.5 mM $CaCl_2$, 0.5 mM DTT, 0.1 mM EDTA, 7% glycerol (v/v), 120 ng carrier pBR322 DNA and 2−5 ng (10 000−20 000 c.p.m.) of $^{32}$P end-labeled DNA fragment. The binding reaction was initiated by addition of 14 $\mu$l of each protein fraction (~7 $\mu$g protein). After 5 min incubation at 25°C, freshly diluted DNase I (2.4 ng) was added, and DNA digestion allowed to proceed for 40 s at 25°C. The reaction was stopped by addition of 80 $\mu$l of a stop solution containing 12.5 mM EDTA, 12.5

μg/ml proteinase K, 125 μg/ml yeast tRNA, 0.1% SDS (v/v), and the mixture incubated 15 min at 37°C. The DNA fragments were extracted once with a phenol:chloroform:isoamylalcohol mixture (24:24:1), precipitated with ethanol and 0.3 M NaOAc, resuspended in 5 μl of 90% formamide (v/v), and denatured for 3 min at 90°C prior to electrophoresis on a 0.35 mm thick 8% polyacrylamide sequencing gel. The gels were autoradiographed wet on Kodak-X OMAT S films at −70°C with Cronex HI plus intensifying screens. The G+A and C+T chemical degradation reactions of the DNA probes and the electrophoresis were done as described by Maxam and Gilbert (1980).

*Analysis of protein-DNA complexes by gel electrophoresis*

DNA-protein complexes were formed in 20 μl of footprint buffer for 15 min at 25°C with 5 – 10 fmol of $^{32}$P-labeled DNA fragments (5000 – 15 000 c.p.m.), carrier pBR322 DNA and proteins as indicated in figure legends. After incubation, the samples were directly loaded on a 13 × 11 cm polyacrylamide vertical slab gel (4.94% acrylamide, 0.06% bisacrylamide) made with 20 mM Tris-HCl pH8, 1 mM EDTA, 5% glycerol (v/v) as gel buffer. Reservoir buffer was 20 mM Tris-HCl pH 8, 1 mM EDTA and 5 mM 2-mercaptoethanol. After polymerization a pre-electrophoresis of 1 h was run at constant voltage (200 V, 35 mA), the reservoir buffer was changed and the gel equilibrated overnight. A second pre-electrophoresis of 1 h and change of buffer was repeated before loading the samples. Electrophoresis was run for 3 – 4 h in the cold room, at constant voltage (200 V), with one buffer change after 1.5 h. Bromophenol blue and xylene cyanol were added in a side well to monitor migration. Electrophoresis was stopped when xylene cyanol dye was 2 cm from the bottom of the gel. The dye migrated like DNA fragments of ~470 bp. The gels were autoradiographed wet as above.

To reveal the footprinting activity at the level of the complexes on the gel, the DNase I treatment was performed on the protein-DNA complexes and stopped with EDTA prior to electrophoresis, the complexes and free DNA bands were located by autoradiography, and eluted from the gel by diffusion overnight at 37°C with three changes of a 0.4 ml solution containing 10 mM Tris-HCl pH 7.8, 1 mM EDTA, 0.2% SDS (w/v), 0.3 M NaCl, and 1 μg/ml yeast tRNA. The solutions were filtrated through a membrane (Schleicher and Schüll, FP 030/2) the DNA fragments precipitated with ethanol and subjected to electrophoresis as in the DNase I protection experiment. When possible, approximately the same amount of radioactivity was loaded on the gel (~8000 c.p.m.).

## Acknowledgements

We are grateful to Michael Rosbash (Brandeis University) who generously provided the plasmid containing *RP51A* gene, to Sylvie Camier and Nejib Marzouki for the gift of protein fractions, to Bernard Lescure, Sylvie Camier, and Carl Mann for many discussions. The work was funded in part by a grant from the CNRS (ATP 033-542).

## References

Abovich,N. and Rosbash,M. (1984) *Mol. Cell. Biol.*, **4**, 1871-1879.
Bram,R.J. and Kornberg,R.D. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 43-47.
Camier,S., Gabrielsen,O., Baker,R. and Sentenac,A. (1985) *EMBO J.*, **4**, 491-500.
Cottrelle,P., Thiele,D., Price,V., Memet,S., Micouin,J.-Y., Marck,C., Buhler,J.-M., Sentenac,A. and Fromageot,P. (1985a) *J. Biol. Chem.*, **260**, 3090-3096.
Cottrelle,P., Cool,M., Thuriaux,P., Price,V., Thiele,D., Buhler,J.-M. and Fromageot,P. (1985b) *Curr. Genet.*, **9**, 693-697.
Davison,B.L., Egly,J.-M., Mulvihill,E.R. and Chambon,P. (1983) *Nature*, **301**, 680-686.
Dynan,W.S. and Tjian,R. (1983a) *Cell*, **32**, 669-680.
Dynan,W.S. and Tjian,R. (1983b) *Cell*, **35**, 79-87.
Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505-6525.
Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047-3060.
Giniger,E., Varnum,S.M., Ptashne,M. (1985) *Cell*, **40**, 767-774.
Guarente,L. (1984) *Cell*, **36**, 799-800.
Guarente,L. and Hoar,E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7860-7864.
Guarente,L., Yocum,R.R. and Gifford,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7410-7414.
Johnston,M. and Davis,R.W. (1984) *Mol. Cell. Biol.*, **4**, 1440-1448.
Leer,R.J., Van Raamsdonk-Duin,M.C., Mager,W.H. and Planta,R.J. (1985) *Curr. Genet.*, **9**, 273-277.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
Miyamoto,N.G., Moncollin,V., Wintzerith,M., Hen,R., Egly,J.-M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 8779-8799.
Parker,C.S. and Topol,J. (1984a) *Cell*, **37**, 273-283.
Parker,C.S. and Topol,J. (1984b) *Cell*, **36**, 357-369.
Ruet,A., Camier,S., Smagowicz,W., Sentenac,A. and Fromageot,P. (1984) *EMBO J.*, **3**, 343-350.
Schirmaier,F. and Philippsen,P. (1984) *EMBO J.*,**3**, 3311-3315.
Strauss,F. and Varshavsky,A. (1984) *Cell*, **37**, 889-901.
Strühl,K. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7385-7389.
Strühl,K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7865-7869.
Teem,J.L., Abovich,N., Kaufer,N.F., Schwindinger,W.F., Warner,J.R., Levy,A., Woolford,J., Leer,R.J., van Raamsdonk-Duin,M.C., Mager,W.H., Planta,R.J., Schultz,L., Friesen,J.D., Fried,H. and Rosbash,M. (1984) *Nucleic Acids Res.*, **12**, 8295-8312.
Warner,J.R. (1982) in Strathern,J., Jones,E. and Broach,J. (eds.), *Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression*, Cold Spring Harbor Laboratory Press, NY.
Warner,J.R., Mitra,G., Schwindinger,W.F., Studeny,M. and Fried,H.M. (1985) *Mol. Cell. Biol.*, **5**, 1512-1521.
West,R.W., Yocum,R.R. and Ptashne,M. (1984) *Mol. Cell. Biol.*, **4**, 2467-2478.
Wildeman,A.G., Sassone-Corsi,P., Grundström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129-3133.
Yocum,R.R., Hanley,S., West,R. and Ptashne,M. (1984) *Mol. Cell. Biol.*, **4**, 1985-1998.

*Received on 23 September 1985*