The EMBO Journal vol.8 no.12 pp.3815 - 3824, 1989

The two DNA-binding domains of yeast transcription factor τ as observed by scanning transmission electron microscopy

Patrick Schultz, Nejib Marzouki[1], Christian Marck[1], Anny Ruet[1], Pierre Oudet and André Sentenac[1]

Laboratoire de Génétique Moléculaire des Eucaryotes, 11 rue Humann, 67085 Strasbourg Cedex and [1]Service de Biochimie du Département de Biologie, Bat 142, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette Cedex, France

Communicated by R.Cortese

Yeast transcription factor τ interacts with the intragenic promoter of tRNA genes, binding to both the A and B block elements. Affinity-purified τ factor and τ-tDNA complexes were examined by scanning transmission electron microscopy to analyze the structural features of free and DNA bound factor. The free factor appeared as two tightly associated globular domains of roughly similar size (10 nm in diameter) and mass (~300 kd). A combination of these two domains results in a mass for the factor of 510–670 kd. When τ was allowed to interact with recombinant $tRNA_3^{Leu}$ genes with variable A block–B block spacing, different structures were observed. With short genes, the two globular domains were not resolved and τ appeared as a large particle covering the A and B block region. On the other hand, with genes having a larger A–B distance (53 or 74 bp), mostly dumb-bell-shaped complexes were formed with individualized factor domains bound separately to the A and B blocks. A smaller proportion of the complexes appeared to consist of a large particle bound at only one site, essentially on the B block. Mapping of the binding domains in the DNA showed a good correlation with the respective positions of the A and B promoter elements. Factor binding did not induce a noticeable DNA bending, although with extended genes apparent DNA shortening and cases of DNA looping were observed. Upon cleavage of the $tRNA_3^{Leu}$ gene between the A and B blocks after or prior to complex formation, the two factor domains remained attached to the same DNA fragment (mostly the B-DNA fragment). In addition, images of protein-linked, reconstituted full-length genes were also observed. These different conformational states of the τ-tDNA complexes probably reflect the dynamic aspect of the interaction of the factor with its DNA target.

Key words: RNA polymerase C/TFIIIC/tRNA genes

Introduction

The transcription machinery for tRNA genes is a multiprotein apparatus that requires two general transcription factors to direct proper binding and accurate initiation by RNA polymerase C (Geiduschek and Tocchini-Valentini, 1988). The selective recognition of the tRNA genes by this machinery rests on the presence of two conserved intragenic sequence elements, the A and B blocks, which correspond to the invariant bases of the D and TΨCG loops of tRNAs (Ciliberto *et al.*, 1983; Geiduschek and Tocchini-Valentini, 1988). As a primary step of gene activation, the promoting sequences are recognized by transcription factor TFIIIC, or τ in the yeast system, in the absence of any other transcriptional component (Lassar *et al.*, 1983; Ruet *et al.*, 1984). Factor TFIIIB contributes to the stability of TFIIIC–DNA complexes and is required for transcription complex formation. It does not bind to DNA by itself (Klekamp and Weil, 1986; Waldschmidt *et al.*, 1988). Yeast factor τ and its interaction with tDNA have been extensively studied (Klementz *et al.*, 1982; Ruet *et al.*, 1984). Nuclease protection and methylation experiments (Camier *et al.*, 1985; Stillman *et al.*, 1985a), as well as point mutation analyses (Baker and Hall, 1984; Baker *et al.*, 1986), have demonstrated the specific binding of τ to the A and B blocks, the involvement of critical bases and the prominent role of the B block sequence in complex formation. Analysis of a series of $tRNA_3^{Leu}$ gene recombinants having variable A to B distances have demonstrated the remarkable flexibility of τ–DNA interaction (Baker *et al.*, 1987; Fabrizio *et al.*, 1987). Variation in A–B distance mostly affected A block binding, but the relative helical orientation of the A and B binding sites was found to be unimportant. The results suggest distinct roles for the two promoter elements: the B block is responsible for anchoring τ to the gene and thereby enhancing its binding to the A element. Interaction with the A block is the determining step in directing initiation by RNA polymerase C.

While the polypeptide composition of the factor is not yet precisely known, a multisubunit structure can be inferred from the large size of τ factor, estimated to be equivalent to a globular protein of 300–400 kd by sedimentation analysis (Ruet *et al.*, 1984; Stillman *et al.*, 1985b). Using specific antibodies and UV cross-linking, two distinct polypeptides of 145 and 100 kd have been found to interact directly with the coding strand of the tRNA gene (Gabrielsen *et al.*, 1989). Selective proteolysis experiments have indicated that τ factor contains two separate binding domains, called τ_A and τ_B, each interacting with one of the two promoter elements (Marzouki *et al.*, 1986). This structure of the factor has suggested a model in which τ would bind alternatively at the B and A regions to allow transcription of the intragenic promoter by RNA polymerase C. The situation in mammalian cells appears somewhat different in that the cognate transcription factor could be resolved into two functionally distinct components, C1 and C2 (Yoshinaga *et al.*, 1987). These activities may correspond to τ_A and τ_B. Like τ_B, TFIIIC2 interacts by itself with the B region of VA1 RNA genes. Addition of TFIIIC1 is required for A block binding and transcription, although TFIIIC1 does not bind by itself to DNA (Yoshinaga *et al.*, 1987). Cross-linking studies have identified in TFIIIC2 one polypeptide

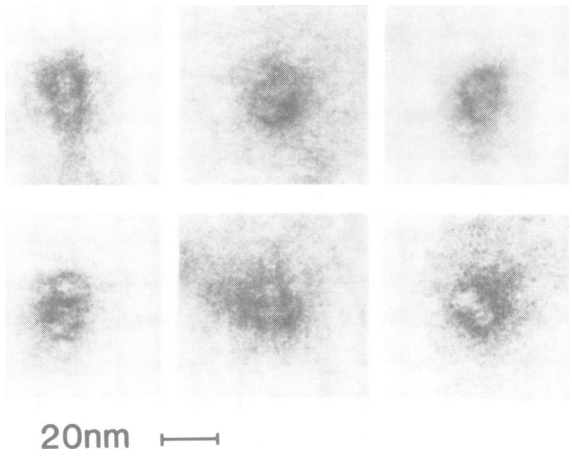

Fig. 1. Negative staining of purified factor τ preparation. Purified τ factor was adsorbed undiluted and negatively stained. The selected images show the two-domain structure of the factor. The otherwise observed particles appeared to be heterogeneous, but were compatible in size with variable orientations of the factor on the supporting carbon film. Mixed dark-field and bright-field images. Bar indicates 20 nm.

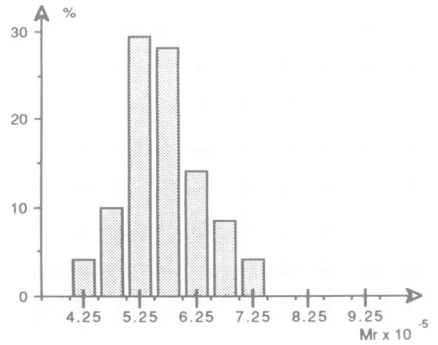

Fig. 2. Histogram representing the mass distribution of τ factor. The factor preparation was diluted 5-fold in double-distilled water prior to adsorption on glow-discharged grids and air-drying. The mass distribution is centered at 570 ± 50 kd and is heterogeneous (n = 71).

of 250 kd that interacts with DNA, but a multisubunit structure for TFIIIC2 remains a possibility in view of its large size (~400 kd) (Boulanger *et al.*, 1987).

In the present work we used STEM (scanning transmission electron microscopy) to visualize directly the molecular structure of τ factor as well as its mode of interaction with tRNA genes. The observation of purified τ factor revealed an organization of τ into two domains of roughly similar size. The mass and morphology of τ factor and τ–tDNA complexes were characterized using different $tRNA_3^{Leu}$ gene recombinants with variable A–B distances.

Results

Direct observation and mass determination of purified τ factor

Factor τ was purified by conventional ion-exchange chromatography and, as a last step, by specific binding to a tDNA–agarose column. The purified factor preparation contained several polypeptide chains, two of which (145 and 100 kd) have been shown to interact directly and specifically with tRNA genes (Gabrielsen *et al.*, 1989). The purified τ factor was negatively stained and observed by STEM. The

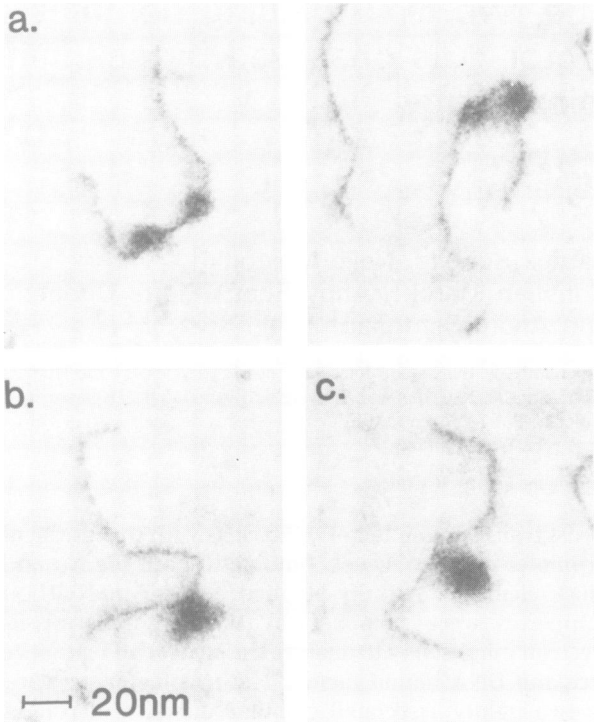

Fig. 3. Different types of nucleoprotein complexes observed with wild-type $tRNA_3^{Leu}$ (Leu-74) gene. τ factor preparation (10 ng of protein) was incubated with 2 ng of DNA in a 10 mM TEA, pH 7.4, 190 mM ammonium acetate, 10% (v/v) glycerol, 1 mM EDTA buffer for 30 min at room temperature, and the complexes were observed by STEM. (**a**) The predominant type of complexes (66%) contains two globular particles of ~10 nm in diameter which both interact with DNA. These complexes are therefore referred to as doublet or dumb-bell-shaped structures. (**b**) The second type of complexes (34%) appears to be constituted by a single and large globular structure of 14–16 nm size. These single particle complexes could correspond to partially collapsed doublet complexes. (**c**) Some single particle complexes appear often as laterally elongated structures presenting two domains, one of which is DNA bound. Mixed dark-field and bright-field signals. The bar represents 20 nm.

only reproducibly observed structures appeared as two globular particles of similar size (10–12 nm), arranged as a dimer or as a 20 × 12 nm ovoid-shaped particle with a central septum (Figure 1). Smaller particles (10–12 nm) were also observed which could correspond to dissociated factor, contaminants or to an unfavorable orientation of the molecule on the support.

In order to characterize the mass of the particles using analytical STEM, τ preparations were adsorbed and air-dried. The unstained τ factor appeared to be similar in size to the stained particle (data not shown). The separation between the two domains was no longer visible, probably due to imperfect structural preservation during air-drying. Although the structural information about these unstained structures is limited, the direct counting of the scattered electrons provides the mass distribution of the particles (Ohtsuki *et al.*, 1979; Freeman and Leonard, 1981; Engel *et al.*, 1982). The histogram obtained by measuring the relative electron scattering of 71 individual molecules showed that 80% of the particles have a mass ranging from 510 to 630 kd (Figure 2).

Binding of τ factor to the $tRNA_3^{Leu}$ gene

Purified τ factor was incubated with the wild type $tRNA_3^{Leu}$ gene (Leu-74; A–B distance: 74bp) under salt conditions

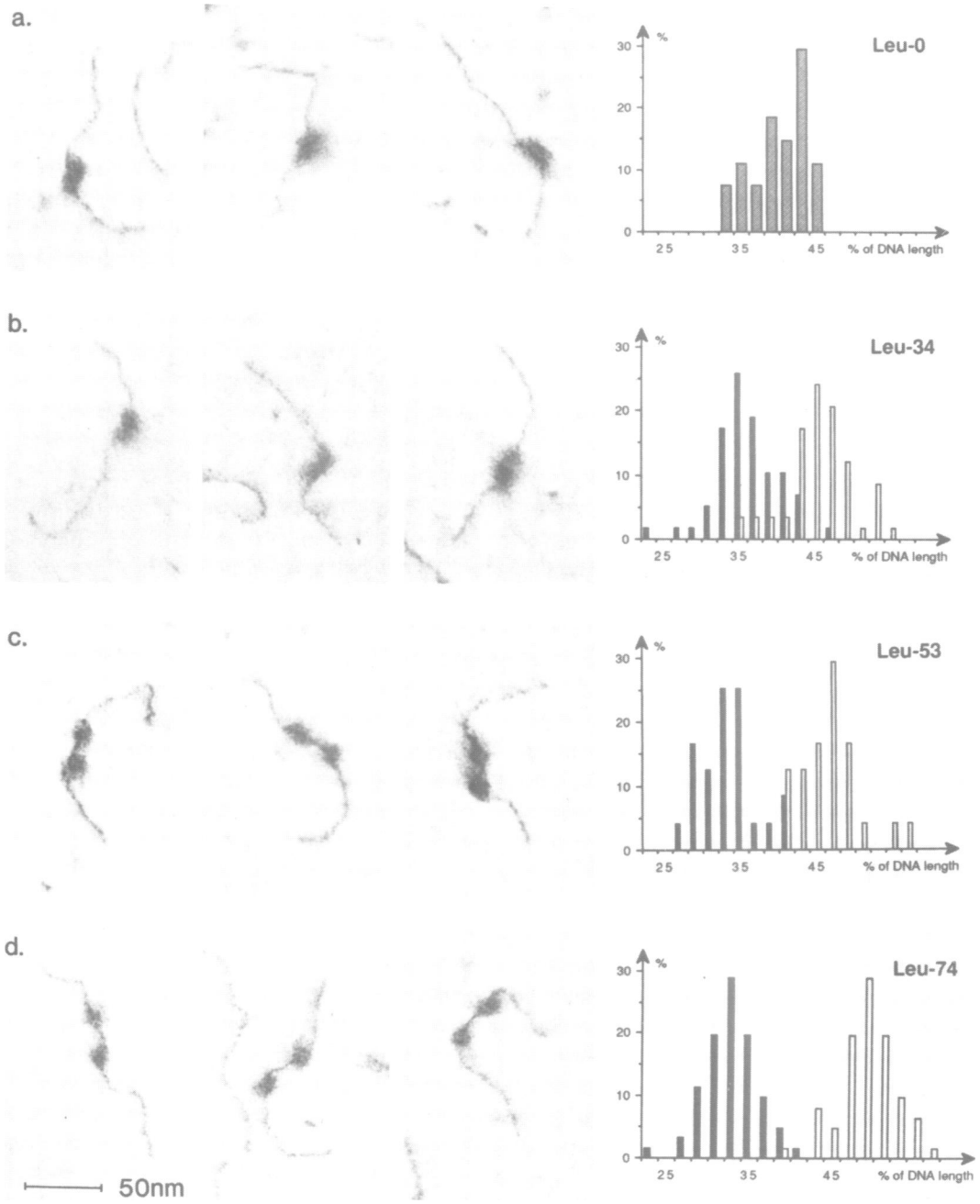

Fig. 4. Gallery of dumb-bell-shaped complexes observed upon interaction of factor τ with recombinant $tRNA_3^{Leu}$ genes having variable distances between the two promoter elements. The A block–B block distance is 0, 34, 53 and 74 bp in **a**–**d** respectively (Baker *et al.*, 1987). The histogram facing each gallery represents the distribution (in %) of the particles along the DNA. The location of the bound particle is expressed as a percentage of total DNA length. In the case of Leu-0 (n = 27), only single particle complexes were observed. When two globular particles were observed for Leu-34 (n = 58), Leu-53 (n = 24) and Leu-74 (n = 61), the binding location histogram of the particle interacting with the closest DNA end is shown in black, whereas that of the second particle is blank. Mixed dark-field and bright-field signals. The bar represents 50 nm.

(190 mM ammonium acetate) optimal for efficient and specific complex formation, as determined in gel-retardation assays (data not shown). The nucleoprotein complexes were stained with uranyl acetate and observed in STEM dark-field mode. About 15–20% of the DNA molecules were complexed to a globular, protein-like structure of 10–30 nm in diameter. DNA binding of these particles appeared to be very specific since 90% of the observed complexes mapped within 20% of the total DNA length (n = 70). The residual 10% aspecific binding was either random, or, for 70% of the cases examined, at the very end of DNA filaments and reflected probably a low affinity of τ or of contaminant proteins for free DNA ends.

The specific nucleoprotein complexes observed under these conditions presented essentially two characteristic forms. The most abundant species (66% of the specific complexes) consisted of two clearly individualized globular particles of similar size (10–15 nm, n = 35) interacting with the same DNA fragment at two different sites (Figure 3a). This type of complex is hereafter referred to as doublet or dumb-bell-shaped complex. The remaining specific complexes were composed of a single elongated

Table I. Binding specificity

tRNA gene recombinant	DNA length (bp)	Position of the promoter elements	Experimental binding sites in fractional distance from the closest end
Leu-74	492	block A: 194 bp (32.7%) block B: 280 bp (50.2%)	33 ± 3% (n = 61) 49 ± 3% (n = 61)
Leu-53	471	block A: 194 bp (34.2%) block B: 259 bp (48.0%)	33 ± 3% (n = 24) 45 ± 3% (n = 24)
Leu-34	452	block A: 194 bp 35.6%) block B: 240 bp (45.8%)	35 ± 3% (n = 58) 45 ± 3% (n = 58)
Leu-0	418	block A: 194 bp (38.5%) block B: 206 bp (41.4%)	39 ± 3% (n = 27)

protein particle of larger size (14 × 16 ±2 nm, n = 20; Figure 3b). These large particles sometimes appeared to be the doublet complexes that were bound to the DNA at one site only (Figure 3c). Complexes formed by a single small particle or by two large particles were not observed in significant amounts. These observations identify morphologically specific complexes formed upon interaction of τ factor with $tRNA_3^{Leu}$.

Binding of τ factor to different sized $tRNA_3^{Leu}$ genes

The specificity of interaction of τ with the internal promoter sequences was investigated by mapping the binding sites of the doublet complexes in the wild-type $tRNA_3^{Leu}$ gene (Leu-74) and in the deleted $tRNA_3^{Leu}$ genes Leu-53, Leu-34 and Leu-0 (see Materials and methods and Baker *et al.*, 1987). Characteristic views of the dumb-bell-shaped complexes obtained with these different genes are shown in Figure 4. Remarkably, the distance between the binding sites of the two particles of the doublet increased with the distance between the A and B blocks in the Leu-34, Leu-53 and Leu-74 genes. The Leu-0 gene appeared to be bound exclusively to a single large particle, probably because the two binding domains were next to each other. The binding sites on the DNA correlated well with the respective positions of the promoter elements (Figure 4, right, and Table I). These experiments clearly demonstrate that the dumb-bell-shaped complexes correspond to the two binding domains of factor τ, functionally defined as τ_A and τ_B, each of which recognizes a specific promoter element.

Length distribution of the different $tRNA_3^{Leu}$ genes complexed with τ

The mean length of the Leu-74 DNA filament (492 bp) was 151 nm (σ = 6 nm, n = 78; Figure 5a) (corresponding to 0.31 nm/bp). This length distribution was compared to that of DNA complexed with τ factor in order to explore the possibility of a protein-induced compaction. The apparent length of DNA in the particles was measured by following the shortest distance between entry and exit sites. The dumb-bell-shaped complexes formed with Leu-74, Leu-53 or Leu-34 showed a DNA length distribution identical to that of naked DNA (Figure 5b). On the other hand, the DNA length distribution of Leu-74 bound to a single large particle was heterogeneous and presented a bimodal distribution, with either normal size or shortened DNA molecules (Figure 5c). The major peak (70% of these single particle complexes) had the same DNA length (150 nm, n = 26) as naked DNA or DNA complexed with a doublet particle. This major subpopulation of complexes could be further divided into two classes according to the location of the particle binding site: in 75% of these complexes, the particle mapped near the B block, whereas only 25% of the complexes were located near the A block (data not shown). We suggest that this subpopulation of single particle complexes corresponds to the binding of τ to only one of the promoter elements: either to the A block or, preferentially, to the B block. The apparent larger size and often elongated shape of these particles (see Figure 3b and c), as well as their mass (see below), are compatible with partially collapsed dumb-bell structures, suggesting that all of the nucleoprotein complexes that were observed have an identical protein content.

The presence of a minor peak (30% of the single particle complexes) corresponding to an apparently shorter DNA (130 nm; Figure 5b) suggested a compaction of the nucleic acid upon binding of the large particle. This compaction may result from the formation of a DNA loop caused by protein−protein interactions between the two domains of the doublet particles. To investigate this possibility, $tRNA_3^{Leu}$ genes longer than Leu-74 were constructed to help make possible DNA loops easier to distinguish. The stretched genes Leu-82, Leu-99 and Leu-131 were transcriptionally active *in vitro* and formed stable complexes with factor τ, as was shown by gel-retardation assays (A.Ruet and C.Marck, unpublished results). As shown in Figure 6, where τ interacts with Leu-82 gene, when the A and B blocks were separated by a sufficiently long DNA segment (>8 helical turns), DNA loops could be visualized. Such protruding DNA loops were not observed for genes with shorter A−B distances, probably because they were hidden by the factor itself.

STEM mass determination of the nucleoprotein complexes

Mass measurements of the $tDNA_3^{Leu}-\tau$ complex were performed using the Leu-53 gene, where the A and B block binding domains are well separated. The unstained complexes appeared to be very similar to stained complexes: the DNA was observed with a low contrast, but the dumb-bell-shaped and single large particle complexes could be clearly distinguished (data not shown).

The mass distribution of all the observed specific complexes (defined as the mass of the protein particle and that of the DNA it covers) ranged predominantly from 600 to 860 kd and appeared to be heterogeneous (Figure 7a). The mass distribution of the single large particle complexes largely overlapped with that of the dumb-bell-shaped

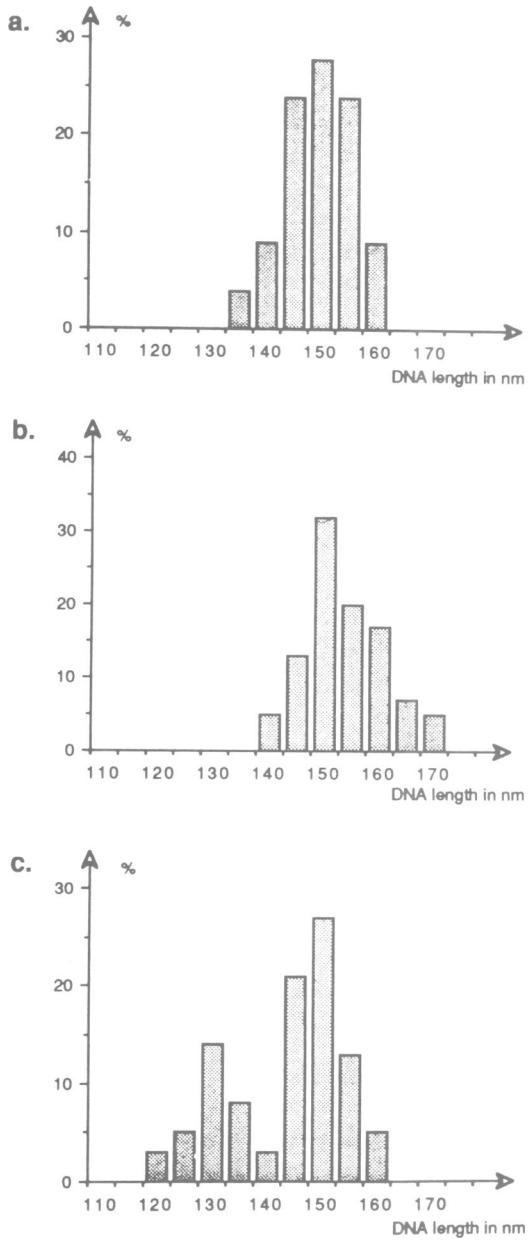

Fig. 5. Histograms of the length distribution of complexed $tDNA_3^{Leu}$ gene. The length of the spread DNA filaments was determined with a map-measurer on high magnification prints. (**a**) The length distribution of naked Leu-74 tDNA presents a peak centered at 151 ± 6 nm (n = 78). (**b**) The length distribution of Leu-74 in dumb-bell-shaped complexes is not significantly different than that of naked DNA (156 ± 9 nm, n = 60). (**c**) The length distribution of Leu-74 DNA in single particle complexes is bimodal. The major peak is centered at 150 ± 5 nm (n = 26), as is the case of the naked DNA. The second peak, centered at 130 ± 6 nm (n = 13), suggests a compaction of the DNA molecule by the bound particle.

complexes (Figure 7b), suggesting a close similarity in protein composition. The mass per unit length of naked DNA (1.1 kd/bp; data not shown) was significantly higher, in our experimental conditions, than expected (0.6 kd/bp), a fact that was probably due to salt accumulation. Therefore, the mass of the covered DNA (~100 bp) can, at least in part, account for the difference (100 kd) between the mass of the complex and that of the free factor.

Dumb-bell-shaped complexes could be analyzed in more

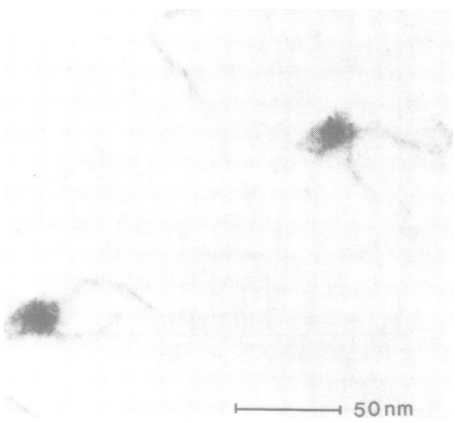

Fig. 6. Single particle complexes formed upon interaction of τ factor with Leu-82 tDNA frequently show an extended DNA loop protruding out of the globular particle. The two domains of the factor may interact with and constrain the DNA in this peculiar structure. Mixed dark-field and bright-field signals. The bar represents 50 nm.

detail, since the two binding domains were individualized and the location of each protein domain could be mapped to near the A or B block. The factor domain bound to the A block had essentially the same mass distribution as that bound to the B block. This mass distribution, which was somewhat heterogeneous, was centered at 370 kd (including the DNA) (Figure 7c). The two DNA binding domains of τ therefore have similar masses of ~300 kd.

Binding of τ to separated A and B promoter elements

We further investigated the binding specificities of the two factor domains and whether they could bind independently to excised A and B blocks (i.e. present on distinct DNA molecules). The $tDNA_3^{Leu}$ Leu-53 gene was cut with *Bam*HI restriction enzyme within the intervening region, yielding a small 180 bp fragment containing the A block (A-DNA) and a longer 291 bp fragment containing the B block (B-DNA). Both the A and B blocks mapped close to one end of their respective fragments. Factor τ incubated with each individual fragment was observed to bind at one end of the A-DNA and B-DNA fragments (Figure 8a and b respectively). Binding to the end of the A-DNA fragment was ~3–4 times less frequent than to the B-DNA fragment, thus reflecting a higher affinity or stability of B-DNA–τ interaction (Table II).

Factor τ was also incubated with an equimolecular mixture of A- and B-DNA fragments which could be easily identified by their difference in length. In addition to the individual complexes described above, the factor was also observed to interact simultaneously with one A-DNA and one B-DNA, thus reconstituting a pseudo full-length gene structure on 30% of the examined nucleoprotein complexes (Figure 8c). In contrast, τ was not observed to bridge two homologous DNA fragments in significant amounts (see Table II), even when incubated separately with either the A- and B-DNA fragment alone. The observation that τ links preferentially two heterologous DNA fragments strongly suggests that it contains two functionally individualized binding domains (τ_A and τ_B) that are specific for the A and B block respectively. In these experiments, the binding of one factor molecule bridging A- and B-DNA was ~3 times higher than the binding to A-DNA alone. This suggests an increased

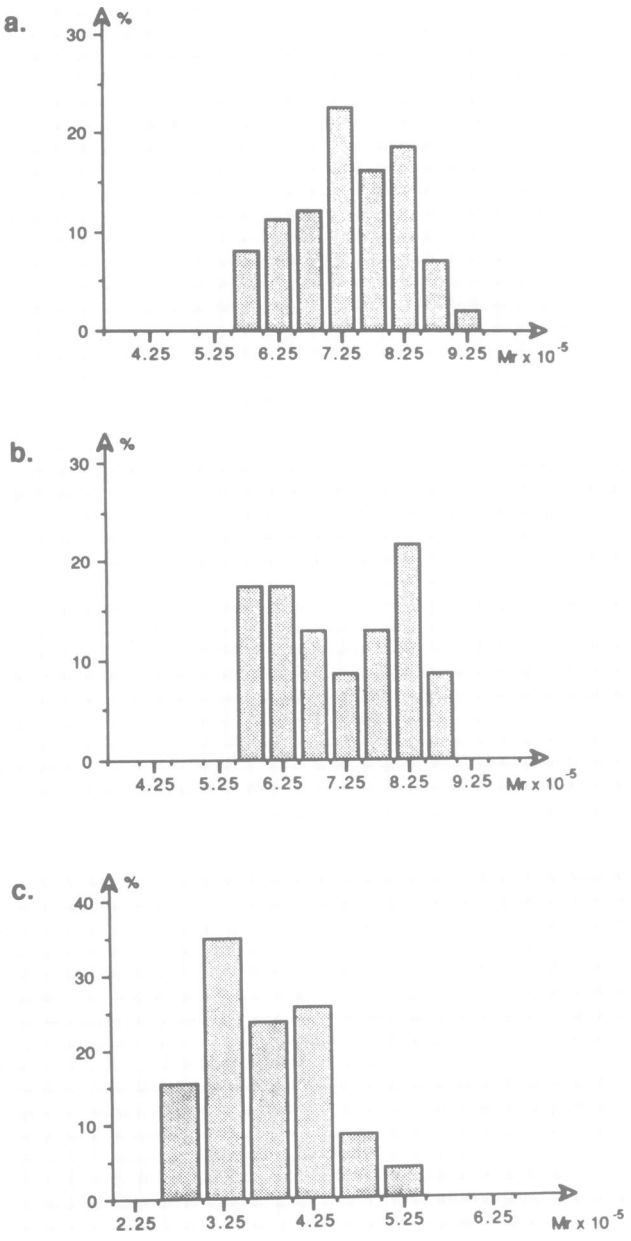

Fig. 7. Histograms of the mass distribution of τ factor bound to the Leu-53 tDNA. (**a**) Mass distribution of the dumb-bell-shaped complexes including the underlying DNA (~100 bp). The histogram is centered at 730 kd and shows heterogeneity (n = 88). (**b**) Mass distribution of single particle complexes. It is comparable to that of the dumb-bell-shaped complexes, suggesting a similar protein content (n = 20). (**c**) Mass distribution of each globular domain of factor τ bound to either the A or B block in dumb-bell-shaped complexes. The mass distribution of both domains was identical and centered at 360 kd (n = 188).

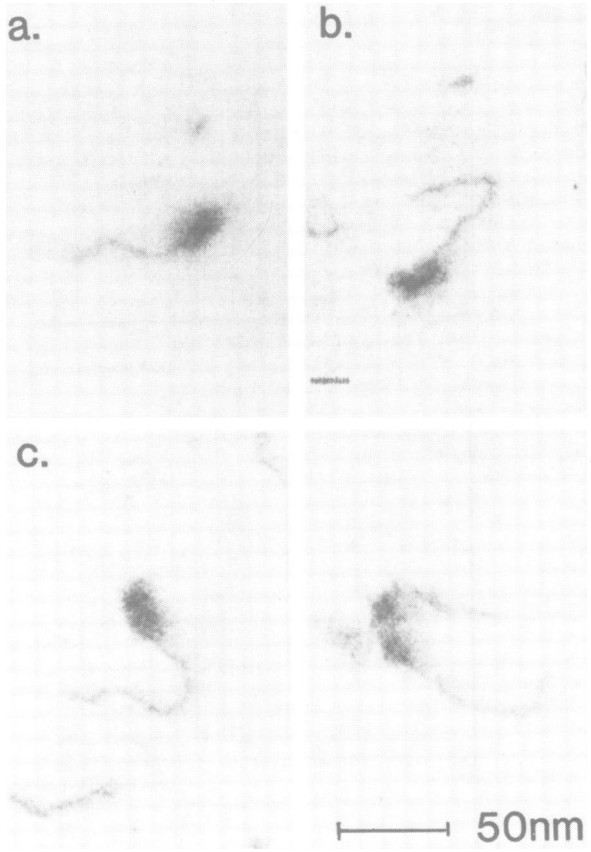

Fig. 8. Factor τ interaction with excised A and B blocks. Factor τ was incubated with the different DNA fragments containing either the A or the B block located close to one end, or with an equimolar mixture of the two fragments. (**a**) When incubated with the A block containing DNA (A-DNA), little binding was observed, and the binding that did occur took place mainly at the ends of the DNA molecules. The elongated shape of the bound protein entity suggests that both domains remain attached. (**b**) The incubation with B-DNA produced essentially the same complexes as in (a) except that the protruding DNA is longer and that the occurrence of complexes is higher. (**c**) In the presence of both A- and B-DNA, the complexes described in (a) and (b) were observed. In addition, 30% of the nucleoprotein complexes appeared as pseudo full-length DNAs in which factor τ bridges A- and B-DNA (right). Mixed bright-field and dark-field signals. The bar represents 50 nm.

affinity or stabilization of the A-DNA$-\tau$ interaction when the factor first interacts with the B promoter element. According to shape and size, it was always the entire factor molecule ($\tau_A-\tau_B$) that interacted with the separated halves of the Leu-53 gene and not individual domains, which suggests again a strong interaction between the two DNA-binding domains.

Band shift assays (Baker *et al.*, 1986) were used to examine further the binding of τ factor to Leu-74 and Leu-34 genes, which either induced the separation of $\tau_A-\tau_B$ (Leu-74) or left them closely associated (Leu-34) (see Figure 4b and d). Cleavage of the gene at the *Bam*HI site between the A and B blocks was performed either prior to or after complex formation. The results were essentially identical for the two genes (Figure 9). Incubation of τ with a mixture of A and B fragments gave rise to a retarded band of complex (lanes 5) that migrated like the control complexes formed with the uncleaved full-length genes (lanes 2). These complexes contained essentially the B-DNA fragments and very small amounts of A-DNA (data not shown). Cleavage of the DNA after complex formation again did not affect complex migration (lanes 3). Analysis of the DNA retained in the complexes revealed that the complexes contained 62% uncleaved DNA, 30% B-DNA and <5% A-DNA. This experiment confirmed and extended the conclusions derived from electron microscopy studies. In this experiment, factor τ was bound to the B-DNA fragments as a $\tau_A-\tau_B$ complex, since a dissociation of the two domains would generate a fast migrating complex, as described in the case of protease-resistant τ_B (Marzouki *et al.*, 1986). Even the cleavage of Leu-74 DNA after complex formation did not lead to the

Table II. Interaction of τ with excised promoter elements

Binding of τ factor to A-DNA	Percentage of specific nucleoprotein complexes relative to free DNA: 10% (8%, n = 73; 11%, n = 70) One A-DNA bound per τ factor: 95% Two A-DNAs bound per τ factor: 5%
Binding of τ factor to B-DNA	Percentage of specific nucleoprotein complexes relative to free DNA: 26% (35%, n = 42; 25%, n = 223) One B-DNA bound per τ factor: 93% Two B-DNAs bound per τ factor: 7%
Binding of τ factor to an equimolar mixture of A-DNA and B-DNA	Percentage of specific nucleoprotein complexes relative to free DNA: binding to A-DNA: 1% (<1%, n = 54; 2%, n = 50) binding to B-DNA: 8.5% (7%, n = 71; 10%, n = 61) Proportion of the different specific complexes One B-DNA bound per τ factor: 57% (57%, n = 78; 57%, n = 66) Two B-DNA bound per τ factor: 4% (4%, n = 78; 4%, n = 66) One A-DNA bound per τ factor: 9.5% (11%, n = 78; 8%, n = 66) Two A-DNA bound per τ factor: not detected One A-DNA and one B-DNA bound per τ factor: 29.5% (30%, n = 78; 29%, n = 66)

dissociation of the factor, although the two domains were apparently separated in the complex observed by STEM. Therefore, either the two domains remained physically linked in solution by a hinge region, or they rapidly reassociated (preferentially with the B block) after DNA cleavage.

Discussion

Transcription factor τ is one of the most complex DNA binding proteins, with its large size, its ability to footprint over eight helical turns of DNA, two distinct binding domains recognizing different sequence elements and, in addition, its remarkable feature of adapting to variable gene sizes. The intragenic location of the split promoter gives rise to additional complexity, as the factor is not released during multiple cycles of transcription.

Direct examination of factor τ and τ–tDNA complexes shed some light on its structure and on its mode of interaction with tRNA genes. Our main observations can be summarized as follows: (i) the factor appears to be a large macromolecule composed of two strongly interacting structural domains of ~300 kd each which, when bound to DNA, correspond functionally to the A and B block binding domains τ_A and τ_B; (ii) this doublet structure interacts preferentially with both promoter elements simultaneously, but also with the B block alone, or, less frequently, with the A block; (iii) with large genes, each domain appears to interact separately with one promoter element as if the factor domains were dissociated, although the existence of a connecting 'hinge' region could not be excluded; (iv) finally, in ~10% of the nucleoprotein complexes, factor binding induces a DNA conformational change which may be observed as DNA looping, especially with extended genes. DNA bending was not observed to a significant extent upon factor τ binding.

Structure of transcription factor τ

STEM has proven to be a reliable and accurate method which presents the unique advantage of simultaneously visualizing biological macromolecules and providing an estimate of their mass from electron scattering data (Ohtsuki *et al.*, 1979; Schultz *et al.*, 1986). Analyzed in this way, the structure of τ was characterized by the presence of two globular domains that were visible either after negative staining or when τ was bound to DNA. Their sizes (10–12 nm) were

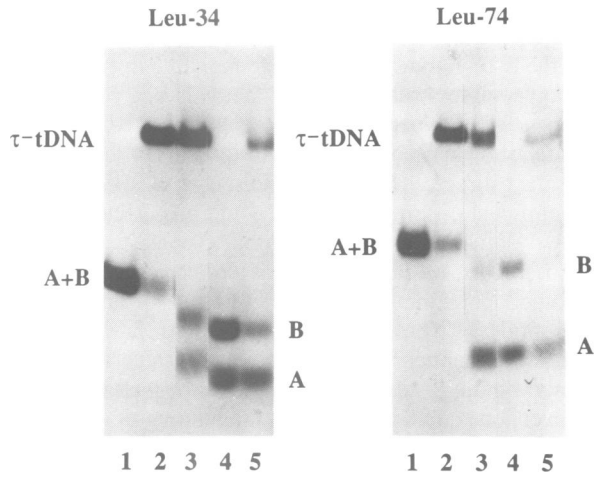

Fig. 9. Interaction of factor τ with cleaved tRNA genes as seen by gel retardation. Factor τ was incubated with the tRNA genes Leu-34 or Leu-74 either prior to or after cleavage of the gene between the A and B blocks with a restriction enzyme, and the complexes were analyzed by electrophoresis. **Lane 1**, uncleaved control DNA (A + B); **lane 2**, complex formation with uncleaved DNA; **lane 3**, DNA cleavage after complex formation; **lane 4**, control cleaved DNA (B and A fragments); **lane 5**, complex formation after DNA cleavage. Note that the migration rate of τ–DNA complexes is unchanged. The band of complex in lane 3 contained uncleaved DNA and B fragments; in lane 5, the band of complexes contained essentially the B fragment.

similar, and the significance of the small differences observed in their internal structure remains to be proven statistically.

The mass distribution of DNA-bound factor τ, once the mass of the underlying DNA had been subtracted, ranged from 510 to 760 kd and presented a large dispersion (250 kd). The observed τ factor appeared therefore to represent an heterogeneous population of particles, since a resolution of 50 kd could be expected from our previous results (Schultz *et al.*, 1986). Most of the heterogeneity could be related to the presence or absence of a putative entity of ~65 kd interacting with either domain. This variability may have different origins. It may reflect an instability of the factor itself and the loose binding of some component(s), or, alternatively, it could arise from the weak association of co-purified unrelated proteins. This mass distribution was similar to that obtained from free τ factor (510–630 kd), suggesting that the protein content of τ was not drastically

modified upon DNA binding. These concordant mass determinations are significantly higher than previous estimations derived from sedimentation data (Ruet *et al.*, 1984; Stillman *et al.*, 1985b), but are coherent with the large size of the observed bound factor and with gel-filtration experiments (P.Kenigsberg, unpublished results). The shape of the molecule or possibly the dissociation of the factor domains may have influenced the previous estimations. The structural heterogeneity observed with negatively stained factor corresponded probably to different orientations of the adsorbed complexes rather than to a real structural variability.

The mass distribution of the factor domains was determined when τ was bound to the Leu-53 gene which provided a good separation of the two globular entities. With less extended genes or in the case of free factor, the factor domains appeared to be too tightly associated to allow significant measurements. Their masses, which were somewhat heterogeneous, corresponded to a population of 320 ± 50 kd once the mass of the underlying 50 bp long DNA segment had been subtracted. To a first approximation, the two binding domains have the same mass. Whereas the polypeptide structure of the factor is not yet precisely known, two polypeptide chains of 145 and 100 kd were recently shown by antibody binding and UV crosslinking to contact directly tDNA (Gabrielsen *et al.*, 1989). Since no significant difference in mass could be detected between the two domains, a minimum model, consistent with our measurements, would be that τ has a dimeric structure, each monomer containing a copy of the 145 and 100 kd polypeptides. (Note that a model where τ_A and τ_B are structurally identical is apparently in contradiction with the distinct binding specificity of the two factor domains; however, one could imagine that the binding of one domain of τ to the B block induces a conformation change of the factor which commits the second domain to bind exclusively to the A block thus behaving as τ_A.) The opposite repartition (2 × 145 kd in one domain, 2 × 100 kd in the other one) also continues to be a possibility, especially since anti-100 antibodies did not interact with proteolyzed τ_B, whose B block binding activity was preserved, whereas anti-145 antibodies did (Gabrielsen *et al.*, 1989). However, the factor may be more complex and additional components may interact with these structures. A definitive molecular model must await more data on the polypeptide composition and stoichiometry of the factor.

DNA–protein and protein–protein interactions

The nucleoprotein complexes observed upon interaction of τ factor preparations with $tRNA_3^{Leu}$ genes presented characteristic binding properties that allowed the functional identification of the factor: protein binding was mapped at, or close to, the promoter elements in 80% of the complexes; and the preferential interaction with the B block versus the A block was also typical of τ.

Two different classes of structures were observed. The predominant structures were composed of typical dumb-bell-shaped molecules in which each domain interacted with one specific promoter element, as could be shown by the correlation between the observed binding sites and the known location of these elements in the DNA. Thus, our experiments allowed us to resolve the A and B block binding domains, τ_A and τ_B. The stretched complexes formed with Leu-74 or Leu-53 suggested that the two binding domains can be dissociated. The interaction was possibly weakened upon simultaneous binding to the A and B blocks, since separated domains were not observed with the free factor or when the factor was bound to DNA fragments containing a single promoter element.

The second class of complexes was composed of single large particles which, as judged from their size and mass, probably corresponded to the association of the two factor domains. The domains were probably not resolved because of their superposition or close association. In most cases, the factor was bound to the B block or, more rarely, to the A block. Therefore interdomain interactions predominated over the interaction with the second promoter element. In some cases (10% of all the complexes), the presence of a single particle complex was accompanied by an apparent DNA compaction. With very long genes, like Leu-82 or Leu-99, the observation of a small DNA loop protruding from the nucleoprotein complex suggested more directly the existence of an induced DNA constraint, with the factor apparently folding the DNA into a loop in order to interact simultaneously with the A and B blocks without disrupting its interdomain interactions. Some complexes appeared to be similar to the pictures obtained by Stillman *et al.* (1985b). However, these authors found that the majority of complexes that they observed had sharply bent DNA. This was not confirmed in the present study. Particulary, DNA bending was not observed when τ was bound to the DNA through one domain only.

Our results suggest the existence of two specialized domains, each of which recognizes a specific promoter element. The existence of specialized domains (τ_A and τ_B) was previously postulated after the isolation of a partially proteolyzed τ that had only B block binding activity (Marzouki *et al.*, 1986). That these two domains interact strongly was clearly evidenced by their simultaneous binding as a large particle to a single promoter element and by their striking ability to bridge two DNA fragments containing the two necessary promoter elements or to induce the formation of DNA loops. The predominant dumb-bell-shaped complexes with clearly separated particles formed with Leu-74 and Leu-53 suggested that the two domains can dissociate upon DNA binding.

The τ–DNA complex is therefore characterized by two distinct protein–DNA interactions coupled with a strong protein–protein interaction that connects the two DNA-binding domains. The combination of these types of interaction can account for the different structures observed by electron microscopy study. These different conformational states probably reflect the dynamic aspect of the interaction between τ and its DNA target.

On the molecular mechanism of τ action

A comprehensive explanation of τ action in promoting the early steps of tRNA gene transcription calls for an answer to, at least, two puzzling questions. (i) How does the factor accommodate the different A–B spacing of natural tRNA genes (from 33 to 91 bp) and the variety of relative helical orientations of the A and B blocks? (ii) How can the factor remain bound to tDNA during multiple cycles of transcription? Taking into account the present observations, our feeling is that these two questions are linked and that the answer may be unique. Many lines of evidence, including

the present work, indicates that the τ_A and τ_B domains strongly interact. At first sight, this makes it difficult to understand the geometric adaptability of τ–DNA interaction. However, the present work also shows the aptitude of the factor to undergo an apparent dissociation reaction upon binding to the A and B blocks, as if DNA binding could compete for or weaken $\tau_A-\tau_B$ interactions. Whether τ_A and τ_B are fully separated or still connected when bound to large genes like Leu-74 remains unsolved. Whatever the answer to this particular question, the strong interaction of τ_A and τ_B and their induced separation upon DNA binding suggest a general mechanism for the adaptability and tight association of the factor to tDNA during transcription, one that is in the line of our previous hypothesis (Baker *et al.*, 1987). A DNA-induced weakening of the otherwise strong $\tau_A-\tau_B$ interaction could account for the required flexibility in a model in which factor τ binds to the B block, after which τ_A binds non-specifically to neighboring DNA and subsequently searches for its specific promoter element, this being facilitated by DNA looping or/and by a reversible dissociation of the τ_A and τ_B domains. The possibility of having the $\tau_A-\tau_B$ complex bound alternatively to either one promoter element or the other may be important for leaving the coding strand accessible for transcription. All of these situations were indeed observed. It is probable that important conformational changes are induced upon DNA binding, and complex interactions are to be expected during the transcription process if one considers the enormous mass (>1200 kd) of the transcription machinery covering eukaryotic tRNA genes.

The present observations may also be relevant to another aspect of τ function that is required for 5S RNA gene transcription (Geiduschek and Tocchini-Valentini, 1988). The internal promoter of 5S RNA genes contains a conserved A box element, whereas no box B element is present. Therefore there was a possibility that only the τ_A domain is involved in assembling the 5S gene transcription complex. However, the recent finding that the B-block inhibits the interaction of τ with the TFIIIA-5S gene (Braun *et al.*, 1989) rather suggests that the whole factor is involved in transcription complex formation.

Materials and methods

Plasmids and DNA fragments

The construction of the plasmids harboring the different mutant $tRNA_3^{Leu}$ genes has been described previously (Baker *et al.*, 1987). The different genes are referred to according to the number of bases between the A and B blocks: Leu-82, Leu-74, Leu-53, Leu-34 and Leu-0. Two elongated genes, Leu-99 and Leu-131, were constructed by insertion of the two *Sau*3AI fragments from bp 3690–3736 and 3136–3214 from pBR322 into the *Bam*HI site of Leu-53. DNA fragments used in the STEM observations were obtained by *Ssp*I digestion of the various plasmids followed by PAGE and electroelution. All fragments harbor the same sequences outside the A–B region. The region to the 5′ side of the A block is 155 bp long, whereas the region 3′ of B is 241 bp long, thus allowing an unambiguous discrimination of the A and B binding domains. The total length of the wild-type Leu-74 *Ssp*I–*Ssp*I fragment is 492 bp.

Factor τ purification and complex formation

Factor τ was purified according to Camier *et al.* (1985) with the modifications recently described (Gabrielsen *et al.*, 1989), which include a phosphocellulose adsorption step and heparin and DEAE–Sephadex chromatography, followed by affinity chromatography on a tDNA–agarose column. The purified factor preparation contained several polypeptide chains, two of which (145 and 100 kd) were shown to be involved in DNA binding (Gabrielsen *et al.*, 1989). For complex formation, factor τ (~10 ng of protein) was incubated with 2 ng of DNA in 15 μl of incubation buffer containing 10 mM TEA, pH 7.4, 190 mM ammonium acetate, 10% (v/v) glycerol and 1 mM EDTA for 30 min at room temperature. The factor preparations used for the electron microscopy experiments were selected to form exclusively the specific retarded band of complex in gel-retardation assays in the absence of non-specific competitor DNA. Therefore, no competitor DNA was added during complex formation. For the gel-retardation experiments shown in Figure 9, factor τ (200 ng) was incubated with the 228 bp *Mae*II–*Hga*I DNA fragment harboring Leu-34 or with the 314 bp *Mae*II–*Hpa*II fragment harboring ^{32}P-end-labeled Leu-74 (5–10 fmol, 15 000 c.p.m.) in 10 μl mixtures containing 20 mM Tris–HCl, pH 8, 150 mM ammonium acetate, 20 mM ammonium sulfate, 5 mM $MgCl_2$ 25 μg bovine serum albumin, 0.75 μg poly d(I-C) and 5% (v/v) glycerol. After 10 min incubation at 25°C, the tRNA genes were cleaved with a restriction endonuclease, either *Bam*HI (5 U, for Leu-34) or *Hpa*I (5 U, for Leu-74), at a unique site located in between the A and B blocks (Baker *et al.*, 1987), for 20 min at 25°C. Then the mixtures were directly loaded onto a polyacrylamide gel for analysis of factor–DNA complexes by electrophoresis (Marzouki *et al.*, 1986). Alternatively, when indicated, the tDNAs were first cleaved with the restriction enzyme before complex formation and electrophoresis. The lengths of the restricted fragments were, for the *Mae*II–*Bam*HI fragment of Leu-34 (with the A block) 73 bp; for the *Bam*HI–*Hga*I fragment (B block), 155 bp.

Specimen preparation

Thick perforated grids were prepared (Fukami and Adachi, 1965) and covered with a thin (1.5–3 nm) carbon film. Ten microliters of free τ factor (~100 ng protein) was adsorbed on an air glow discharged grid and negatively stained with a 2% uranyl acetate solution. For mass measurements, the specimen was diluted 5 times in a 30 μg/ml tobacco mosaic virus (TMV) suspension in double-distilled water and allowed to dry.

The nucleoprotein complexes were applied onto an amylamine glow-discharged grid as described originally by Dubochet *et al.* (1971) and positively stained with uranyl acetate. The incubation mixture was diluted 10 times to a final DNA concentration of 0.2 μg/ml in the incubation buffer without glycerol (dilution buffer). Mild glutaraldehyde fixation (0.1–0.3%, 5 min at room temperature) did not modify the appearance of the complexes or the relative proportion of their different types. At higher glutaraldehyde concentrations (1%), the complexes appeared to be multi-crosslinked and presented an unrecognizable globular structure. For mass determination TMV was added to the incubation mixture in dilution buffer, and the specimen was then processed similarly except for the staining step, which was omitted.

Electron microscopy

STEM observations were performed on a Vacuum Generator HB5 microscope operating at 100 kV with a cold-field emission gun. The beam current was kept between 1 and 5×10^{-11} A. All the observations were performed at −130°C using a specially designed cold stage (Homo *et al.*, 1980). Analog STEM imaging used a mixed signal mode combining both bright-field and dark-field signals (Jones and Leonard, 1978). Images were recorded on standard 24 × 36 Ilford PanF roll films from the photographic screen at a magnification of 200 000×.

The mass was determined essentially as described previously (Schultz *et al.*, 1986). The digital dark-field images, whose intensities correspond to the number of scattered electrons that are detected and counted, were recorded at a magnification of 200 000×. Length measurements were performed using a map measurer on highly magnified prints. The instrumental error was determined to be ±2% by repetitive measurements of the same DNA filament (150 ± 3 nm, n = 30). The magnification was calibrated using catalase crystals.

Acknowledgements

We thank R.Baker and B.Hall (University of Washington, Seattle) for providing the recombinant tRNA genes. We are grateful to R.Swanson for critical reading of the manuscript. This work was supported by a grant from the French ministry of Research and Technology.

References

Baker,R.E. and Hall,B.D. (1984) *EMBO J.*, **3**, 2793–2800.

Baker,R.E., Gabrielsen,O.S. and Hall,B.D. (1986) *J. Biol. Chem.*, **261**, 5275–5282.

Baker,R.E., Camier,S., Sentenac,A. and Hall,B. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8768–8772.

Boulanger,P.A., Yoshinaga,S.K. and Berk,A.J. (1987) *J. Biol. Chem.*, **262**, 15098–15108.
Braun,B.R., Riggs,D.L., Kassavetis,G.A. and Geiduschek,E.P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 2530–2534.
Camier,S., Gabrielsen,O., Baker,R. and Sentenac,A. (1985) *EMBO J.*, **4**, 491–500.
Ciliberto,G., Castagnoli,L. and Cortese,R. (1983) *Curr. Top. Dev. Biol.*, **18**, 59–88.
Dubochet,J., Ducommun,M., Zollinger,M. and Kellenberger,E. (1971) *J. Ultrastruc. Res.*, **35**, 145–167.
Engel,A., Baumeister,W. and Saxton,W.O. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4050–4054.
Fabrizio,P., Coppo,A., Fruscoloni,P., Benedetti,P., Di Segni,G. and Tocchini-Valentini,G.P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8763–8767.
Freemann,R. and Leonard,K.R. (1981) *J. Microsc.*, **122**, 275–286.
Fukami,A. and Adachi,K. (1965) *J. Electron Microsc.*, **14**, 112–116.
Gabrielsen,O.S., Marzouki,N., Ruet,A., Sentenac,A. and Fromageot,P. (1989) *J. Biol. Chem.*, **264**, 7505–7511.
Geiduschek,E.P. and Tocchini-Valentini,G.P. (1988) *Annu. Rev. Biochem.*, **57**, 873–914.
Homo,J.C. (1988) *Electron Microsc.*, **1**, 92.
Jones,A.V. and Leonard,K.R. (1978) *Nature*, **271**, 659–660.
Klekamp,M.S. and Weil,P.S. (1986) *Arch. Biochem. Biophys.*, **246**, 783–800.
Klementz,R., Stillman,D.J. and Geiduschek,E.P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6191–6195.
Lassar,A.B., Martin,P.L. and Roeder,R.G. (1983) *Science*, **222**, 740–748.
Marzouki,N., Camier,S., Ruet,A., Moenne,A. and Sentenac,A. (1986) *Nature*, **323**, 176–178.
Ohtsuki,M., Isaacson,M.S. and Creewe,A.V. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1228–1232.
Ruet,A., Camier,S., Smagowicz,W., Sentenac,A. and Fromageot,P. (1984) *EMBO J.*, **3**, 343–350.
Schultz,P., Weiss,E., Colin,P., Régnier,E. and Oudet,P. (1986) *Chromosoma*, **94**, 189–198.
Stillman,D.J., Caspers,P. and Geiduschek,E.P. (1985a) *Cell*, **40**, 311–317.
Stillman,D.J., Better,M. and Geiduschek,E.P. (1985b) *J. Mol. Biol.*, **815**, 451–455.
Waldschmidt,R., Jahn,D. and Seifart,K.H. (1988) *J. Biol. Chem.*, **263**, 13350–13356.
Yoshinaga,S.K., Boulanger,P.A. and Berk,A.J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8768–8772.

Received on July 18, 1989; revised on August 7, 1989