

# Identifying a Transcription Factor Interaction Site on RNA Polymerase II

ANDREA M. SKANTAR<sup>1</sup> AND ARNO L. GREENLEAF<sup>2</sup>

*Department of Biochemistry, Duke University Medical Center, Durham, NC 27710*

We have generated a series of fusion proteins carrying portions of subunit IIC, the second largest subunit of *Drosophila* RNA polymerase I, and have used them in a domain interference assay to identify a fragment of the IIC subunit that carries the binding site for a basal transcription factor. Fusion proteins carrying a subunit IIC fragment spanning residues Ala<sup>519</sup>-Gly<sup>992</sup> strongly inhibit promoter-driven transcription in both unfractionated nuclear extracts and in reconstituted systems. The same fusion proteins similarly inhibit dTFIIF stimulation of Pol II elongation on dC-tailed templates, suggesting that the IIC(A519-G992) fragment, which carries conserved regions D-H, interferes with transcription by binding to dTFIIF. Finally, dTFIIF can be specifically cross-linked to a GST-IIC(A519-G992) fusion protein or to subunit IIC in intact Pol II.

RNA polymerase II    TFIIF    Transcription factors    Domain interference    *Drosophila*  
Fusion proteins

---

RNA polymerase II is a complex, multisubunit enzyme that plays a central role in eukaryotic gene expression (Sawadogo and Sentenac, 1990; Young, 1991, for recent reviews). Like the other nuclear RNA polymerases I and III, RNA polymerase II consists of two large subunits and eight or more small subunits. The largest subunit (ca. 215 kDa) is usually referred to as IIa (RPB1 in yeast), and the second largest subunit (ca. 140 kDa) is usually referred to as IIC (RBP2 in yeast) (Greenleaf, 1992). Although a complete picture of how individual subunits function during the transcription cycle has not yet emerged, the importance of subunits IIa and IIC during catalysis and in the regulation of transcription is clear from several lines of evidence. For example, sequence analysis of the two largest subunits from several organisms has revealed extensive evolutionary conservation (Palenik, 1992, and above references). Eukaryotic subunit IIa shares eight regions of significant homology with the  $\beta'$  subunit of eubacteria as well as with the corresponding subunit of archaeobacteria and the largest subunits of eukaryotic RNA polymerases I and III. The

conserved regions tend to concentrate toward the N-terminus. Eukaryotic subunit IIC contains nine regions homologous to the bacterial  $\beta$  subunit (and the archaeobacterial and polymerase I and III counterparts), with the regions of strongest similarity concentrated toward the C-terminus.

Functional similarity between the prokaryotic and eukaryotic enzymes, suggested by the striking evolutionary conservation, has been borne out by biochemical analysis. For example, both large subunits were shown by antibody inhibition or UV cross-linking to contact the DNA template and the nascent transcript (Bartholomew et al., 1986; Breant et al., 1983; Dissinger and Hanna, 1990; Gundelfinger, 1983). Further, immunoprecipitation of epitope-tagged yeast enzyme revealed that formation of "premature core" yeast RNA polymerase II, containing the yeast counterparts to *E. coli*  $\beta$ ,  $\beta'$ , and  $\alpha$ , proceeds in a manner analogous to bacterial RNA polymerase assembly (Kolodziej and Young, 1991). In addition, *E. coli*  $\beta$  subunit and the second largest subunits of wheat germ pol II and yeast polymerases I, II, and III were affinity cross-linked to nucleotide substrate analogues,

---

Received November 9, 1994; revision accepted March 9, 1995.

<sup>1</sup>Present address: Department of Plant Pathology, North Carolina State University, Box 7616, Raleigh, NC 27695.

<sup>2</sup>Address correspondence to Arno L. Greenleaf.

suggesting that the second largest subunits of prokaryotic and eukaryotic enzymes participate in binding the nucleotide substrates and in phosphodiester bond formation (Grachev et al., 1987; Grachev et al., 1989; Mustaev et al., 1991; Riva et al., 1987). Taken together, these results suggest that numerous aspects of RNA polymerase molecular architecture and function have been conserved throughout evolution.

Largely because of facile genetic selection schemes, much is known about possible functions of the  $\beta$  subunit. For example, in vitro studies of reconstituted bacterial RNA polymerase containing mutant subunits have shown that defects in the C-terminal one-third of  $\beta$  subunit cause biochemical defects such as decreased promoter clearance (Lee et al., 1991), changes in promoter selectivity (Glass et al., 1986b; Nene and Glass, 1984), altered polymerase propagation (Sagitov et al., 1993), changes in enzyme pausing and termination (Landick et al., 1990a; Landick et al., 1990b), or the inability to bind  $\sigma$  factor (Glass et al., 1986a).

In view of the structural similarities between bacterial  $\beta$  subunit and eukaryotic subunit IIC, it is reasonable to speculate that the eukaryotic subunit carries out functions similar to those of its bacterial counterpart. In contrast to the bacterial case, however, it has not yet been possible to dissociate any eukaryotic RNA polymerase into individual subunits and reconstitute active enzyme in vitro. On the other hand, mutations generated in yeast and *Drosophila* have provided some clues as to the role of the second largest subunit. For instance, certain mutations in the C-terminus of yeast subunit IIC (RPB2) are responsible for gene-specific transcriptional defects in vivo (Scafe et al., 1990b); this may indicate altered interactions with a transcription factor or factors. Other RPB2 mutations were isolated as suppressors of a mutation in the *RPB1* gene, which suggests that a physical interaction may occur between the affected regions of the two largest subunits (Martin et al., 1990). Similarly, suppressors of a mutation in the largest subunit of *Drosophila* RNA polymerase II mapped to the *RpIII40* gene that encodes subunit IIC (Mortin, 1990). Other single-base substitutions in the C-terminal one-third of *RpIII40* caused developmental defects or were homozygous lethal, implying that alterations to these domains are critical to the enzyme's structure or function in vivo (Chen et al., 1993; Mortin et al., 1992).

RNA polymerase II requires several accessory factors for accurate initiation and effective elongation in vitro (Zawel and Reinberg, 1993, for review); several of these interact directly with Pol II.

Notably, TFIIF (RAP30/74) was originally isolated from HeLa cells on the basis of its binding to immobilized RNA polymerase (Sopta et al., 1985). The 30-kDa RAP30 subunit of TFIIF binds to RNA polymerase II in the absence of the 74-kDa RAP74 subunit (Killeen and Greenblatt, 1992) and is sufficient to recruit the enzyme into preinitiation complexes (Flores et al., 1991). Sequences required for binding to RNA polymerase II have been localized to the central portion of RAP30, a stretch of which shows homology to the  $\sigma$  subunit of *E. coli* RNA polymerase (Sopta et al., 1989; Yonaha et al., 1993). Binding of TFIIF, RAP30, or the homologous rat liver factor  $\beta\gamma$  to RNA polymerase II prevents the enzyme from associating with nonspecific DNA sequences in gel mobility shift assays, a function also exhibited by  $\sigma$  factors (Conaway and Conaway, 1990; Killeen and Greenblatt, 1992). Further, TFIIF can bind to *E. coli* RNA polymerase, and this binding is competed by  $\sigma 70$  (McCracken and Greenblatt, 1991). These results suggest that TFIIF probably binds to a region of RNA polymerase II that is conserved between prokaryotes and eukaryotes and that aspects of the mechanism of RNA polymerase II recruitment to the promoter may also be conserved. However, the RNA polymerase II subunits that are the direct target of TFIIF action have not been investigated previously, and the subunit of RNA polymerase II to which RAP30 binds has not been identified.

The *Drosophila* counterpart to mammalian TFIIF was originally isolated from *Drosophila* Kc cell nuclear extracts and called factor 5 (Price et al., 1987). Factor 5 is composed of two subunits of 86 and 34 kDa; it is essential for transcription initiation in vitro, it stimulates the elongation rate of pure RNA polymerase II in vitro, and it binds to RNA polymerase II (Price et al., 1989). Based on these properties and the sequence of a cloned gene for the large factor 5 subunit, it is clear that factor 5 is the *Drosophila* homolog of TFIIF (Kephart et al., 1993). In this report, we will refer to factor 5 as dTFIIF.

Although dTFIIF has a high affinity for free RNA polymerase II, template competition experiments and gel filtration of elongating ternary complexes have shown that in vitro it interacts transiently with paused RNA polymerase II, then dissociates upon reentry of the enzyme into a productive elongation cycle (Price et al., 1989). The mechanistic details of dTFIIF's role in initiation and elongation, including its interactions with subunits of RNA polymerase II, are not clear. As an initial step toward clarifying details of these

events, we sought to identify the dTFIIF binding site on Pol II. We hoped to identify not only the subunit to which the factor binds, but also the specific portion of that subunit involved in the interaction.

Our approach was based on the following ideas and facts. First, we reasoned that a recombinant fusion protein carrying some or all of the dTFIIF binding site, if added in excess to a transcription reaction, might compete with intact Pol II for dTFIIF binding and thereby inhibit factor-dependent transcription; this "domain interference assay" thus could potentially identify the Pol II subunit, or fragment thereof, carrying the dTFIIF-interacting domain. Next, genetic and biochemical experiments in *E. coli*, briefly summarized above, suggested the IIC subunit (the  $\beta$  homologue) as a likely candidate for interacting with dTFIIF. Further, the cloned gene for the *Drosophila* IIC subunit was available. Finally, a precedent for this general approach was in the literature (Rappaport et al., 1988).

In the work we describe here, fusion proteins containing portions of *Drosophila* RNA polymerase II subunit IIC fused to  $\beta$ -galactosidase or glutathione-S-transferase (GST) were expressed in *E. coli*, purified, and used as probes to examine the interaction of dTFIIF with this subunit of Pol II. We demonstrate that fusion proteins containing a particular region of the IIC subunit interfere with the activity of *Drosophila* dTFIIF during initiation and elongation. We also show that dTFIIF can be cross-linked to a fusion protein containing this region of subunit IIC as well as to intact subunit IIC within native RNA polymerase II.

## MATERIALS AND METHODS

### Chemicals

Restriction enzymes and terminal deoxynucleotidyl transferase were from BRL or Boehringer Mannheim. [ $\alpha$ - $^{32}$ P]CTP and [ $^{125}$ I]Bolton-Hunter reagent [3-(4-hydroxyphenyl) propionic acid *N*-hydroxysuccinimide ester] were from ICN Biomedical. Ribonucleoside triphosphates were from Pharmacia. HEPES [4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid, free acid], phenylmethylsulfonyl fluoride (PMSF), and dithiothreitol (DTT) were from Calbiochem. Nitrocellulose used for blotting of iodinated GST-IIC(A519-G992) was from Bio-Rad. Nitrocellulose used for chemiluminescent detection of horseradish peroxidase-conjugated antibodies was from Amersham. Ethylene glycol-bis (succinimidyl succinate), or EGS,

was from Pierce. All other reagents were from Sigma.

### Expression and Purification of $\beta$ -Galactosidase-IIC Fusion Proteins

A 2.0-kb *EcoRI* fragment from *Drosophila RpIII40* genomic DNA clone pBHgt-31 (Hamilton et al., 1993) was used to probe a  $\lambda$ gt11 library of 0–12 h *Drosophila* embryo cDNA (from Tao Hsieh, Duke University) at high stringency. A phage containing a 2.0-kb cDNA fragment, corresponding to the C-terminal two-thirds of the *RpIII40* coding region, was obtained from this screen. A 2.0-kb partial *EcoRI* fragment containing the cDNA was subcloned from this phage into pUC19 to give pAS-D1. A 1.9-kb *SalI-HindIII* fragment from pAS-D1 was then subcloned into the expression vector pUR290 to give pGAL-IIC(A519-K1157). The construct pGAL-IIC(A519-G992) was generated by cloning a 1.5-kb *BamHI* subfragment from pGAL-IIC(A519-K1157) back into pUR290. The plasmid pAS-W1 consists of a 1.0-kb *EcoRI* cDNA fragment from the 3' end of *RpIII40*, cloned into pUC19. Digestion of this plasmid with *BamHI* generated a 540-bp fragment that was subcloned into pUR291, resulting in plasmid pGAL-IIC(G992-T1175).

The  $\beta$ -galactosidase ( $\beta$ -gal) fusion proteins were prepared by the following procedure. All steps were at 4°C or on ice unless noted otherwise. Five milliliters of an overnight bacterial culture containing the expression plasmid were inoculated into 500 ml LB + ampicillin (50  $\mu$ g/ml) and grown at 37°C until mid-log phase (1.5–2 h), then induced with 1 mM IPTG for 2 h. Cells were pelleted at 4000 rpm in a GSA rotor for 10 min, then frozen at –80°C for 15 min to overnight. Buffer L [20 mM Tris-HCl-Cl, pH 7.2, 200 mM KCl, 10 mM EDTA, 0.1 mM DTT, 0.1% PMSF (of a saturated solution in isopropanol)] was added at 3 ml per gram of wet cells and incubated on ice until thawed. Cells were sonicated with 3–4  $\times$  30-s bursts with a microtip at maximum setting. Nine more cell volumes of buffer L were added and mixed gently. The diluted lysate was spun at 12,000 rpm for 30 min in an SA600 rotor. The pellet fraction was solubilized in buffer S (8 M urea, 50 mM Tris-HCl-Cl, pH 7.9, 1 mM EDTA, 50 mM KCl, 0.1% PMSF) and homogenized in a glass homogenizer (Dounce or Wheaton). Solid DTT was added to 50 mM. The solution was gently stirred at 4°C for 1–2 h. The protein was then dialyzed vs. 25 mM HEPES, pH 7.6, 15% glycerol, 0.1 mM EDTA, 0.1 mM DTT, 0.1%

PMSF (HGEDP) overnight, with at least one change of buffer. After dialysis, the fusion protein solution was spun at 10,000 rpm for 10 min to remove any insoluble protein.

Further purification of  $\beta$ -gal fusion proteins was carried out on DEAE cellulose (DE52, Whatman). The salt of the fusion protein solution was increased to 100 mM KCl by the addition of HGKE (1 M KCl in buffer HGE). This solution was loaded onto a DE52 column equilibrated with 100 mM HGKEDP, washed with 5 column volumes of the same buffer, and step-eluted with 350 mM HGKEDP. The peak of protein was pooled, 25% w/v ammonium sulfate was added, and the solution stirred for 30 min. After centrifugation for 13 min at 45,000 rpm in a T865 rotor at 0°C, the pellets were resuspended in and dialyzed overnight versus HGEDP. The KCl concentration was then adjusted to 0.1 M with 1 M HGKEDP, and the fusion protein fractions were loaded onto a Pharmacia Superose 6 FPLC column equilibrated with 0.1 M HGKEDP. The extent of proteolysis of eluted fractions was assessed by SDS-PAGE. Selected fractions were pooled and concentrated in Centricon microconcentrators (Amicon). The protein concentration was determined using the Bio-Rad protein assay reagent. Commercially made  $\beta$ -gal used in the transcription assays was from Sigma. Lyophilized protein was denatured by suspension in buffer S and treated the same way as GAL-IIC(A519-K1157), except that  $\beta$ -gal was not subjected to chromatography.  $\beta$ -Galactosidase activity was measured by the method of Miller (1972).

#### *Expression and Purification of GST-IIC Fusion Proteins*

Plasmids based on the pGEX vectors (AMRAD Corp., Pharmacia) were generated, employing standard procedures analogous to those used for the  $\beta$ -gal fusions, to encode fragments of the IIC subunit linked C-terminally to GST (details in Skantar, 1993). The constructs analyzed are summarized in Fig. 1C.

Bacterial strains expressing GST fusion proteins were grown and induced generally as described by Smith and Johnson (1988). GST-IIC fusion proteins were found almost exclusively in inclusion bodies; thus, it was necessary to solubilize them from this insoluble fraction. This was accomplished basically as for the  $\beta$ -gal fusions, above. GST-IIC(E901-Q1041) and GST-IIC(D825-G992) were further subjected to GSH-agarose chromatography (Smith and Johnson, 1988). Sol-

ubilized GST-IIC(G776-I859) was isolated from a preparative SDS-polyacrylamide gel as described (Hager and Burgess, 1980), except that the acetone-precipitated protein eluted from the gel slice was resuspended in buffer S, then dialyzed against 50 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM DTT, 0.1% PMSF, to renature the protein. GST-IIC(G776-I859) was then subjected to GSH-agarose affinity chromatography. The GST used for in vitro transcription assays was purified from cell supernatant (Smith and Johnson, 1988). Each of these proteins was concentrated in Centricon-10 microconcentrators (Amicon) in order to remove the glutathione as well as to concentrate the fractions, so that a broad molar excess of fusion proteins could be examined in the transcription assays.

The binding of GST-IIC(A519-G992) to glutathione-agarose was poor under all conditions examined, most likely due to the size of the RNA polymerase IIC fragment relative to GST. Therefore, the inclusion bodies containing most of this fusion protein were prepared essentially as described (Frankel et al., 1991; Lin and Cheng, 1991). These protocols both involve hypotonic lysis of the outer membranes, believed to be the greatest source interference with proper fusion protein folding. The resulting inclusion body fraction was suspended in 10 ml buffer S and incubated at 4°C for 30 min, centrifuged at 60,000  $\times g$  for 30 min, then dialyzed against several exchanges of HGEDP and frozen at -80°C.

#### *Templates for In Vitro Transcription*

The preparation of templates A2 (carrying the *Act5C* promoter) and p*Ball-E* for transcription, was described previously (Coulter and Greenleaf, 1985; Price et al., 1987). The template pPCP consists of a *Bam*HI/*Eco*RI fragment of the *Drosophila* actin 5C gene subcloned into the vector pSP73, and was the kind gift of Dr. David Price (University of Iowa). *Pst*I-digested template plasmids were dC-tailed using terminal deoxynucleotidyl transferase as described previously (Price and Parker, 1984).

#### *Inhibition of KcN-Mediated and Reconstituted Transcription*

A complete description of in vitro transcription assays and the procedures for preparing *Drosophila* Kc cell nuclear extract (KcN) and transcription factors can be found in Price et al. (1987). Reconstituted transcription assays typically contained factor 3 (the *Drosophila* homolog of TFIIE) and a

fraction eluting from phosphocellulose between 0.3 and 0.4 M KCl (referred to as the P11-0.4 M step). This fraction contained transcription factors 5 (dTFIIIF), DmS-II, and an RNaseH activity, in addition to other unidentified factors required for accurate initiation, such as TBP. Although not a transcription factor, addition of a DNase inhibitor purified from *Drosophila* extracts was also necessary to prevent digestion of template by nucleases present in less pure transcription factor fractions (Sluder et al., 1987). For some inhibition reactions, the P11-0.4 M step fraction was preincubated with GAL-IIc(A519-K1157) or  $\beta$ -gal for 10 min at 25°C before the other required factors were added. In other reactions, all of the transcription factors were added together (see figure legends). A cocktail containing template, nucleotides, and other cofactors was added to start the reactions. After 20 min at 25°C, the reactions were stopped, extracted, and electrophoresed.

#### *Elongation Assay Using dC-Tailed DNA Templates*

Elongation assays were carried out essentially as described (Price et al., 1989; Sluder et al., 1988) except that the concentration of cold CTP was increased to 100 mM. dTFIIIF used for elongation assays was purified essentially as described (Price et al., 1989) with some minor modifications. Reactions contained 10 mg/ml dC-tailed pPCP and 1–4  $\mu$ Ci [ $\alpha^{32}$ P]CTP per reaction. A master reaction was prepared with 5–10 units of purified RNA polymerase II, template, GTP, CTP, and ATP, and buffer. This mixture was incubated under UTP-less conditions on ice for 5 min to allow the polymerase to bind to dC-tailed template. At the same time, dTFIIIF was preincubated with fusion proteins or control proteins at 25°C for 5 min, in tubes that contained UTP and enough KCl and HGE to give 50–100 mM KCl in the final 12.5- $\mu$ l reaction volume. After the preincubations, the master polymerase mixture was aliquotted into the dTFIIIF fusion protein reactions, and elongation from the dC-tailed templates was allowed to proceed for 3–10 min. Reactions were stopped and samples prepared as described for the specific transcription assay.

Inclusion of  $\mu$ g quantities of the IIc fusion proteins sometimes resulted in inconsistent recovery of the RNA after transcription. Thus, a radioactively labeled, 186 bp size and recovery standard was made by PCR amplification of *RpIII140* cDNA pAS-D1 with the primers [5'-CCGTACCGCGTACACATCTGCAAC and 5'-

AAGCTTATCGATCGGCGCAATGTT-3']. The PCR products were radioactively labeled during amplification with approximately 1 [ $\alpha^{32}$ P]dAMP incorporated per DNA strand. The labeled standard was phenol extracted, ethanol precipitated, washed with 70% ethanol, dried, and resuspended in TE. An aliquot of standard was added to the Sarkosyl solution used to stop the transcription reactions. This provided an internal standard for monitoring nucleic acid recovery during the subsequent extraction steps.

#### *Iodination of GST-IIc(A519-G992) With Bolton-Hunter Reagent*

The iodination of GST-IIc(A519-G992) was carried out as described (Bolton and Hunter, 1973). After passage through a Sephadex G50 column (buffer 0.1 M HGEDK), BSA (50  $\mu$ g/ml) was added to the peak fractions as a stabilizer. A portion of each peak fraction was denatured by adding an equal volume of 10 M urea, 15 mM DTT, and 0.1% of PMSF. The protein was allowed to denature on ice for 15 min, then was dialyzed versus HGEDP, and stored at 4°C.

#### *Chemical Cross-Linking of dTFIIIF to GST-IIc(A519-G992) and RNA Polymerase II*

dTFIIIF was incubated with iodinated GST-IIc(A519-G992) or with purified RNA polymerase II for 15–60 min at 25°C, under conditions similar to in vitro transcription reactions, including 25 mM HEPES, pH 7.6, 12% glycerol, 0.1 mM EDTA, and 50–150 mM KCl. A solution of 50 mM EGS was prepared fresh in DMSO; working dilutions were prepared in HGE such that 2 ml added to the 15-ml reactions would give the desired final concentration of EGS. After cross-linking for 15 min, the reactions were quenched with 75 mM lysine; 6  $\times$  SDS gel sample buffer was added immediately. The samples were heated to 100°C for 5 min, spun briefly, then loaded onto an SDS-polyacrylamide gel (Laemmli, 1970). The gels were electroblotted for 2 h as described (Towbin et al., 1979; Weeks et al., 1982). When iodinated GST-IIc(A519-G992) was used, the blots were air dried, wrapped in Saran Wrap, and autoradiographed. Before immunodetection of cross-linked products, the nitrocellulose membranes were dried under a heat lamp for 10 min per side to ensure protein adsorption to the membrane.

#### *Affinity-Purified Antibodies*

The preparation of Pol II subunit IIa antibodies was described previously (Lee et al., 1991). An-

tibodies to domains of RNA polymerase IIc were obtained by subcutaneous injection of 0.5 mg denatured IIC(A519-G992) into a rabbit, 0.5 mg IIC(G992-T1175) into another rabbit, and 1 mg each protein into a goat as described (Weeks et al., 1982). After a second injection, serum was obtained from the animals and stored at  $-20^{\circ}\text{C}$ . The titer of antibodies was checked by Western blotting (Towbin et al., 1979). Fusion protein columns were constructed by binding 9.1 mg GAL-IIC(A519-G992) or 4.7 mg GAL-IIC(G992-T1175) to 1-ml aliquots of Reacti-gel (Pierce) according to the manufacturer's instructions. Affinity purification of the antibodies was performed as described (Robbins et al., 1984).

#### *Western Blotting and Chemiluminescent Detection*

Western blotting was as previously described (Weeks et al., 1993). Identical blots of dTFIIF cross-linked to RNA polymerase II were probed with antibodies affinity purified against the 215-kDa (anti-exon 2) or 140-kDa subunits. The blots were incubated with a 1:5000 dilution of primary antibody overnight at  $4^{\circ}\text{C}$ . Bound antibody was detected by incubation for 1 h with 1:10,000 dilution of horseradish peroxidase-conjugated swine anti-goat IgG (Tago). The blots were developed using ECL chemiluminescent antibody detection system according to the manufacturer's instructions (Amersham).

## RESULTS

### *Fusion Protein GAL-IIC(A519-K1157) Inhibits Promoter-Dependent Basal Transcription In Vitro*

*Drosophila* Kc cell nuclear extracts (KcN) support specific, promoter-driven transcription in vitro. To identify critical transcription factor-RNA polymerase II interactions, we first examined the ability of a fusion protein carrying a large fragment of the RNA polymerase IIc subunit to inhibit this process. The nuclear extract was incubated with a molar excess of GAL-IIC(A519-K1157) (Fig. 1) or  $\beta$ -gal for 5 min, before the addition of DNA template, nucleotides, and other cofactors (see Materials and Methods). A 50–75% reduction in the 450-base run-off transcript from the actin A2 promoter was observed when fusion protein GAL-IIC(A519-K1157) was present in a 500–1000-fold molar excess over the endogenous Pol II (Fig. 2, lanes 7–9). A similar excess of  $\beta$ -gal

caused slight if any reduction in the level of transcript (Fig. 2, lanes 5–6). Because neither fusion protein nor  $\beta$ -gal exhibited RNase activities (Fig. 2, lanes 2–3), this experiment clearly shows that the RNA polymerase II domain present in the GAL-IIC(A519-K1157) fusion protein interferes with some component of the KcN extract that is required for transcription in vitro.

We attribute the high molar excess of fusion protein required for significant inhibition in this and subsequent experiments to the necessity to denature and then renature insoluble fusion proteins (Skantar, 1993, and Materials and Methods). Presumably only a fraction of the polymerase portion of the fusion protein adopts its correct conformation after this treatment. We base this statement on measurements of the specific activity of the  $\beta$ -gal moiety of the fusion protein, which we monitored frequently and often found to be  $\leq 25\%$  of that of commercial  $\beta$ -gal treated similarly (Lee and Greenleaf, 1991; Skantar, 1993). Note that we always subjected the control proteins ( $\beta$ -gal, GST) to the same denaturation/renaturation treatment as the fusion proteins.

To determine which factor within the *Drosophila* extract was inhibited by the fusion protein, the effect of GAL-IIC(A519-K1157) on reconstituted in vitro transcription was investigated. Initial experiments were done using partially purified transcription factors obtained from fractionation of *Drosophila* Kc cell nuclear extracts (Price et al., 1989). Figure 3 shows that including up to a 300-fold molar excess of GAL-IIC(A519-K1157) over the polymerase in the reaction reduced the level of reconstituted transcription by as much as 75% (Fig. 3, lanes 2–8). This inhibition was enhanced 30–50% by preincubation of the fusion protein with the P11-0.4 M fraction before addition of the other factors (Fig. 3, lanes 10–12). Inclusion of  $\beta$ -gal did not significantly inhibit transcription, even at 3000-fold excess over RNA polymerase II (Fig. 3, lanes 8–9). Taken together, these results show that the subunit IIC segment carried by GAL-IIC(A519-K1157) interacts with one or more factors in the P11-0.4 M fraction that are required for in vitro transcription.

### *Localization of the Transcription Factor Interacting Domain of RNA Polymerase II Subunit IIC*

To determine whether fusion proteins containing subfragments of GAL-IIC(A519-K1157) were capable of interfering with transcription factors during promoter-dependent transcription, we

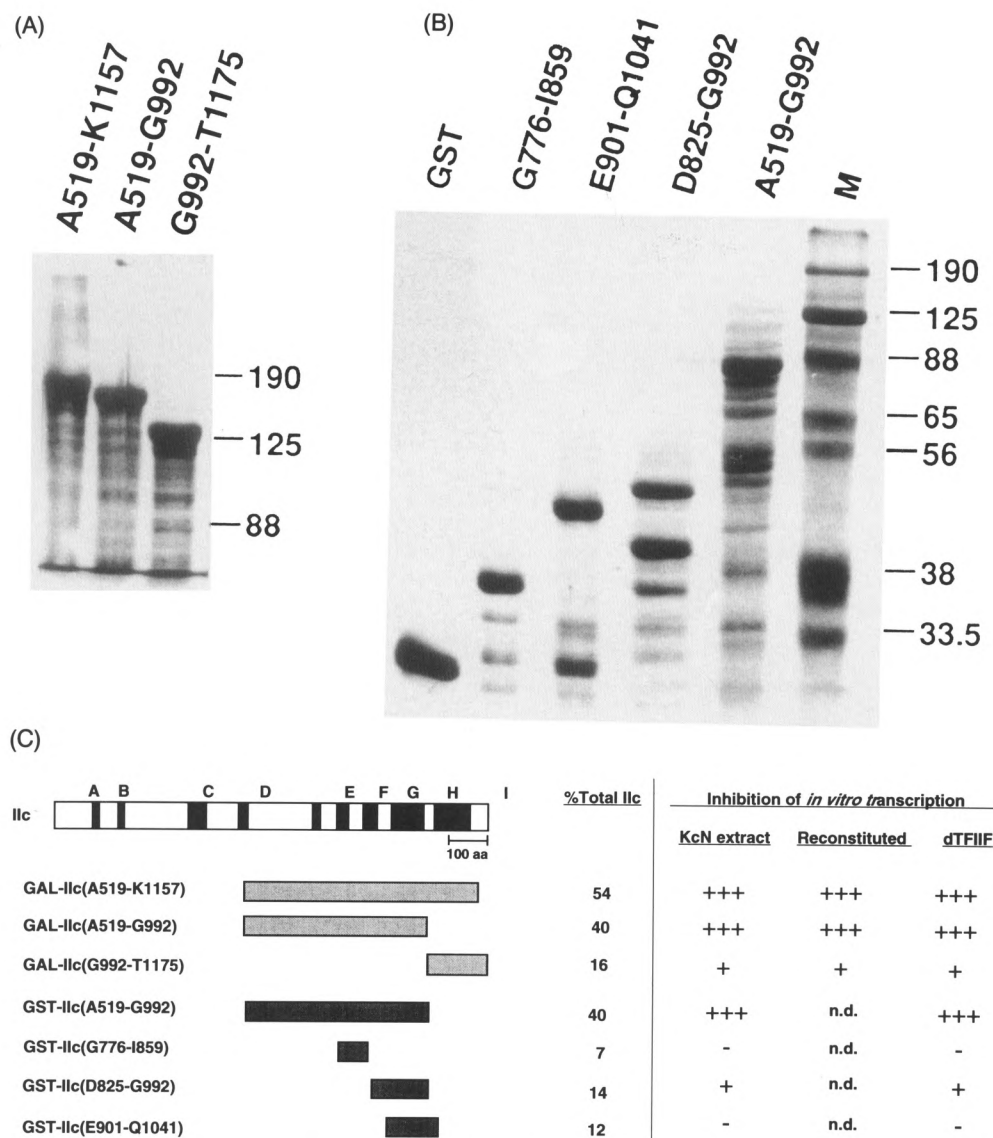
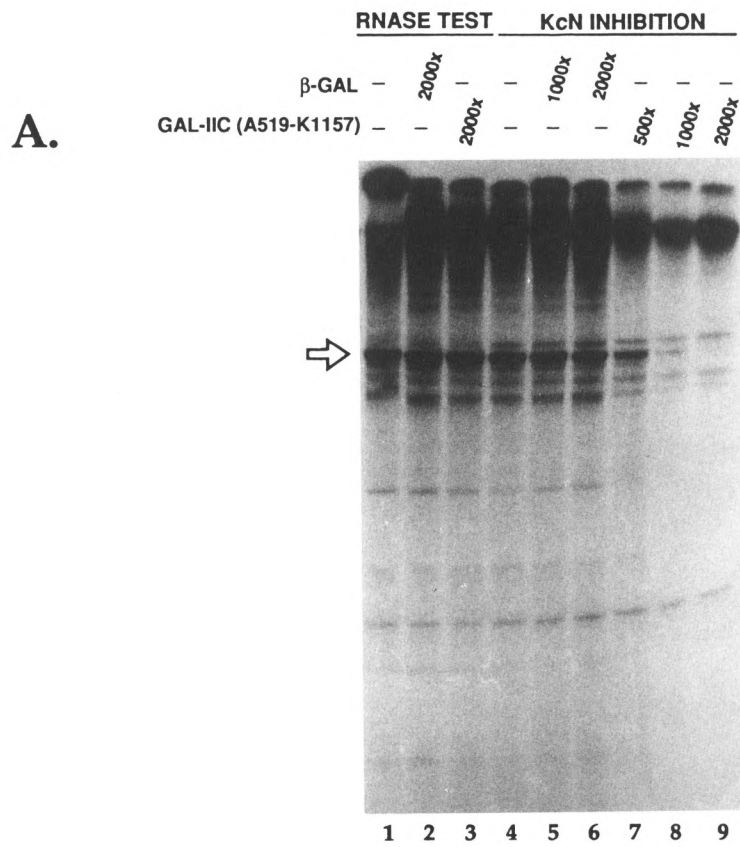


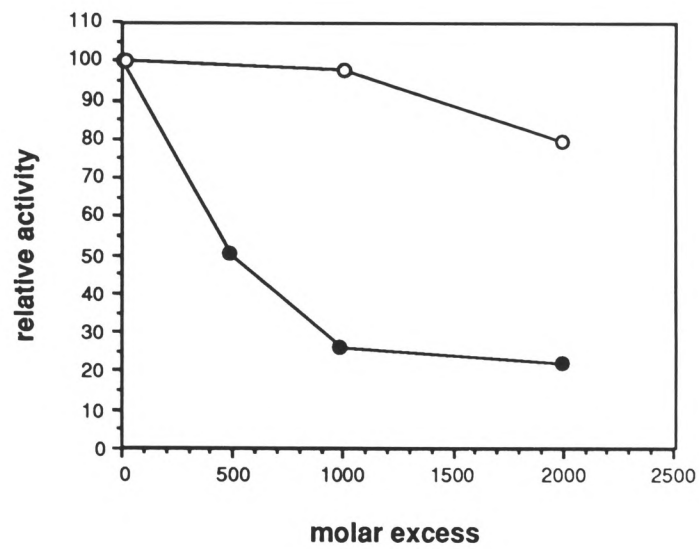
FIG. 1. Fusion proteins used for inhibitions. Coomassie blue stained SDS-polyacrylamide gels of (A) the  $\beta$ -galactosidase-IIC fusion proteins (10  $\mu$ g each lane) and (B) glutathione-S-transferase-IIC fusion proteins (2  $\mu$ g unproteolyzed) used for inhibition studies. Purification of the fusion proteins is described in Materials and Methods. (C) Schematic diagram of the RNA polymerase IIC subunit and fusion proteins in this study. Black boxes denote regions conserved between eukaryotic subunits IIC and eubacterial  $\beta$  subunit. Gray boxes denote the extent of GAL-IIC and GST-IIC fusion proteins, compared to the full-length IIC sequence. The percent of the whole subunit in each fusion protein is indicated. To the right of the vertical line, the degree of inhibition by the listed fusion proteins is summarized (data from Results section). Inhibition of "KcN extract" and of "reconstituted" transcription refers to promoter-directed transcription of the A2 template; inhibition of "dTFIIF" refers to stimulation of elongation in the dC-tailed template assay. n.d. = not done.

studied the effect of GAL-IIC(A519-G992) and GAL-IIC(G992-T1175) on the P11-0.4 M step fraction during reconstituted *in vitro* transcription assays. Increasing amounts of these fusion proteins were preincubated with the P11-0.4 M fraction as described above. Figure 4 shows that a 28-fold molar excess of GAL-IIC(A519-G992) over the RNA polymerase II present in the reaction is sufficient to achieve a 50% reduction in the

amount of run-off transcript. In the presence of a 138-fold or greater excess GAL-IIC(A519-G992), less than 20% of the control activity remains. In contrast, GAL-IIC(G992-T1175) inhibited transcription weakly, requiring more than seven times as much fusion protein to diminish the amount of transcript by 50%. Thus, we have identified an RNA polymerase subunit IIC segment required for a functional interaction between a basal transcrip-



**B.**





tion factor and RNA polymerase during promoter-dependent transcription.

*GAL-IIC(A519-K1157) and GAL-IIC(A519-G992) Inhibit dTFIIF (Factor 5) Stimulation of RNA Polymerase II Elongation*

Having localized the inhibitory region of subunit IIC to within the fragment defined by amino acids 519–992, we next sought to determine the basal transcription factor that was the target of the inhibition. Price et al. had purified *Drosophila* transcription dTFIIF (factor 5) from the P11-0.4 M step and showed that this factor was required for initiation of transcription (Price et al., 1987). Further characterization of highly purified dTFIIF revealed that it could also dramatically stimulate the elongation rate of RNA polymerase II during a promoterless transcription assay, and that it could bind stably to free RNA polymerase II in glycerol gradients (Price et al., 1989). Given the cosedimentation of dTFIIF and RNA polymerase II on glycerol gradients, its requirement for initiation, and its stimulatory effects during elongation, we decided to investigate whether interactions between GAL-IIC fusion protein and dTFIIF could lead to inhibition of dTFIIF activity during in vitro transcription. The dC-tailed template assay provides a way to evaluate the effect of individual transcription factors on the Pol II elongation rate, in the absence of initiation from a promoter. To establish that the fusion proteins did not interfere with the elongation rate of RNA polymerase II itself, GAL-IIC(A519-K1157), GAL-IIC(A519-G992), GAL-IIC(G992-T1175), or  $\beta$ -gal were added to dC assays in the absence of dTFIIF. RNA polymerase II was allowed to bind to the dC-tailed template and then fusion proteins or  $\beta$ -gal were incubated with the complex, and then elongation was initiated with UTP (see Materials and Methods). None of the proteins reduced

the elongation rate of RNA polymerase II alone; rather, they all slightly stimulated the number of transcripts seen. In addition, the transcripts were not degraded by fusion protein fractions either in the presence or absence of dTFIIF, indicating an absence of RNase activities (not shown).

GAL-IIC(A519-K1157), GAL-IIC(A519-G992), and GAL-IIC(G992-T1175) were tested individually over a wide range of molar excess relative to RNA polymerase II for reduction of dTFIIF stimulation. Increasing amounts of GAL-IIC(A519-K1157) caused a decrease in dTFIIF stimulation of the Pol II elongation rate. Although transcription was slightly stimulated at the lowest levels of GAL-IIC(A519-K1157) tested, at higher amounts the production of dTFIIF-dependent RNAs in the 1000–2000 base size range was dramatically reduced. In contrast, production of short, largely dTFIIF-independent RNAs was not affected. Fusion protein-dependent reduction in transcription leveled off between 80-fold or greater GAL-IIC(A519-K1157) (data not shown). Similarly, when excess GAL-IIC(A519-G992) was incubated with dTFIIF, between 1- and 20-fold GAL-IIC(A519-G992) slightly stimulated elongation (Fig. 5, lanes 5–8). However, at greater than 20-fold excess GAL-IIC(A519-G992), dTFIIF stimulation was dramatically reduced (Fig. 5, lanes 9–12). Similar to the results with GAL-IIC(A519-K1157), greater than 500-fold GAL-IIC(A519-G992) caused no further inhibition of dTFIIF activity (not shown). GAL-IIC(A519-G992) primarily diminished the production of long RNA, but a slight increase in some shorter transcripts was seen with 500-fold excess GAL-IIC(A519-G992) (Fig. 5, lane 12). Note that the level of transcription in the presence of 500-fold GAL-IIC(A519-G992) was only slightly higher than in the presence of  $\beta$ -gal but no dTFIIF, indicating that the fusion protein almost completely abolished the dTFIIF stimulatory activity (Fig. 5,

FACING PAGE

FIG. 2. Inhibition of specific initiation by GAL-IIC(A519-K1157) or  $\beta$ -galactosidase. The amount of RNA polymerase II (mass  $\approx$  600,000 Da) in Kc cell nuclear extract (KcN) was estimated to be 10 ng/ $\mu$ l (A. Sluder, personal communication). Thus, reactions containing 3  $\mu$ l KcN had an estimated 50 fmol of RNA polymerase II. The molar excess of fusion protein added to transcription reactions was calculated using the concentration of unproteolyzed GAL-IIC(A519-K1157) determined by comparison to a known amount of pure  $\beta$ -galactosidase on a Coomassie blue-stained SDS-polyacrylamide gel (not shown). (A) Run-off transcription assay. Lanes 1–3: test for RNase activity of the fusion proteins. Identical transcription reactions were incubated for 20 min, stopped by the addition of  $\alpha$ -amanitin (2  $\mu$ g/ml final), then incubated for an additional 5 min with the indicated proteins. In lanes 4–9, HGE buffer,  $\beta$ -galactosidase, or fusion protein GAL-IIC(A519-K1157) as indicated were incubated with 3  $\mu$ l KcN extract for 5 min before addition of the mixtures to DNA, nucleotides, and cofactors as described in Materials and Methods. Transcriptions were allowed to proceed for 20 min. The 450-base run-off transcript (linearized A2 template) is indicated by the arrow. In (B), the amount of run-off transcript was quantitated by densitometer scanning of the autoradiograph on an ImageQuant computing densitometer (Molecular Dynamics). The entire autoradiograph was scanned, and the desired bands were marked for quantitation. An equal-sized area next to the RNA bands was chosen as the background value and subtracted from each sample. The values from the uninhibited transcription reactions were set at 100%. The relative activity of inhibition reactions was calculated by dividing by this value.

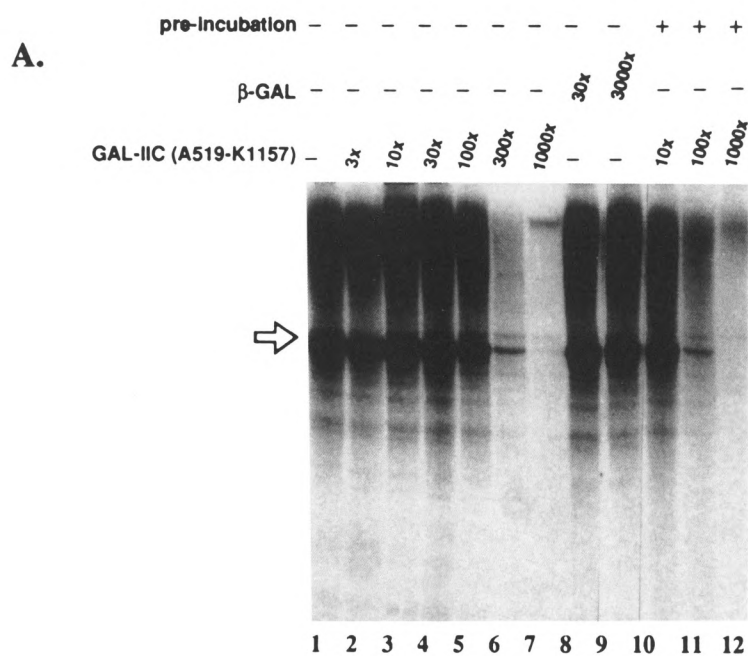
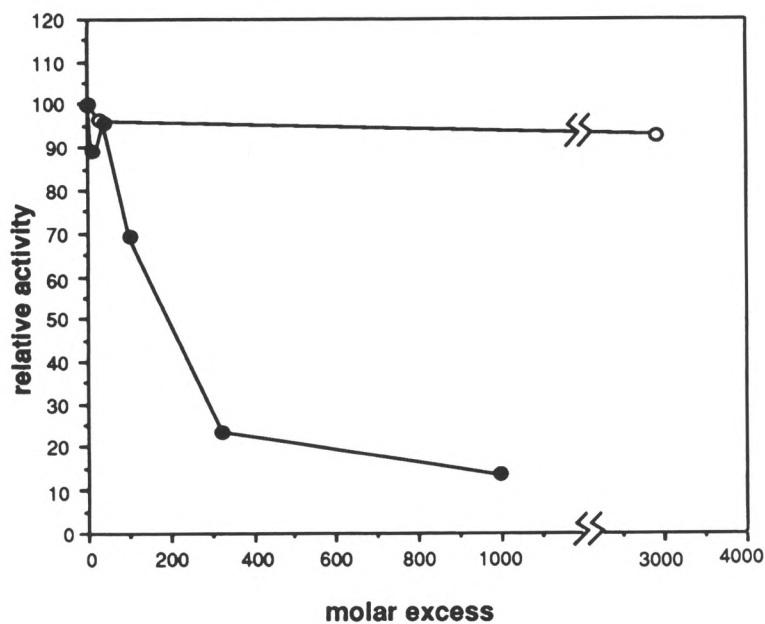
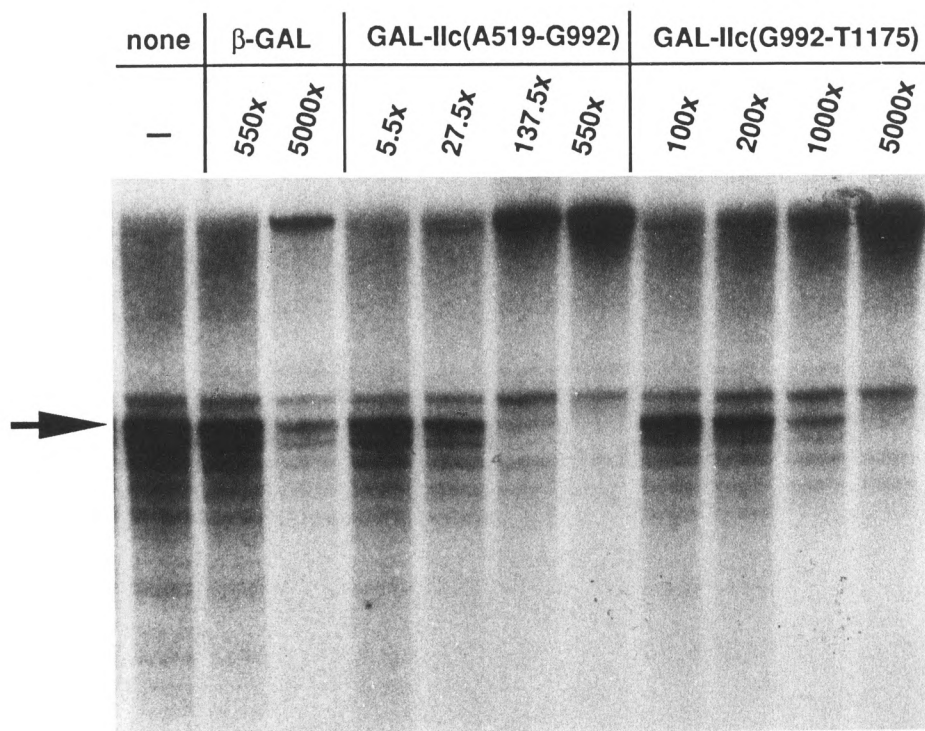
**B.**

FIG. 3. Inhibition of reconstituted transcription by GAL-IIC(A519-K1157). Each reaction contained 5  $\mu$ g/ml linear A2 template, 10 units (4.8 fmol) embryo RNA polymerase II, 0.0625  $\mu$ l DNase inhibitor (Sluder et al., 1987, MonoQ fraction), 0.5  $\mu$ l factor 3 (MonoQ fraction), 2  $\mu$ l P11-0.4 M step, and the indicated molar excess of GAL-IIC(A519-K1157) or  $\beta$ -galactosidase over RNA polymerase II. The 450-base actin run-off transcript is indicated by the arrow. (A) Run-off transcription assay. In lanes 1-9, GAL-IIC(A519-K1157) or  $\beta$ -galactosidase was added to reactions just before the addition of the transcription factors. In lanes 10-12, GAL-IIC(A519-K1157) was preincubated with the P11-0.4 M fraction for 10 min as described in Materials and Methods. (B) Quantitation of transcripts by scanning laser densitometry.

**A.**



**B.**

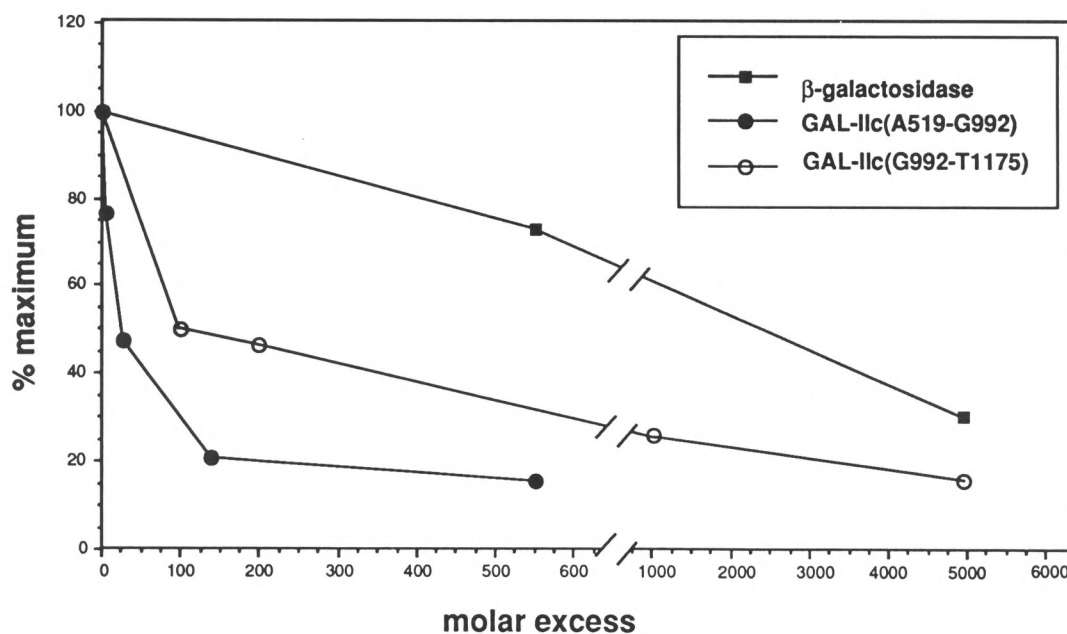


FIG. 4. Inhibition of reconstituted transcription by GAL-IIc(A519-G992) and GAL-IIc(G992-T1175). (A) Run-off transcription assays were conducted as described in Fig. 3. Transcriptions contained 10 units (4.8 fmol) RNA polymerase II, 0.1 μl RNase inhibitor, 0.5 μl factor 3, and 3 μl P11-0.4 M fraction. In these reactions, the P11-0.4 M fraction was preincubated for 5 min at 25°C with lane 1, no addition; lanes 2-3, β-galactosidase; lanes 4-7, GAL-IIc(A519-G992); lanes 8-11, GAL-IIc(G992-T1175). The actin run-off transcript is indicated by the arrow. (B) Quantitation of run-off transcript produced in (A).

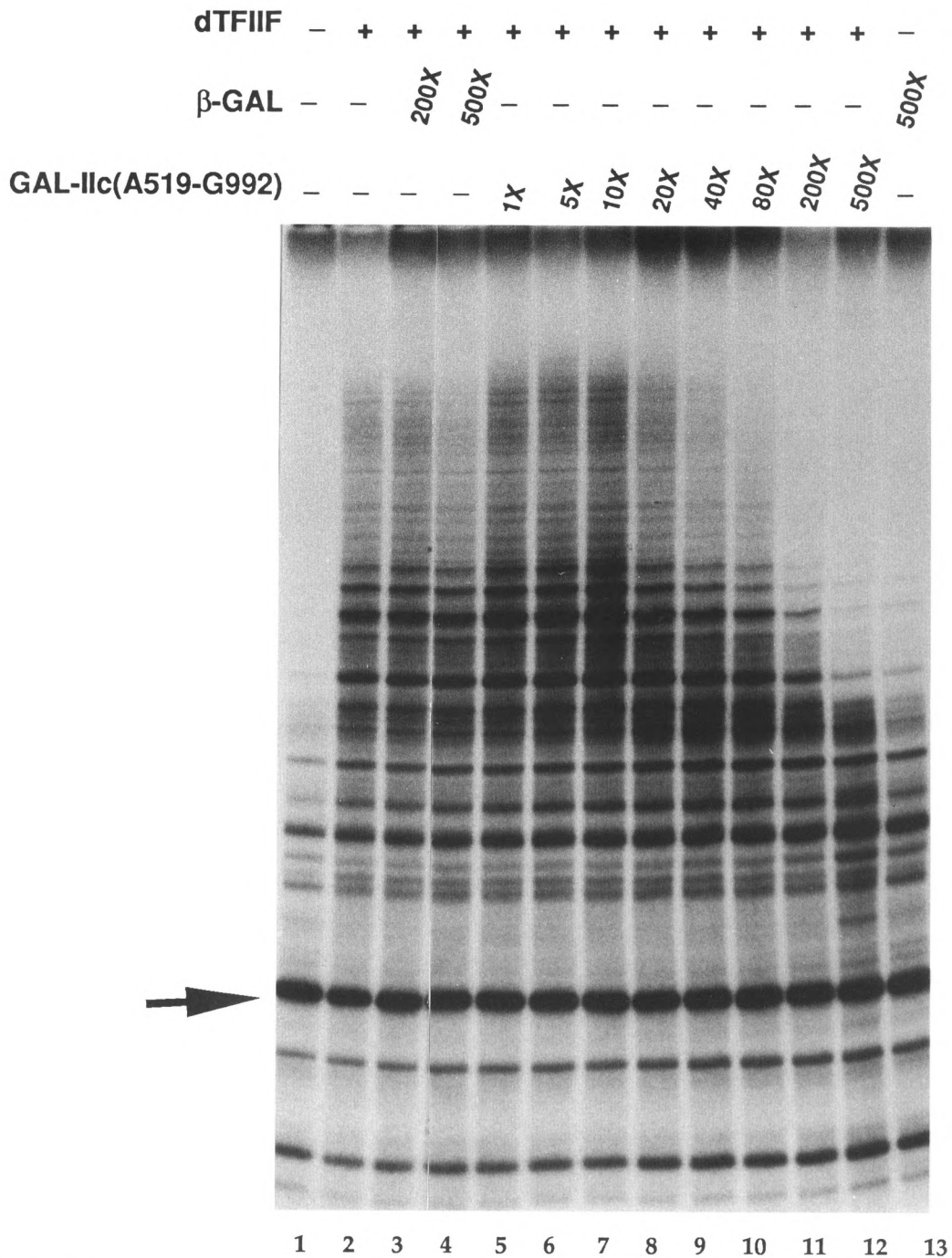


FIG. 5. Inhibition of dTFIIIF elongation stimulation during transcription of dC-tailed template. The final KCl concentration in each reaction was 80 mM. Each reaction contained 4.8 fmol Pol II (29 ng), 10  $\mu$ g/ml template dC-PCP, 140 ng MonoQ fraction DNase inhibitor, 5 units RNase inhibitor (Boehringer Mannheim), and 1.1  $\mu$ g  $\beta$ -galactosidase. The indicated reactions contained 100 ng of MonoQ-purified dTFIIIF. GAL-IIc(A529-G992) was preincubated with dTFIIIF for 5 min at 25°C. Following preincubations of Pol II with template and dTFIIIF with fusion protein, the master polymerase mixture was aliquotted into the individual reaction tubes, and transcripts were elongated for 3 min. Lane 1: Pol II only; lane 2: Pol II + dTFIIIF; lane 3-4: same as lane 2 with indicated amount of  $\beta$ -galactosidase in molar excess over Pol II; lanes 5-12: same as lane 2 with indicated amount of GAL-IIc(A519-G992); lane 13: Pol II +  $\beta$ -galactosidase. The arrow indicates the position of a 186 bp size and recovery standard added with the transcription stop solution. Note that the length of the longest transcripts in lanes 5-8 is ca. 1.5 kb.

compare lanes 12 and 13). In contrast to these results, GAL-IIC(G992-T1175) did not efficiently inhibit dTFIIF stimulation of RNA polymerase II. Greater than 500-fold excess of this fusion protein was required to see appreciable reduction in the level of transcription (data not shown).

*Inhibition of Promoter-Dependent Basal Transcription and dTFIIF-Mediated Stimulation of Elongation by GST-IIC Fusion Proteins*

An investigation of the physical association between transcription factors and RNA polymerase II is essential to a clear understanding of how gene expression is controlled. In particular, localization of *Drosophila* transcription dTFIIF binding to a particular RNA polymerase II subunit or subunit domain would provide some insight into the factor's mode of action and the role played by that subunit during the transcription cycle. We have demonstrated that the fragment of RNA polymerase subunit IIC consisting of amino acids 519–992 can inhibit dTFIIF elongation-stimulatory activity, thus providing strong, albeit indirect, evidence for a specific interaction between dTFIIF and this subunit domain. The following experiments were designed to obtain direct physical evidence for the binding of dTFIIF to the IIC(A519-G992) domain. The large, tetrameric structure of  $\beta$ -gal-IIC fusion proteins precluded their use in most types of binding experiments and, in part, motivated the construction of the GST-IIC fusion proteins.

The GST fusion proteins are summarized schematically in Fig. 1 (and see Materials and Methods). GST-IIC(A519-G992) contained the IIC fragment from GAL-IIC(A519-G992) cloned into the GST vector. In addition, smaller portions of IIC were constructed as GST fusions, to see if the interacting domain could be localized further. The conserved domains F, G, H, and I were targeted for these fusion proteins, due to the genetic and biochemical evidence, which suggests that these regions are critical for the enzyme's function (see Discussion).

The demonstration of equivalent dTFIIF inhibition by GST-IIC(A519-G992) and GAL-IIC(A519-G992) would provide persuasive evidence for the specificity of this effect. Indeed, we found that GST-IIC(A519-G992) and GAL-IIC(A519-G992) interfere similarly with Kc cell nuclear extract-mediated transcription from the actin promoter (not shown). Parallel inhibition by the two fusion proteins was also observed during dTFIIF-dependent stimulation of RNA polymerase II elongation. Figure 6 shows that comparable

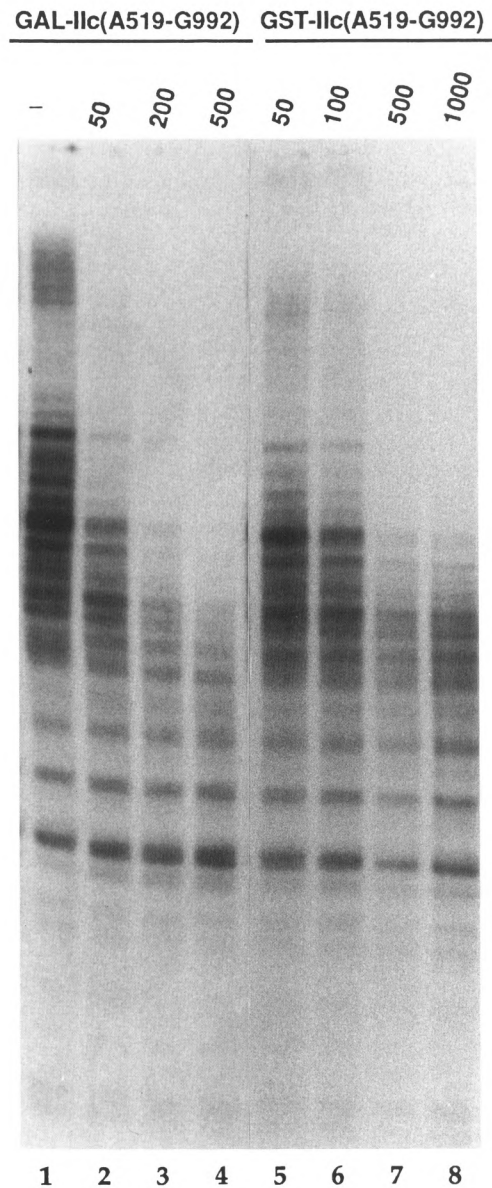


FIG. 6. Comparison of GAL-IIC(A519-G992) and GST-IIC(A519-G992) inhibition of dTFIIF-mediated elongation stimulation. The final KCl concentration was 80 mM; transcripts were elongated for 3 min. Each reaction contained 4.8 fmol Pol II (29 ng), 140 ng MonoQ fraction DNase inhibitor, 5 units RNase inhibitor (Boehringer Mannheim), and 20  $\mu$ g/ml BSA. All reactions contained 100 ng of MonoQ-purified dTFIIF. Lane 1: Pol II + dTFIIF; lanes 2–4: same as lane 1 plus the indicated excess of GAL-IIC(A519-G992); lanes 5–8: same as lane 1 plus the indicated excess of GST-IIC(A519-G992).

levels of GST-IIC(A519-G992) and GAL-IIC(A519-G992) caused similar reductions in the size and number of transcripts elongated. GST-IIC(A519-G992) did not interfere with elongation by RNA polymerase II itself. These data strongly suggest that the IIC(A519-G992) domain plays an

important role in the association of transcription dTFIIF with RNA polymerase II during initiation and elongation.

To determine whether the inhibitory IIC domain could be localized any further, we conducted preliminary tests for inhibition by small GST-IIC fusion proteins in KcN transcription and elongation assays. The results of these studies are summarized in Fig. 1C. None of the smaller fusion proteins significantly inhibited either KcN-mediated transcription from a promoter or dTFIIF-specific stimulation of RNA polymerase II elongation (data not shown). Overall, the inability of GST-IIC(G776-I859), GST-IIC(E901-Q1041), and GST-IIC(D825-G992) to interfere clearly with either KcN-dependent transcription or dTFIIF-mediated elongation stimulation provides evidence that a fairly large portion of the second largest subunit is necessary for the association with dTFIIF, either via a direct interaction or by indirectly stabilizing the interacting domain.

#### *Chemical Cross-Linking of dTFIIF to Iodinated GST-IIC(A519-G992)*

Protein-protein interactions involved in many processes have been studied by an increasingly diverse range of methods, including gel mobility shift assays (Williams et al., 1991), blot-overlay techniques (Homann et al., 1991; Wood et al., 1992), and affinity chromatography (Ing et al., 1992; Koleske et al., 1992). Of course, the intrinsic affinity of the associating proteins ultimately determines whether a complex can be detected by a particular assay method. In addition, the lifetime of association of a complex is affected by differences in reaction pH, salt concentration, and concentrations of the interacting proteins. Binding of mammalian dTFIIF analogue, RAP30/74, to immobilized RNA polymerase II is modest, with a reported dissociation constant of  $2 \times 10^{-8}$  M (Formosa et al., 1991). We had hoped that the association of dTFIIF with the GST-IIC fusion protein might be similarly detected by using immobilized fusion proteins; however, such efforts were unsuccessful (data not shown). We then sought an alternative method to study the association of dTFIIF with RNA polymerase II, namely chemical cross-linking with ethylene glycol bis-succinimidyl succinate (EGS).

Chemical cross-linking has been used to study the spatial relationships between components of complex structures such as virions (Mariani et al., 1990), cytochromes (Font et al., 1991), heat shock factor (Clos et al., 1990), or ribosomes (Peters

and Richards, 1977). In general, the spacing and geometry of the target moieties and the linear dimensions and geometry of the reagent used are critically important for the establishment of a successful cross-link. Naturally, when intermolecular protein-protein interactions are being studied, the concentrations of the reacting species must be sufficiently high for the complexes to be detected. Ideally, the amount of reagent used should be enough to cross-link the proteins of interest without modifying every available reactive group. Although a single modification is unlikely to cause drastic changes, it is possible that modification of several or key hyperreactive groups may markedly affect the protein's physical properties or biological activity. The contribution of most small cross-linking agents to molecular weight differences observed between species is negligible; however, possible changes in hydrodynamic properties of the cross-linked species must not be overlooked.

Radiolabeled GST-IIC(A519-G992) was incubated either alone or with an excess of Mono Q-purified dTFIIF for 1 h before cross-linking of the proteins with EGS. Figure 7 shows that three protein bands, with apparent molecular weights of approximately 190, 170, and 120, had cross-linked to labeled GST-IIC(A519-G992) in the presence of added dTFIIF, but not in its absence. This effect was reproduced in three separate experiments (unpublished results). Factor dTFIIF (factor 5) has subunits of 86 kDa (dTFIIFa or factor 5a) and 34 kDa (dTFIIFb or factor 5b), and a total mass of 120 kDa, whereas GST-IIC(A519-G992) has an estimated mass of 86 kDa also. Although this experiment does not show directly that dTFIIF subunits are present in these cross-linked species, the band sizes are consistent with fusion protein cross-linking to dTFIIF(a + b), dTFIIFa (factor 5a) only, or dTFIIFb (factor 5b) only, respectively. In a separate experiment using unlabeled fusion protein, antibodies specific for the IIC(A519-G992) domain detected the 190-kDa band, confirming the presence of GST-IIC(A519-G992) in the cross-linked complex with dTFIIF (data not shown). The smaller cross-linked bands were not detected by chromogenic detection of the antibodies, probably due to the limited sensitivity of this method.

#### *Chemical Cross-Linking of dTFIIF to RNA Polymerase II*

dTFIIF bound stably to *Drosophila* RNA polymerase II during glycerol gradient sedimentation (Price et al., 1989). Therefore, a complex of

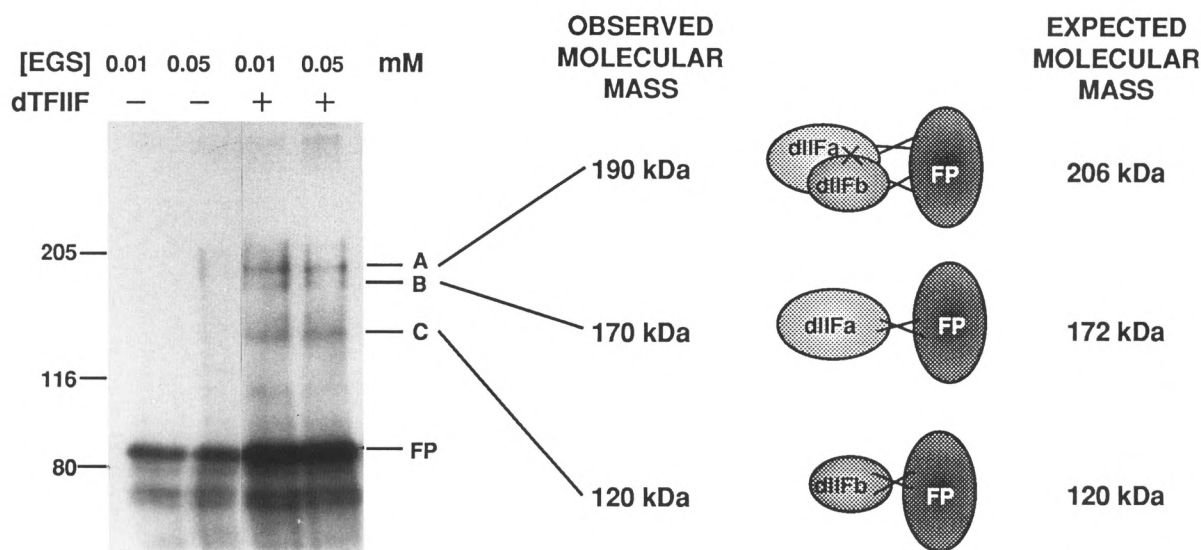


FIG. 7. Chemical cross-linking of iodinated GST-IIC(A519-G992) to dTFIIF. Each 15- $\mu$ l reaction contained 5  $\mu$ l [ $^{125}$ I]GST-IIC(A519-G992) [ $5.6 \times 10^5$  cpm] and either no factor (lanes 1 and 2) or 1.5  $\mu$ g dTFIIF fraction Q15 (lanes 3 and 4) at a final salt concentration of 100 mM KCl. The reactions were incubated for 1 h at 25°C; cross-linking with EGS proceeded for 15 min and was then quenched by the addition of 30 mM lysine. Electrophoresis sample buffer was added (from 6 $\times$  stock), the samples were heated to 100°C for 5 min, and then loaded on a 7.5% polyacrylamide-SDS gel. The gel was blotted to nitrocellulose for 2 h at 0.3 amps, air dried, then exposed to X-ray film. Reactions contained 0.01 mM EGS (lanes 1 and 3) or 0.05 mM EGS (lanes 2 and 4). The positions of molecular weight markers are indicated at the left. The observed and expected molecular masses of GST-IIC(A519-G992) cross-linked to one or both subunits of dTFIIF are indicated at the right.

dTFIIF and intact RNA polymerase II should be detectable by chemical cross-linking. RNA polymerase II and an excess of TFIIF (two different preparations) were incubated together for 15 min, cross-linked with EGS, and subjected to SDS-PAGE. Identical Western blots of these gels were then incubated with antibodies to exon 2 of the *Drosophila* RNA polymerase II 215-kDa subunit IIa (Fig. 8A) or with antibodies to the IIC(A519-G992) domain of the 140-kDa subunit (Fig. 8B). Two major and several minor bands were detected only by antibodies to the 140-kDa subunit and not by antibodies to the 215-kDa subunit. As was shown for dTFIIF cross-linking to the fusion protein, the estimated sizes of some of the bands detected by anti-IIC antibodies are consistent with the sizes predicted for cross-linking of one or both dTFIIF subunits to RNA polymerase IIC [dTFIIF + IIC = 260 kDa (bands "A"); dTFIIFa + IIC = 226 kDa (band "B1"); dTFIIFb + IIC = 174 kDa (bands "C")]. Note that the intensity of bands C is  $\gg$  that of B1, an observation that is consistent with dTFIIFb (the RAP 30 homologue) directly contacting RNA polymerase II (see the Introduction).

Although some of the cross-linking may have been due to specific or nonspecific binding to other proteins in the dTFIIF fractions, some, if not all, of the bands are due to specific cross-

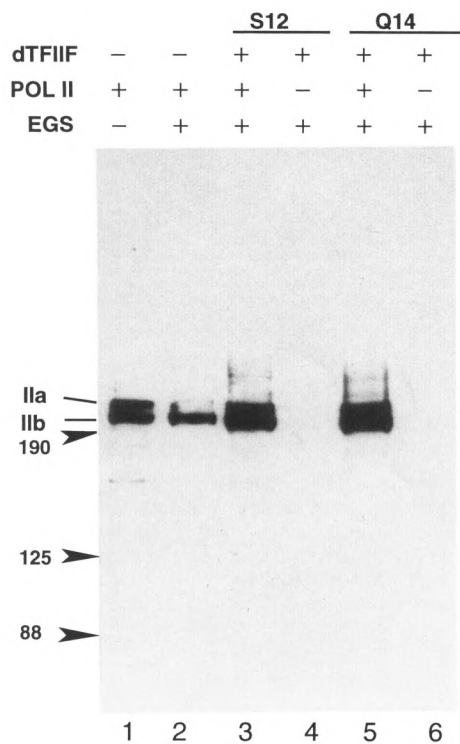
linking of the 140-kDa subunit to dTFIIF. First, two different dTFIIF preparations gave virtually identical results. Also, when an equivalent amount of BSA was incubated with RNA polymerase II in the presence of EGS, no BSA-dependent cross-linking was detected by either large subunit antibody (data not shown). Furthermore, that anti-IIC(A519-G992) antibodies affinity purified from two different goats, one injected with purified native RNA polymerase II and one injected with bacterially expressed IIC(A519-G992), detected the same pattern of cross-linked bands shows that the effect was not due to nonspecific cross-reaction of antibodies to proteins in the dTFIIF fraction (data not shown). Taken together, all of these results point to the specific cross-linking of dTFIIF to subunit IIC.

## DISCUSSION

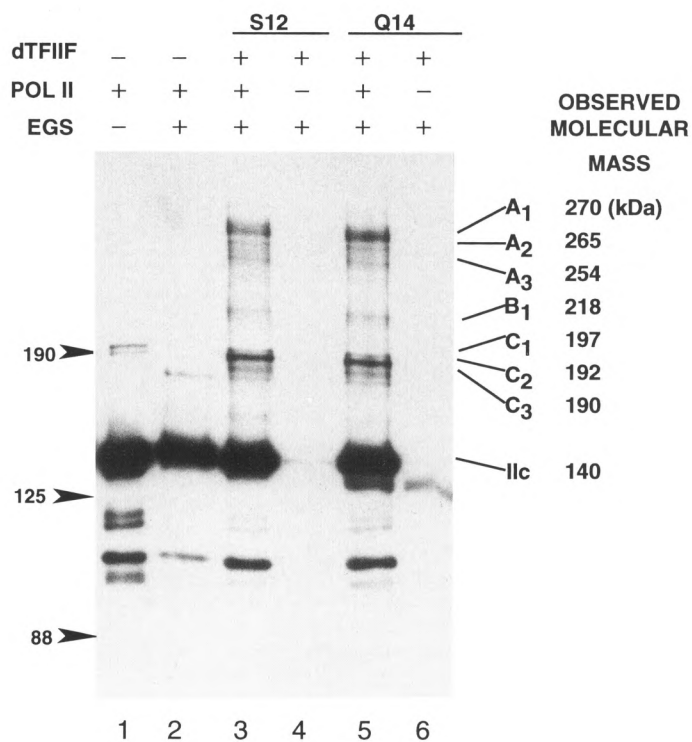
### *Domain Interference and the Role of dTFIIF During Transcription*

The elongation efficiency of both prokaryotic and eukaryotic RNA polymerases is limited significantly by the amount of time they spend at pause sites in the DNA. Previous work from this lab and others has shown that RNA polymerase II pauses frequently during transcription on dC-tailed tem-

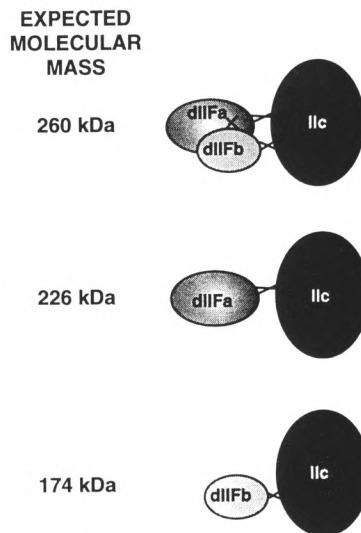
**A.  $\alpha$ -IIa**



**B.  $\alpha$ -IIc(A519-G992)**



**C.**





plates, with a very strong, apparently sequence-independent pause occurring at base 13 (formerly called the 14-mer) and many sequence-dependent pauses occurring downstream (Kadesh and Chamberlin, 1982; Coulter and Greenleaf, 1985; Gao and Price, 1993; Reines, 1992; Sluder et al., 1989). Unlike factor DmS-II, which can suppress pausing, dTFIIF does not alter the ability of RNA polymerase II to recognize a particular site as a pause; it simply reduces the amount of time Pol II spends at the pause. Sluder et al. (1989) also found that dTFIIF increased the elongation rate of RNA polymerase II through the 13-mer region, and demonstrated the ability of dTFIIF to increase the total number of polymerase molecules productively engaged in transcription on dC-tailed templates (reproduced here in, for example, Fig. 5A, lanes 1 and 2). They also discovered that dTFIIF reduced the residence time of Pol II at multiple sites downstream of the 13-mer, causing a three- to fivefold stimulation of the overall elongation rate (Price et al., 1989; Sluder, 1988). However, dilution experiments and size fractionation of actively transcribing complexes have shown that dTFIIF is not stably bound to RNA polymerase II during elongation on dC-tailed templates (Price et al., 1989). Therefore, under these conditions dTFIIF apparently interacts independently with polymerase each time it pauses, and an excess of dTFIIF over RNA polymerase II is necessary to ensure that this encounter is not rate limiting.

In the studies reported here, we observed inhibition of dTFIIF-stimulated transcription of dC-tailed templates by fusion proteins carrying portions of RNA polymerase subunit IIC. This observation is consistent with the proposals that the described effects of dTFIIF on elongation occur by virtue of its binding to a site on RNA polymerase that includes portions of the IIC subunit, and that excess fusion protein present in the reaction titrates dTFIIF, diminishing its productive interactions with the transcribing polymerase. We

note that the effective fusion proteins, GAL-IIC(A519-K1157), GAL-IIC(A519-G992), and GST-IIC(A519-G992), did not significantly change the relative strength or pattern of pause sites observed with RNA polymerase II alone or with added dTFIIF during elongation from either the dC-Ball-E or dC-PCP templates (Figs. 5 and 6B, and data not shown). Thus, the fusion proteins apparently exerted no additional effects on the elongation reaction other than to prevent dTFIIF from productively engaging the polymerase. In addition, that GAL-IIC(A519-K1157), GAL-IIC(A519-G992), and GST-IIC(A519-G992) most dramatically affected the production of long transcripts in these dC assays illustrates that the inhibition by these three fusion proteins is neither instantaneous nor irreversible.

That large molar excesses of these fusion proteins were required to produce a clear reduction in elongation rate is consistent with the notion that dTFIIF acts via a transient interaction with this Pol II domain. Similarly, the inability to show association between dTFIIF and the IIC fusion proteins in the absence of cross-linker may also reflect the dynamic nature of this interaction. It may also be that part of the necessity for excess fusion proteins can be attributed to only a fraction of the fusion protein being folded correctly. All of the fusion proteins we generated were insoluble as expressed in *E. coli*, and obtaining soluble protein necessitated denaturation, renaturation, and dialysis into appropriate buffers. As one approximation of correct refolding, we monitored the specific activity of  $\beta$ -gal in the GAL-IIC constructs. However, it is not possible to know how accurately this reflects the refolding of the IIC portion of the fusions.

The domain interference results are consistent with the findings of Price and coworkers (Kephart et al., 1992), who showed that three agents known to disrupt protein-protein interactions inhibited the activity of dTFIIF during transcription from

#### FACING PAGE

FIG. 8. Cross-linking of RNA polymerase II to dTFIIF. Identical reactions containing 23.5 ng RNA polymerase II were incubated alone or with one of two different preparations of dTFIIF, fraction S12 (500 ng protein of a MonoS fraction) or fraction Q14 (1 mg protein of a MonoQ fraction) (Skantar, 1993, Fig. A-4), and cross-linked with 0.02 mM EGS as described in the legend to Fig. 7. Identical nitrocellulose blots were blocked and probed with goat antibodies affinity purified to the second exon of RNA polymerase subunit IIA (anti-exon 2) (Lee and Greenleaf, 1991; Weeks et al., 1993) or to IIC(A519-G992). The blots were incubated with a 1:5000 dilution of primary antibody overnight at 4°C, then for 1 h with 1:10,000 horseradish peroxidase-conjugated swine anti-goat IgG. The blots were developed using ECL chemiluminescent antibody detection system (Amersham). (A) Blot probed with goat anti-exon 2 (directed against exon 2 of the 215-kDa subunit IIA). (B) Blot probed with anti-IIC(A519-G992). To the left of each figure, the positions of molecular size markers are indicated. In (A), the positions of 215-kDa subunit including the carboxy-terminal domain (IIa) or missing this feature (IIb) are also indicated. In (B), the cross-linked species that appear only in the presence of the dTFIIF fractions and Pol II are grouped A, B, and C, along with their observed molecular masses. The molecular composition and sizes expected for cross-linking of dTFIIF subunits to IIC are illustrated in (C). Note that the 260-kDa species may result from individual cross-linking of the dTFIIF subunits to IIC, or by cross-linking of dTFIIF subunits to each other with only one being cross-linked directly to IIC.

dC-tailed templates. In the presence of 250 mM KCl, 0.3% Sarkosyl, or 1 mg/ml heparin, dTFIIF stimulation of RNA polymerase II elongation rate was nearly or completely eliminated. That disruptive agents or an excess of dTFIIF binding sites can inhibit dTFIIF activity to a similar degree confirms that this factor is not tightly bound to RNA polymerase II during elongation as measured here in vitro.

The transcription inhibition data strongly suggest that dTFIIF makes direct contact with the second largest subunit of RNA polymerase II. On the other hand, our attempts to demonstrate binding between fusion protein GST-IIC(A519-G992) and dTFIIF in the absence of cross-linking agents were not successful (data not shown). The dissociation constant measured for RAP30/74 binding to immobilized RNA polymerase II is  $2 \times 10^{-8}$  M (Formosa et al., 1991), which is probably a reasonable approximation of dTFIIF binding as well. However, one might predict that the absence of certain IIC domains or additional RNA polymerase subunits that stabilize dTFIIF binding could increase the dissociation constant, as could altered folding or stability of the fusion proteins. Therefore, we propose that the binding of dTFIIF to the fusion proteins is likely much weaker than to native polymerase, rendering it undetectable in the absence of cross-linking agents.

The cross-linking of dTFIIF to [<sup>125</sup>I]GST-IIC(A519-G992) (Fig. 7) or to subunit IIC of native *Drosophila* RNA polymerase II (Fig. 8) adds some physical evidence to support a direct involvement of this subunit in the modulation of RNA polymerase II activity by dTFIIF. As shown in the Results, some of the shifted bands containing fusion protein or subunit IIC had apparent molecular masses in SDS-PAGE gels consistent with cross-linking to one or both dTFIIF subunits. Due to the complexity of cross-linking reactions, we cannot say whether both dTFIIF subunits bound directly to IIC domains or whether only one dTFIIF subunit was directly attached. Both possibilities are illustrated in the schematic representations of expected cross-linking products (Figs. 7 and 9C). However, others have shown that mammalian RAP30 (the dTFIIFb analogue) can bind to RNA polymerase II in the absence of RAP74 (the dTFIIFa analogue) (Killeen and Greenblatt, 1992). Therefore, it is possible that dTFIIFb bound directly to and was cross-linked to subunit IIC, and that dTFIIFa cross-linked to dTFIIFb. dTFIIFa may also cross-link directly to IIC if its attachment to dTFIIFb brings it into proximity with IIC, such that cross-linking can take place. Interestingly,  $\sigma$  factor and the  $\beta$  subunit of *E. coli*

RNA polymerase cross-link to the *lac* UV5 promoter only five nucleotides apart (Paddon and Hartley, 1987). Given the apparent sequence and functional similarity between RAP30 and  $\sigma 70$  (McCracken and Greenblatt, 1991), and between *Drosophila* Pol IIC and  $\beta$  (Falkenburg et al., 1987), it is perhaps not surprising that dTFIIF was cross-linked to IIC or that the IIC fusion proteins interfered with dTFIIF function.

#### *Defining Functionally Significant Regions of RNA Polymerase II*

Table 1 summarizes critically important amino acid residues or domains defined by genetic manipulations or functional analyses of second largest subunits from several different organisms, focusing on mutations that overlap the *Drosophila* IIC sequences that inhibited dTFIIF in this study. Interestingly, mutations in many of these sites affected the propagation of RNA polymerase (i.e., promoter clearance and pausing, Sagitov et al., 1993), or interfered with interactions of transcription factors GreB (Borukhov et al., 1993) or  $\sigma$  (Glass et al., 1988). Other mutations have been localized near the enzyme's catalytic center (Scafe et al., 1990a; Chen et al., 1993; Kashlev et al., 1990; Lee et al., 1991). Additionally, alterations in *Drosophila RpIII40* or yeast *RPB2* suppress mutations in the large subunit of RNA polymerase II, thus implicating domains E and I in subunit-subunit contacts (Table 1).

The inability of the smaller fusion proteins GST-IIC(G776-I859), GST-IIC(D825-G992), or GST-IIC(E901-I041) to interfere with dTFIIF activity in the elongation assay as effectively as GST-IIC(A519-G992) suggests that much of the region encompassed by residues 519-992 is important for the interaction of dTFIIF with IIC (Fig. 7). However, at this time, some combinations of domains D-I have not been investigated, and we therefore do not know whether certain other fusion proteins would inhibit dTFIIF as effectively as GST-IIC(A519-G992). Because conserved domain I has been implicated in functions such as  $\sigma$  factor binding (Glass et al., 1988), it remains possible that this C-terminal domain plays a role, possibly an indirect one, in the dTFIIF interaction with IIC.

#### *Potential Future Applications of the Domain Interference Assay*

Given that mammalian TFIIF (RAP30/74) can be dissociated by salt from immobilized RNA polymerase II columns (Sopta et al., 1985), and that dTFIIF stimulation of elongation on dC-

TABLE 1  
FUNCTIONALLY DEFINED RESIDUES OF SUBUNIT IIC HOMOLOGUES IN DOMAINS THAT OVERLAP  
THE dTFIIF-INTERACTING FUSION PROTEINS

Species	Amino Acids(s)	Domain	Defect or Function	Reference
<i>E. coli</i> -RPOB	813 (G → K)	F	decreased promoter clearance; abortive initiation & pausing	Lee et al., 1991
	several	F & I	altered termination and pausing	Landick et al., 1990a
	1065 (K → R)	H	near priming nucleotide; fails to elongate properly	Kashlev et al., 1990
	D1064-K1073	H	polymerase propagation	Sagitov et al., 1993
<i>S. cerevisiae</i> -RPB2	Δ889-1342	I	fails to bind $\sigma$	Glass et al., 1988
	913 (G → D)	G	near proposed purine binding site	Scafe et al., 1990b
	K936	H	active site labelling (priming nucleotide)	Treich et al., 1992
	1142 (G → D)	I	gene-specific transcription defects	Scafe et al., 1990a
	1145 (S → L)	I	suppress <i>rpb1-1</i> mutation	Martin et al., 1990
	1136 (D → N)	I	suppress <i>rpb1-1</i> mutation	Martin et al., 1990
<i>Drosophila</i> -RpII140	S1, 728 (S → C)	E	suppress <i>RpII215<sup>WJK1</sup></i>	Mortin, 1990, and personal communication
	S7, 735 (M → I)	E	suppress <i>RpII215<sup>WJK1</sup></i>	Mortin, 1990, and personal communication
	Z19, 1028 (E → K)	I	recessive lethal	Chen et al., 1993
	A5, Δ1047-1051	I	recessive lethal	Chen et al., 1993
	Z43, 940 (R → H)	H	enhance <i>Ubx</i> phenotype	Chen et al., 1993
	Z45, 1098 (S → F)	I	recessive lethal	Chen et al., 1993
	M39, 992 (G → E)	H/I	recessive lethal	Chen et al., 1993

tailed templates is inhibited by 0.3% Sarkosyl, 1 mg/ml heparin, or 250 mM salt (Kephart et al., 1993), it seems likely that the binding of TFIIF to RNA polymerase II involves electrostatic interactions. Intriguingly, long stretches of charged residues have been located in RAP74 (Finkelstein et al., 1992; Kephart et al., 1993), and this TFIIF subunit is also believed to be phosphorylated in vivo (Sopta et al., 1985). It will be important to know what role if any these properties play in the association of the factor with the enzyme, and preliminary functional dissections of this TFIIF subunit have been carried out (Kephart et al., 1994, and refs. therein). When both dTFIIF subunit genes are cloned, it should be possible to exploit their availability in combination with the IIC constructs and approaches we have described in this paper to carry out more detailed investigations of interactions between dTFIIF and Pol II.

Several groups have examined the entry of basal transcription factors and RNA polymerase II into initiation complexes (Buratowski et al., 1989; Killeen et al., 1992; Lu et al., 1992). As pointed out recently, in order to understand the architecture of the initiation complex it will be important to map the factor-polymerase interactions to specific Pol II subunits (Buratowski, 1994). The domain interference technique could be quite useful in helping to determine critical protein contacts within the initiation complex. In addition, Pol II subunit domain fusion proteins

might be used as probes to determine whether transcriptional regulatory proteins act via interactions with specific polymerase II subunit domains.

Finally, application of the domain interference assay to the study of existing mutant *Drosophila* RNA polymerase II subunits might complement genetic studies of these altered enzymes. Due to the homozygous lethality of most known *RpII140* mutations, these mutant enzymes are not available for direct biochemical analysis (Chen et al., 1993, and refs. therein). However, it would be possible to express as fusion proteins subunit domains containing characterized mutations; these could then be analyzed by the techniques described in this work. For example, interactions between mutant domains and transcription factors could be examined, potentially pinpointing defective protein-protein interactions responsible for lethality in vivo. Additional point mutations, deletions, or sequence substitutions could be introduced into critical regions by in vitro mutagenesis, providing an array of altered subunit domains whose properties could then be examined. Ultimately such a combination of in vivo and in vitro approaches will contribute to an increased understanding of factor-polymerase interactions and of mechanisms involved in transcribing genetic information.

#### ACKNOWLEDGEMENTS

We express our gratitude to Deborah Danaher, Muh-Lin Lee, and Meei Liu for expert technical

assistance; and to John Weeks, David Price, and Steven Hardin for helpful discussions. This work

was supported by National Institutes of Health Grant GM 28078.

## REFERENCES

- B. Bartholomew, M. E. Dahmus, and C. F. Meares (1986), *J Biol Chem* 261, 14226-14231.
- A. E. Bolton and W. M. Hunter (1973), *Biochem J* 133, 529-539.
- S. Borukhov, V. Sagitov, and A. Goldfarb (1993), *Cell* 72, 459-466.
- B. Breant, J. Huet, A. Sentenac, and P. Fromageot (1983), *J Biol Chem* 258, 11968-11973.
- S. Buratowski (1994), *Cell* 77, 1-3.
- S. Buratowski, S. Hahn, L. Guarente, and P. A. Sharp (1989), *Cell* 56, 549-561.
- Y. Chen, J. Weeks, M. A. Mortin, and A. L. Greenleaf (1993), *Mol Cell Biol* 13, 4214-4222.
- J. Clos, J. T. Westwood, P. B. Becker, S. Wilson, K. Lambert, and C. Wu (1990), *Cell* 63, 1085-1097.
- J. W. Conaway and R. C. Conaway (1990), *Science* 248, 1550-1553.
- D. E. Coulter and A. L. Greenleaf (1985), *J Biol Chem* 260, 13190-13198.
- S. Dissinger and M. M. Hanna (1990), *J Biol Chem* 265, 7662-7668.
- D. Falkenburg, B. Dworniczak, D. M. Faust, and E. K. Bautz (1987), *J Mol Biol* 195, 929-937.
- A. Finkelstein, C. F. Kostrub, J. Li, D. P. Chavez, B. Q. Wang, S. M. Fang, J. Greenblatt, and Z. F. Burton (1992), *Nature* 355, 464-467.
- O. Flores, H. Lu, M. Killeen, J. Greenblatt, Z. F. Burton, and D. Reinberg (1991), *Proc Natl Acad Sci USA* 88, 9999-10003.
- B. Font, D. Goldschmidt, J. F. Chich, M. Thieffry, J. P. Henry, and D. C. Gautheron (1991), *FEBS Lett* 279, 105-109.
- T. Formosa, J. Barry, B. M. Alberts, and J. Greenblatt (1991), *Methods Enzymol* 208, 24-45.
- S. Frankel, R. Sohn, and L. Leinwand (1991), *Proc Natl Acad Sci USA* 88, 1192-1196.
- H. Gao and D. H. Price (1993), *J Biol Chem* 268, 18762-18770.
- R. E. Glass, A. Honda, and A. Ishihama (1986a), *Mol Gen Genet* 203, 492-495.
- R. E. Glass, S. T. Jones, V. Nene, T. Nomura, N. Fujita, and A. Ishihama (1986b), *Mol Gen Genet* 203, 487-491.
- R. E. Glass, N. T. Ralphs, N. Fujita, and A. Ishihama (1988), *Eur J Biochem* 176, 403-407.
- M. A. Grachev, T. I. Kolocheva, E. A. Lukhtanov, and A. A. Mustaev (1987), *FEBS Lett* 103, 113-121.
- M. A. Grachev, E. A. Lukhtanov, A. A. Mustaev, E. F. Zaychikov, M. N. Abdukayumov, I. V. Rabinov, V. I. Richter, Y. S. Skoblov, and P. G. Chistyakov (1989), *Eur J Biochem* 180, 577-585.
- A. L. Greenleaf (1992), in *Transcriptional Regulation*, vol. 1 (K. Yamamoto and S. McKnight, eds.), Cold Spring Harbor Press, Cold Spring Harbor, NY, pp. 55-80.
- E. D. Gundelfinger (1983), *FEBS Lett* 157, 133-138.
- D. A. Hager and R. R. Burgess (1980), *Anal Biochem* 109, 76-86.
- B. J. Hamilton, M. A. Mortin, and A. L. Greenleaf (1993), *Genetics* 134, 517-529.
- H. E. Homann, W. Willenbrink, C. J. Buchholz, and W. J. Neubert (1991), *J Virol* 65, 1304-1309.
- N. H. Ing, J. M. Beekman, S. Y. Tsai, M.-J. Tsai, and B. W. O'Malley (1992), *J Biol Chem* 267, 17617-17623.
- T. R. Kadesh and M. J. Chamberlin (1982), *J Biol Chem* 257, 5286-5295.
- M. Kashlev, J. Lee, K. Zalenskaya, V. Nikiforov, and A. Goldfarb (1990), *Science* 248, 1006-1009.
- D. D. Kephart, N. F. Marshall, and D.H. Price (1992), *Mol Cell Biol* 12, 2067-2077.
- D. D. Kephart, M. P. Price, Z. F. Burton, A. Finkelstein, J. Greenblatt, and D. H. Price (1993), *Nucleic Acids Res* 21, 1319.
- D. D. Kephart, B. Q. Wang, Z. F. Burton, and D. H. Price (1994), *J Biol Chem* 269, 13536-13543.
- M. T. Killeen, B. Coulombe, and J. Greenblatt (1992), *J Biol Chem* 267, 9463-9466.
- M. T. Killeen and J. F. Greenblatt (1992), *Mol Cell Biol* 12, 30-37.
- A. J. Koleske, S. Buratowski, M. Nonet, and R. A. Young (1992), *Cell* 69, 883-894.
- P. A. Kolodziej and R. A. Young (1991), *Mol Cell Biol* 11, 4669-4678.
- U. K. Laemmli (1970), *Nature* 227, 680-685.
- R. Landick, A. Colwell, and J. Stewart (1990a), *J Bacteriol* 172, 2844-2854.
- R. Landick, J. Stewart, and D. N. Lee (1990b), *Genes Dev* 4, 1623-1636.
- J. Lee, M. Kashlev, S. Borukhov, and A. Goldfarb (1991), *Proc Natl Acad Sci USA* 88, 6018-6022.
- J. M. Lee and A. L. Greenleaf (1991), *Gene Expr* 1, 149-167.
- K.-H. Lin and S. Y. Cheng (1991), *Biotechniques* 11, 748-751.
- H. Lu, L. Zawel, L. Fisher, J. M. Egly, and D. Reinberg (1992), *Nature* 358, 641-645.
- C. Mariani, M. De Beuckeleer, J. Truettner, J. Leemans, and R. B. Goldberg (1990), *Nature* 347, 737-741.
- C. Martin, S. Okamura, and R. Young (1990), *Mol Cell Biol* 10, 1908-1914.
- S. McCracken and J. Greenblatt (1991), *Science* 253, 900-902.
- J. H. Miller (1972), in *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 398-404.
- M. Mortin (1990), *Proc Natl Acad Sci USA* 87, 4864-4868.
- M. Mortin, R. Zuerner, S. Berger, and B. J. Hamilton (1992), *Genetics* 131, 895-903.

- A. Mustaev, M. Kashlev, J. Lee, A. Polyakov, A. Lebedev, K. Zalenskaya, M. Grachev, A. Goldfarb, and V. Nikiforov (1991), *J Biol Chem* 266, 23927-23931.
- V. Nene and R. Glass (1984), *Mol Gen Genet* 194, 166-172.
- C. J. Paddon and R. W. Hartley (1987), *Gene* 53, 11-19.
- B. Palenik (1992), *Curr Opin Genet Dev* 2, 931-936.
- K. Peters and F. M. Richards (1977), *Annu Rev Biochem* 46, 523-551.
- D. H. Price and C. S. Parker (1984), *Cell* 38, 423-429.
- D. H. Price, A. E. Sluder, and A. L. Greenleaf (1987), *J Biol Chem* 262, 3244-3255.
- D. H. Price, A. E. Sluder, and A. L. Greenleaf (1989), *Mol Cell Biol* 9, 1465-1475.
- J. Rappaport, K. Cho, A. Saltzman, J. Prenger, M. Golomb, and R. Weinmann (1988), *Mol Cell Biol* 8, 3136-3142.
- D. Reines (1992), *J Biol Chem* 267, 3795-3800.
- M. Riva, A. R. Schaffner, A. Sentenac, G. R. Hartmann, A. A. Mustaev, E. F. Zaychikov, and M. A. Grachev (1987), *J Biol Chem* 262, 14377-14380.
- A. Robbins, W. S. Dynan, A. L. Greenleaf, and R. Tjian (1984), *J Mol Appl Genet* 2, 343-353.
- V. Sagitov, V. Nikiforov, and A. Goldfarb (1993), *J Biol Chem* 268, 2195-2202.
- M. Sawadogo and A. Sentenac (1990), *Annu Rev Biochem* 59, 711-754.
- C. Scafe, C. Martin, M. Nonet, S. Podos, S. Okamura, and R. A. Young (1990a), *Mol Cell Biol* 10, 1270-1275.
- C. Scafe, M. Nonet, and R. A. Young (1990b), *Mol Cell Biol* 10, 1010-1016.
- A. M. Skantar (1993), PhD Thesis, Duke University.
- A. E. Sluder (1988), PhD Thesis, Duke University.
- A. E. Sluder, D. H. Price and A. L. Greenleaf (1987), *Biochimie* 69, 1199-1205.
- A. E. Sluder, D. H. Price, and A. L. Greenleaf (1988), *J Biol Chem* 263, 9917-9925.
- A. E. Sluder, A. L. Greenleaf, and D. H. Price (1989), *J Biol Chem* 264, 8963-8969.
- D. B. Smith and K. S. Johnson (1988), *Gene* 67, 31-40.
- M. Sopta, R. W. Carthew, and J. Greenblatt (1985), *J Biol Chem* 260, 10353-10360.
- M. Sopta, Z. F. Burton, and J. Greenblatt (1989), *Nature* 341, 410-414.
- H. Towbin, T. Staehelin, and J. Gordon (1979), *Proc Natl Acad Sci USA* 76, 4350-4354.
- J. R. Weeks, D. E. Coulter, and A. L. Greenleaf (1982), *J Biol Chem* 257, 5884-5892.
- J. R. Weeks, S. E. Hardin, J. Shen, J. M. Lee, and A. L. Greenleaf (1993), *Genes Dev* 7, 2329-2344.
- F. E. Williams, U. Varanasi, and R. J. Trumbly (1991), *Mol Cell Biol* 11, 3307-3316.
- E. R. Wood, O. B. McDonald, and N. Sahyoun (1992), *J Biol Chem* 267, 14138-14144.
- M. Yonaha, T. Aso, Y. Kobayashi, H. Vasavada, Y. Yasukochi, S. M. Weissman, and S. Kitajima (1993), *Nucleic Acids Res* 21, 273-279.
- R. A. Young (1991), *Annu Rev Biochem* 60, 689-715.
- L. Zawel and D. Reinberg (1993), *Prog Nucleic Acid Res Mol Biol*, 67-108.