

## A Key Role for the $\alpha 1$ Helix of Human RAP74 in the Initiation and Elongation of RNA Chains\*

Received for publication, June 23, 2002, and in revised form, September 26, 2002  
Published, JBC Papers in Press, September 26, 2002, DOI 10.1074/jbc.M206249200

Janel D. Funk‡, Yuri A. Nedialkov, Dianpeng Xu, and Zachary F. Burton§

From the Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824-1319

RNA polymerase II-associating protein 74 (RAP74) is the large subunit of transcription factor IIF (TFIIF), which is essential for accurate initiation and stimulates elongation by RNA polymerase II. Mutations within or adjacent to the  $\alpha 1$  helix of the RAP74 subunit have been shown to decrease both initiation and elongation stimulation activities without strongly affecting the interactions of RAP74 with the RAP30 subunit or the interaction between TFIIF and RNA polymerase II. In this manuscript, mutations within the  $\alpha 1$  helix are compared with mutations made throughout the neighboring conserved N-terminal domain of RAP74. Changes within the N-terminal domain include disruptions of specific contacts with the  $\alpha 1$  helix, which were revealed in the recently published x-ray crystal structure (Gaiser, F., Tan, S., and Richmond, T. J. (2000) *J. Mol. Biol.* 302, 1119–1127). Contacts between the  $\beta 4$ - $\beta 5$  loop and the  $\alpha 1$  helix are shown to be largely unimportant for  $\alpha 1$  helix function. Other mutations throughout the N-terminal domain are consistent with the establishment of the dimer interface with the RAP30 subunit. The RAP74-RAP30 interface is important for TFIIF function, but no particular RAP74 amino acids within this region have been identified that are required for TFIIF activities. The molecular target of the  $\alpha 1$  helix remains unknown, but our studies refocus attention on this important functional motif of TFIIF.

TFIIF<sup>1</sup> appears to be an  $\alpha\beta$  heterodimer of RAP74 and RAP30 subunits, and previous reports that TFIIF may be an  $\alpha_2\beta_2$  heterotetramer are not supported by the x-ray crystal structure (1). Although not evident from primary sequence, RAP74 and RAP30 subunits are structurally similar, with an intricate series of N-terminal  $\beta$ -sheets that form a RAP74-RAP30 dimer interface. RAP74 and RAP30 also have similar C-terminal regions with winged helix-turn-helix structures (2, 3). The larger size of the human RAP74 subunit can be attributed to an extensive loop rich in Gly, Pro, Ser, Thr, and charged residues separating more structured N- and C-terminal domains (4, 5).

TFIIF is an RNA polymerase II-specific transcription factor restricted to the eukaryotic kingdom. Fig. 1, therefore, shows

an amino acid sequence alignment of the N-terminal regions of several RAP74 homologues spaced throughout eukaryotic evolution. Beneath the alignment, the primary sequence is correlated with regions of secondary structure. Regions of  $\alpha$ -helix and  $\beta$ -sheet are derived from the crystal structure of human TFIIF (1) or else from secondary structure predictions (6–9) in the regions where structural information is not available. The x-ray crystal structure indicates a unique dimer interface made up of RAP74  $\beta$ -sheets  $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ ,  $\beta 6$ ,  $\beta 7$ ,  $\beta 8$ , and the corresponding  $\beta$ -sheets in RAP30. The  $\beta 4$  and  $\beta 5$  sheets of RAP74 interact to form the structured base of a loop that diverges from the dimer core. In the crystal structure, the RAP74  $\alpha 1$  helix makes intimate contacts with the  $\beta 4$  and  $\beta 5$  sheets.

The  $\alpha 1$  helix has been shown previously to be highly sensitive to mutation (10, 11). Several single amino acid changes, particularly in hydrophobic residues, cause significant defects in both accurate initiation and elongation. Because of symmetrical effects on initiation and elongation,  $\alpha 1$  should function by contacting a molecular target common to both processes, which would implicate RNA polymerase II, TFIIF, and DNA as the most likely targets. Mutations in  $\alpha 1$  do not appear to affect interaction with the RAP30 subunit, and no RAP30 contacts are identified for  $\alpha 1$  in the crystal structure (1). As far as can be discerned,  $\alpha 1$  mutations do not have an effect on assembly of gel-shifted preinitiation complexes that include the TATA-binding protein (TBP), TFIIB, RNA polymerase II, and TFIIF (10). Therefore, if the molecular target of  $\alpha 1$  is RNA polymerase II, the  $\alpha 1$  contact is not critical for TFIIF binding, and, from the TFIIF structure,  $\alpha 1$  does not appear to contact RAP30. Particularly sensitive residues within or adjacent to  $\alpha 1$  include Leu-155, Trp-164, Asn-172, Ile-176, and Met-177. Substituting any of these residues with alanine, which is a change consistent with maintaining the  $\alpha$ -helical structure, significantly reduces accurate initiation and elongation stimulation without apparent effects on known protein-protein contacts (10, 11).

Before the availability of the x-ray crystal structure, our laboratory began a systematic mutagenic analysis of conserved regions within the N-terminal domain of RAP74. When the structure became available, additional mutations were made within regions observed to contact the  $\alpha 1$  helix. In the current work, the effects of mutations within the neighboring N-terminal domain are compared with the effects of mutations within  $\alpha 1$ .

### EXPERIMENTAL PROCEDURES

*RAP74 Mutagenesis and Reconstitution of Recombinant TFIIF*—Amino acid substitution mutants were constructed using the Stratagene QuikChange Mutagenesis kit. Primers were designed to incorporate 2–4 clustered, site-directed substitutions via polymerase chain reaction from a pET21a plasmid (Novagen) encoding wild type RAP74 fused to a C-terminal His<sub>6</sub> tag. Wild type and mutant proteins were purified by Ni<sup>2+</sup>-affinity chromatography, and the TFIIF complex was reconstituted as described (12, 13). Substitution mutations were either multiple alanine replacements or radical charge reversals. The <sup>14</sup>YVV<sup>16</sup> →

\* This work was supported by a grant from the National Institutes of Health (to Z. F. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by the Hughes Undergraduate Research Program at Michigan State University.

§ To whom correspondence should be addressed. Tel.: 517-353-0859; Fax: 517-353-9334; E-mail: Burton@msu.edu.

<sup>1</sup> The abbreviations used are: TFIIF, transcription factor IIF; RAP, RNA polymerase II-associating protein; TBP, TATA-binding protein; wt, wild type.

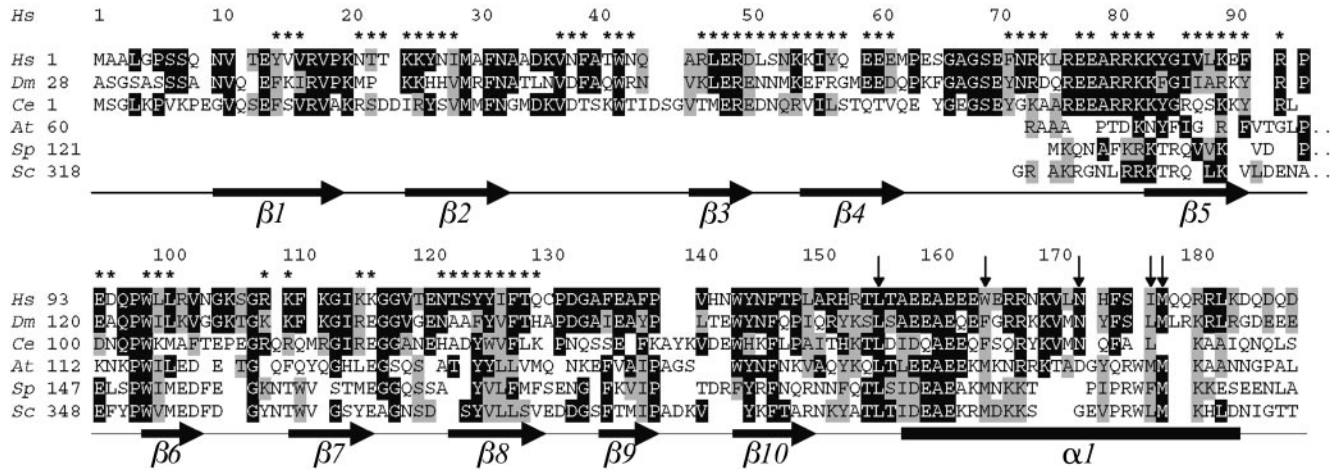


FIG. 1. **Human RAP74 homologues.** A proposed alignment of TFIIF large subunits from human (*Hs*) (4, 5), *Drosophila* (*Dm*) (22, 23), *Caenorhabditis elegans* (*Ce*) (AAB00717), *Arabidopsis thaliana* (*At*) (NP192998), *Schizosaccharomyces pombe* (*Sp*) (T41039) and *Saccharomyces cerevisiae* (*Sc*) (24, 25). Secondary structures from human RAP74 are indicated below the sequences. Not all of the  $\alpha 1$  helix is represented in the crystal structure (1), so the full extent of the helix is predicted from PHD secondary structure analysis (6–9). Asterisks identify amino acids substituted in this study; downward arrows indicate critical amino acids in  $\alpha 1$ . Black shading indicates identity with the human sequence, and gray shading indicates similarity.

<sup>14</sup>AAA<sup>16</sup> mutation contains Y14A, V15A, and V16A substitutions. Similarly, <sup>24</sup>KK<sup>25</sup> → <sup>24</sup>EE<sup>25</sup> contains K24E and K25E substitutions.

**Initiation Assays**—*In vitro* transcription assays were done essentially as described (10). An extract of human HeLa cell nuclei (cells purchased from the National Cell Culture Center, Minneapolis, MN) served as a source of transcription factors. TFIIF was removed from the extract by immunoprecipitation with anti-RAP74 and anti-RAP30 antibodies (5). Activity was reconstituted by the addition of wild type (wt) or mutant human recombinant TFIIF. The buffer for transcription reactions consisted of 12 mM HEPES (pH 7.9), 12% (w/v) glycerol, 0.12 mM EDTA, 0.12 mM EGTA, and 1.2 mM dithiothreitol and contained 60 mM KCl and 12 mM MgCl<sub>2</sub>. The template for transcription was plasmid pML, carrying the adenovirus major late promoter from position –258 to +196. The template was digested with *Sma*I endonuclease at position 217 relative to the transcription start site. 10 pmol of recombinant wt or mutant TFIIF extract and 0.8  $\mu$ g of DNA template were combined in a 20- $\mu$ l volume and incubated for 60 min at 25 °C. Transcription was initiated with 100  $\mu$ M ATP, 100  $\mu$ M CTP, and 5  $\mu$ Ci of [ $\alpha$ -<sup>32</sup>P]UTP (800 Ci per mmol) for 1 min. After pulse labeling, reactions were chased for 30 min with 1 mM each ATP, CTP, GTP, and UTP in the presence of 0.25% Sarkosyl. We find that Sarkosyl treatment increases the yield of full-length transcripts, apparently by blocking factors in the extract that induce pausing and/or termination (11, 14). Reactions were stopped by adding 80  $\mu$ l of stop solution (10 mM Tris-HCl, pH 7.9, 20 mM EDTA, and 0.5% sodium dodecyl sulfate containing 40  $\mu$ g of yeast tRNA per reaction). After extraction with phenol/chloroform and precipitation with ethanol, samples were resuspended in 80% formamide gel loading dye, boiled 2 min, and electrophoresed in a 6% polyacrylamide gel containing 50% (w/v) urea and Tris borate/EDTA buffer. Band intensities were quantitated using an Amersham Biosciences PhosphorImager. Initiation activities are reported as percent of wild type, with each value representing the average of triplicate measurements and error bars reflecting the standard deviation (Fig. 4).

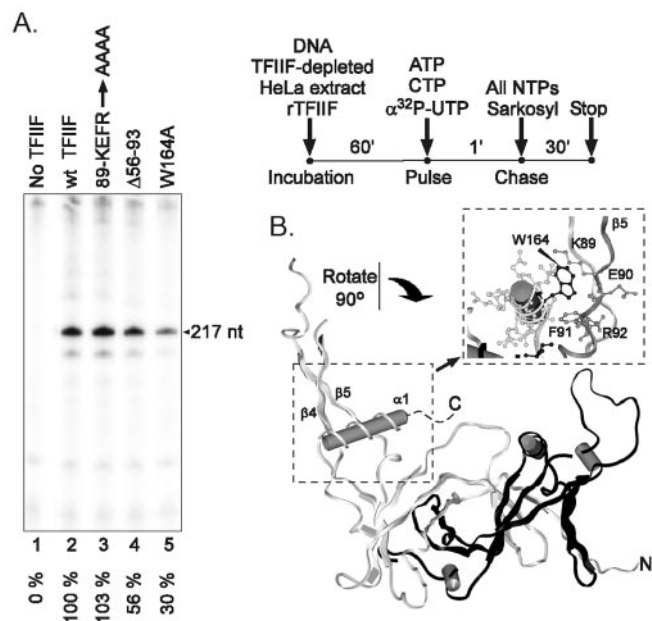
**Elongation Assays**—Elongation assays were done as described previously (11). The source of transcription factors was an extract derived from human HeLa cells. Bead-immobilized U20 templates were synthesized by polymerase chain reaction amplification from the pML20–42 template (15) using an upstream biotinylated primer and immobilized on streptavidin-coated MagneSphere paramagnetic beads (Promega). The C40 DNA is identical to U20, except that C40 has an additional 20-base extension in the CU cassette immediately downstream of the adenovirus major late promoter (<sup>1</sup>ACTCTCTCCCTTCTCTTCTCTCTCTCTCCCTCTCTCCAAAGGCCTTT<sup>50</sup>). The U20 or C40 bead template and extract were combined and incubated for 60 min at 25 °C. Transcription was extended to U20 or C40 with a 10 min pulse with 20  $\mu$ M dATP, 1 mM ApC dinucleotide, 100  $\mu$ M UTP, and 5  $\mu$ Ci of [ $\alpha$ -<sup>32</sup>P]CTP. Complexes were then washed twice with 500  $\mu$ l of transcription buffer containing 1% Sarkosyl, 0.5 M KCl, and 0.003% Nonidet P-40 to remove nascent elongation and termination factors and twice with 500  $\mu$ l of transcription buffer containing 60 mM KCl but lacking

MgCl<sub>2</sub>. Beads were collected with a magnetic particle separator. Complexes were resuspended in an appropriate volume of transcription buffer containing 60 mM KCl and 12 mM MgCl<sub>2</sub> and aliquoted into individual 15- $\mu$ l samples. 10 pmol of wt or mutant TFIIF was added to the reaction mixtures, with transcription buffer added in the no factor control. Elongation was continued from U20 or C40 by the addition of 1 mM each ATP, CTP, UTP, and GTP. Reactions were stopped at 5, 10, 20, or 40 s by the addition of 50  $\mu$ l of 0.5 M EDTA. Beads were collected with a magnetic particle separator, and the supernatant was removed. Beads were resuspended in 10  $\mu$ l of 80% formamide gel-loading buffer, boiled 5 min, and supernatant loaded into a 9% polyacrylamide gel containing 50% (w/v) urea and Tris borate/EDTA buffer. Gel images were visualized using an Amersham Biosciences PhosphorImager.

**Running Start, Two-bond Elongation Assay**—C40 elongation complexes initiated from the adenovirus major late promoter were prepared on bead templates as described above. Each 10- $\mu$ l reaction contained C40 complexes in transcription buffer containing 12 mM MgCl<sub>2</sub>. Reactions were incubated with 12 pmol of wt TFIIF, I176A TFIIF, or buffer for 20–60 min in the presence of 20  $\mu$ M CTP and UTP. The time varies because of the time required to process individual samples using the KinTek Rapid Chemical Quench-Flow instrument (RQF-3). 10  $\mu$ l of 200  $\mu$ M ATP (100  $\mu$ M working concentration) was added in transcription buffer for 30 or 120 s to extend the elongation complex to the A43 position. The time of ATP addition was 30 s in the presence of wt TFIIF, 120 s in the presence of I176A, and 120 s in the absence of the elongation factor. For each protocol, times for ATP addition were optimized for conversion of the C40 complex to A43 and for the efficiency and consistency of subsequent elongation to G44 and G45. After about 20 s the efficiency of elongation from A43 reaches a steady state that is maintained for several minutes; thus, samples processed with different ATP pulse times can be compared. During the short incubation with ATP, 15  $\mu$ l of the reaction mix was rapidly injected into the left sample port of the KinTek RQF-3 instrument. Initiation of the reaction induces mixing of the sample with 15  $\mu$ l of 500  $\mu$ M GTP, which was previously loaded into the right sample port of the RQF-3 instrument (250  $\mu$ M GTP working concentration). Reactions were quenched with 0.5 M EDTA at the times indicated. Beads were collected with a magnetic particle separator and resuspended in formamide gel loading dyes. Samples were electrophoresed in 16% polyacrylamide gels containing 50% (w/v) urea and analyzed as described above. Recovery of samples was sometimes variable; therefore, for each sample lane the sum of all transcripts generated from A43 was determined as 100%.

## RESULTS

**Radical Mutagenesis of the N-terminal Domain of Human RAP74**—Site-directed mutations were constructed throughout the conserved N-terminal region of human RAP74. In all, 68 residues were substituted between amino acids 1–129. Two deletions (29–227 and 62–227) were constructed within a

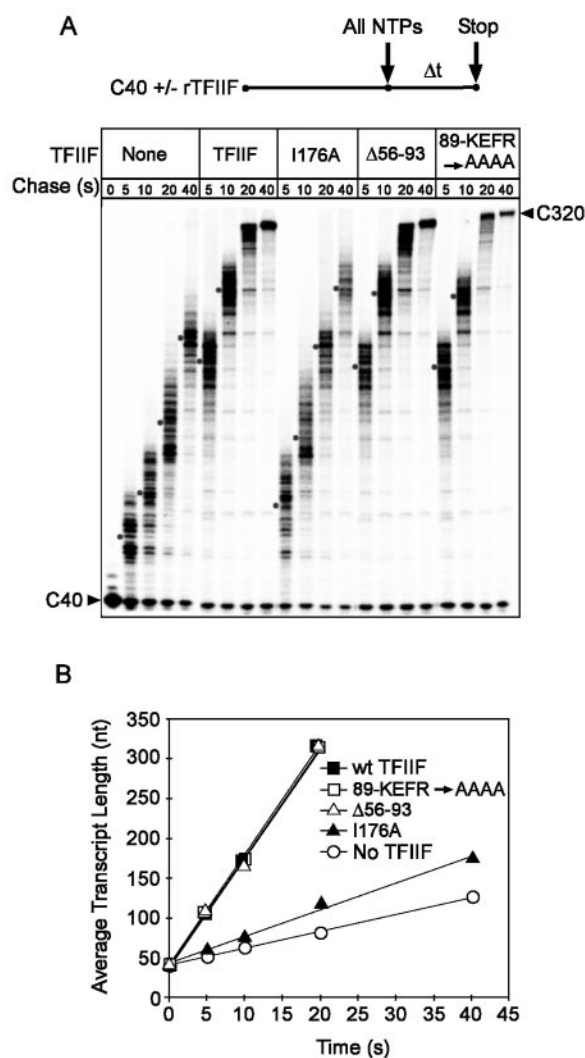


**FIG. 2. Accurate initiation activities of  $\alpha 1$  and  $\beta 4$ - $\beta 5$  loop mutants.** *A*, *in vitro* transcription from the adenovirus major late promoter. The 217-nucleotide (*nt*) runoff transcript is indicated. PhosphorImager quantitation of gel bands is indicated beneath the lane numbers as a percentage of wt. The reaction protocol is summarized above the TFIIIF structure. *B*, TFIIIF x-ray structure (1) highlighting positions of  $\alpha 1$  and the closely associated  $\beta 4$ - $\beta 5$  loop. Trp-164 (W164) in  $\alpha 1$  and Lys-89 (K89), Glu-90 (E90), Phe-91 (F91), and Arg-92 (R92) in  $\beta 5$  are indicated by darker shading in the detailed image of the  $\alpha 1$ - $\beta 5$  interaction.

1–227 C-terminal truncation, because it was previously known that RAP74 (1–227) has near wild type activity in most *in vitro* assays (10, 11, 14). All other mutations were made in full-length RAP74(1–517), including a large internal deletion,  $\Delta 56$ –93, described below.

Mutants were tested for accurate initiation activity in a HeLa cell extract that was depleted of TFIIIF by immunoprecipitation with anti-RAP30 and anti-RAP74 antibodies (Fig. 2). Accurate initiation activity was restored by the addition of human recombinant TFIIIF. NTPs were then added in a pulse-chase protocol in which Sarkosyl is added during the chase to block premature termination and also re-initiation (11, 14). Sample initiation data are compared for wild type TFIIIF, two mutants ( $^{89}\text{KEFR}^{92} \rightarrow ^{89}\text{AAAA}^{92}$  and  $\Delta 56$ –93) in which specific contacts between the  $\beta 4$ - $\beta 5$  loop and the  $\alpha 1$  helix are disrupted, and a critical and representative mutant within  $\alpha 1$  (W164A). The  $^{89}\text{KEFR}^{92} \rightarrow ^{89}\text{AAAA}^{92}$  mutation eliminates all close contacts with Trp-164 observed in the x-ray crystal structure (Fig. 2B), yet  $^{89}\text{KEFR}^{92} \rightarrow ^{89}\text{AAAA}^{92}$  demonstrates near wild type activity in initiation, whereas the  $\alpha 1$  mutation W164A has a much more severe defect (Fig. 2A). Clearly, eliminating strong contacts between the  $\beta 4$ - $\beta 5$  loop and  $\alpha 1$  does not have as great an effect on transcription as making a critical substitution in  $\alpha 1$ , indicating that the  $\beta 4$ - $\beta 5$  loop is not the functional molecular target of the  $\alpha 1$  helix or a necessary participant in  $\alpha 1$  function. In the  $\Delta 56$ –93 mutation, the entire  $\beta 4$ - $\beta 5$  loop was removed. This alteration produces a significant defect in initiation (56% wt), but one that is not quite so severe as the W164A mutation (30% wt).

TFIIIF stimulates the average rate of elongation by RNA polymerase II about 5–6-fold (Fig. 3) (11, 16, 17). Using bead templates, transcription was initiated from the adenovirus major late promoter in the presence of dATP, ApC dinucleotide, [ $\alpha$ - $^{32}\text{P}$ ]CTP, and UTP to synthesize U20 or C40 elongation complexes (RNAs are 20 nucleotides long ending in a 3'-UMP



**FIG. 3. Elongation stimulation activities of  $\alpha 1$  and  $\beta 4$ - $\beta 5$  loop mutants.** *A*, elongation stimulation assay. The protocol is indicated above the panel. RNA polymerase C40 elongation complexes were supplemented with TFIIIF samples as indicated. Elongation times were 0, 5, 10, 20, and 40 s. *Dots* indicate the estimated midpoint of the distribution of RNA lengths. *B*, plots of average transcript length versus elongation time.  $R^2$  for line slopes varied between 0.98 and 1.0.

(U20) or 40 nucleotides long and ending in a 3'-CMP (C40)). Elongation complexes were then stripped of elongation and termination factors with Sarkosyl and salt. After re-equilibration with transcription buffer, recombinant TFIIIF was added back to reactions, and elongation was continued with 1 mM each NTP for 5, 10, 20, or 40 s. In Fig. 3, some strongly paused C40 complexes remain at the C40 position, but most C40 complexes move forward as a fairly synchronous zone of transcripts. No C40 complexes appear to be arrested, as all advance within 2 min (see Fig. 5). The midpoint of the zone was approximated, as indicated by the *black dots* in Fig. 3A. The average elongation rate was then determined as the slope of the average transcript length plotted against the time of transcript extension (Fig. 3B). The average rate of RNA polymerase II elongation is determined to be more than 5-fold as fast in the presence of wt TFIIIF as compared with its absence. A single amino acid change in the  $\alpha 1$  helix (I176A) produced a protein with only about twice the average elongation rate of RNA polymerase II in the absence of factor, demonstrating the severe elongation defect of  $\alpha 1$  mutations but confirming that I176A TFIIIF interacts with and stimulates the RNA polymerase II elongation complex (10, 11). Significantly, however, the

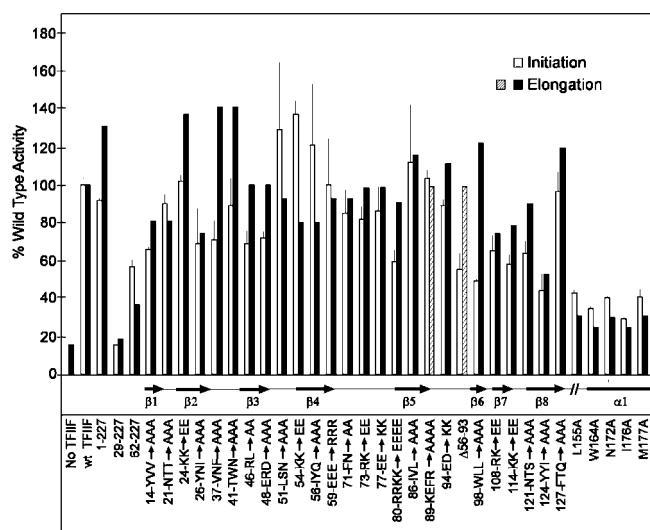


FIG. 4. Comparison of mutations in the N-terminal domain to mutations in  $\alpha 1$ . Accurate initiation assays were done as in Fig. 2, and elongation stimulation assays as in Fig. 3. Initiation assays were done at least in triplicate and are reported as average  $\pm$  S.D. Elongation assay results for  $\beta 4$ – $\beta 5$  loop mutants,  $^{89}\text{KEFR}^{92} \rightarrow ^{89}\text{AAAA}^{92}$  ( $^{89}\text{KEFR} \rightarrow \text{AAAA}$ ) and  $\Delta 56$ – $93$ , are highlighted by striped bars.

$^{89}\text{KEFR}^{92} \rightarrow ^{89}\text{AAAA}^{92}$  and even the large  $\Delta 56$ – $93$  deletion score as wild type in elongation stimulation activity. Despite tight packing of  $\alpha 1$  to  $\beta 4$  and  $\beta 5$  in the crystal structure, the molecular target of  $\alpha 1$  during elongation cannot be the  $\beta 4$ – $\beta 5$  loop, which is damaged or removed in these mutants.

In Fig. 4, initiation and elongation data are shown for a large number of TFIIF mutations within the N-terminal region. Deletion of amino acids 1–28 from the N terminus of RAP74 (29–227) causes a severe defect in both initiation and elongation. Surprisingly, further deletion to create RAP74-(62–227) was not quite so defective in initiation or elongation. Because the larger deletion was not expected to be more active, clones and proteins were reconfirmed. Apparently, both deletions damage the RAP74-RAP30 dimer interface, but the 29–227 fragment may have a more awkward conformation, causing its greater transcriptional defects. Results with these deletion mutations attest to the expected importance of the unique RAP74-RAP30 dimer interface in maintaining TFIIF conformation. Some other mutants, such as triple alanine substitutions  $^{98}\text{WLL}^{100} \rightarrow ^{98}\text{AAA}^{100}$  and  $^{124}\text{YYI}^{126} \rightarrow ^{124}\text{AAA}^{126}$ , have significant defects in transcription, but these mutations are not quite as severe as single alanine substitutions within or immediately adjacent to  $\alpha 1$  (L155A, W164A, N172A, I176A, and M177A). The defects of  $^{98}\text{WLL}^{100} \rightarrow ^{98}\text{AAA}^{100}$  and  $^{124}\text{YYI}^{126} \rightarrow ^{124}\text{AAA}^{126}$  in initiation might be attributed to defects in the RAP74-RAP30 dimer interface and to resulting defects in the transcription complex assembly. Indeed,  $^{98}\text{WLL}^{100} \rightarrow ^{98}\text{AAA}^{100}$  and  $^{124}\text{YYI}^{126} \rightarrow ^{124}\text{AAA}^{126}$  both show defects in the assembly of a TBP-TFIIB-RNA polymerase II-TFIIF complex (data not shown).  $^{98}\text{WLL}^{100} \rightarrow ^{98}\text{AAA}^{100}$  shows no defect in elongation stimulation, although it has a significant defect in initiation. Many multiple alanine substitutions and some multiple radical charge reversals show little effect on transcription. Apparently, most of the function of the RAP74 N terminus (amino acids 1–120) can be accounted for by a role in forming the RAP74-RAP30 dimer interface, as indicated by the crystal structure.

The  $\beta 4$ – $\beta 5$  loop is the most prominent feature of the N-terminal domain of RAP74 that diverges from the dimer interface, and this structure may have a subtle role in initiation that does not appear to be important for elongation. In the crystal structure, the  $\beta 4$ – $\beta 5$  loop interacts with  $\alpha 1$ , but, as we argue

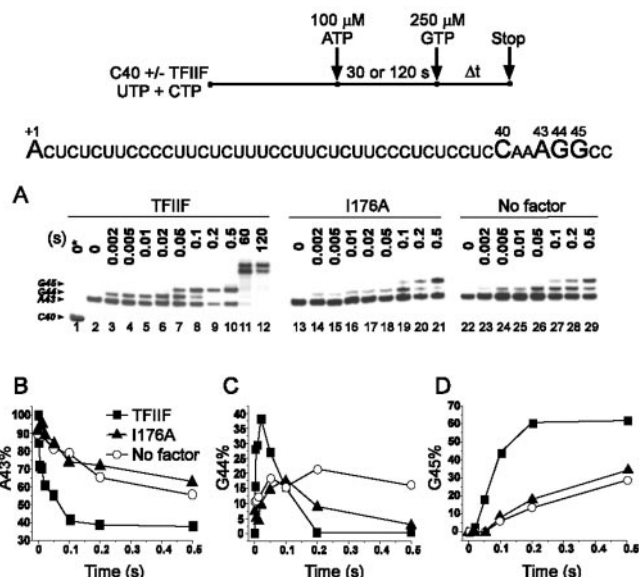


FIG. 5. Elongation defect of an  $\alpha 1$  mutation (I176A). A running start, two-bond elongation assay was done through the RNA sequence  $^{40}\text{CAAAGGCCUU}^{50}$  to monitor rates of G44 and G45 synthesis using a KinTek Rapid Chemical Quench-Flow instrument. Purified C40 elongation complexes were advanced to A43 by addition of  $100 \mu\text{M}$  ATP for 30 s (wt TFIIF) or 120 s (I176A TFIIF or no factor).  $250 \mu\text{M}$  GTP was then added for the indicated times, and reactions were quenched. A, gel data comparing elongation by wt TFIIF (lanes 1–12), I176A TFIIF (lanes 13–21) and no factor (lanes 22–29). GTP chase times are indicated in seconds (s). 0\* indicates no ATP pulse, no GTP chase. B–D, PhosphorImager quantitation of the data shown in panel A. B, rates of A43 disappearance. C, the flow of transcripts through the G44 position. D, rates of G45 synthesis. NTP concentrations just prior to quench addition were  $5 \mu\text{M}$  CTP,  $5 \mu\text{M}$  UTP,  $50 \mu\text{M}$  ATP, and  $250 \mu\text{M}$  GTP. The data shown are representative of several experiments.

above,  $\beta 4$ – $\beta 5$  cannot be the functional molecular target of  $\alpha 1$ . However, the loop charge reversal mutant  $^{80}\text{RRKK}^{83} \rightarrow ^{80}\text{EEEE}^{83}$  and loop deletion  $\Delta 56$ – $93$  have only about 60% wt activity in initiation, although they show approximate wt activity in elongation. It is possible, therefore, that the  $\beta 4$ – $\beta 5$  loop functionally contacts one of the general initiation factors not present in the Sarkosyl-washed elongation complex.

**Analysis of an  $\alpha 1$  Mutant**—Relating the TFIIF crystal structure to our functional analysis of RAP74 mutants refocused our attention on the  $\alpha 1$  helix. In Fig. 5, rates for RNA polymerase II elongation are compared in the presence of wt TFIIF, I176A TFIIF and in the absence of elongation factor. The protocol we employed is referred to as a “running start, two-bond” elongation assay. The running start was necessary to measure the most rapid elongation rates, and two bonds were monitored because rates of G44 and G45 synthesis yield distinct information.<sup>2</sup> A kinetic lag in the first appearance of G45, for instance, indicates a slow step that occurs after formation of the G44 bond but before rapid formation of the G45 bond can commence.

C40 elongation complexes were formed on bead templates and washed with Sarkosyl as described above. C40 complexes were incubated with 12 pmol TFIIF and  $20 \mu\text{M}$  CTP and UTP. The running start is initiated by adding  $100 \mu\text{M}$  ATP for 30 or 120 s to advance C40 to the A41, A42, and A43 positions. The time of ATP incubation was 30 s with wt TFIIF, 120 s with I176A TFIIF, and 120 s with no factor. For each protocol the ATP pulse time was optimized for conversion of C40 to A43. With time delays from 20 s to several minutes, A43 complexes are extended efficiently, so the 30 and 120 s stall times repre-

<sup>2</sup> Y. A. Nedialkov, X. Q. Gong, and Z. F. Burton, unpublished data.

sent a reasonable comparison between samples. During the ATP pulse, complexes were quickly transferred into the sample port of the KinTek Rapid Chemical Quench-Flow instrument and then mixed with 250  $\mu\text{M}$  GTP, allowing further extension to G44, G45, and longer products for the times indicated.

TFIIF stimulates the rates of G44 and G45 synthesis (Fig. 5A). Furthermore, TFIIF accelerates a slow step that occurs after chemistry in the transition between one bond and the next. In the presence of wt TFIIF the following events occur. 1) A significant burst of G44 synthesis is observed at 0.002 s. 2) The peak of G44 accumulation is observed at about 0.02 s. 3) The first appearance of G45 is observed at 0.05 s. 4) The peak of G45 accumulation is observed at 0.1–0.5 s. There appears to be a delay in the first appearance of G45, because G44 is present from the very earliest time points; but the first appearance of G45 is delayed until 0.02–0.05 s. Therefore, the most rapid synthesis rates for the G44 bond must be faster than the most rapid synthesis rates for G45. If this were not the case, the first appearance of G45 would be expected by the 0.005 to 0.01 s time points, at which no G45 is detectable (Fig. 5A; lanes 3–6). The delay in the first appearance of G45 provides evidence for a slow step after phosphodiester bond formation.<sup>2</sup> In the presence of TFIIF carrying the deleterious I176A mutation in the RAP74 subunit or in the absence of TFIIF, the following events occur. 1) The rates of G44 synthesis are slowed. 2) The rates of G44 disappearance are slowed. 3) The first appearance of G45 is delayed.

Quantification of the gel data (Fig. 5, B–D) supports these conclusions and yields further insight into wt TFIIF function and I176A TFIIF defects. Fig. 5B shows the rates of escape from the A43 position. In the presence of wt TFIIF, about 60% of A43 complexes extend rapidly to G44 (within 0.1 s), indicating that these A43 complexes were poised on the active synthesis pathway. On the other hand, about 40% of A43 elongation complexes are strongly paused when GTP is added, but all of these complexes advance within 120 s (Fig. 5A, lanes 11 and 12). TFIIF wt stimulates a much faster rate of A43 disappearance than observed in the absence of an elongation factor (Fig. 5B), indicating, as others have noted (16, 18, 19), that TFIIF suppresses pausing. The I176A mutant does not support the pause-suppression function of TFIIF and instead supports elongation rates similar to those observed in the absence of an elongation factor (Fig. 5B). Fig. 5C shows passage of elongation complexes through the G44 position. TFIIF wt increases the rate of appearance and dramatically enhances the rate of disappearance of G44. By contrast, the I176A mutant supports a G44 synthesis rate that is significantly slower than wild type TFIIF and similar to elongation in the absence of factor. I176A TFIIF, however, does stimulate overall elongation rates (Figs. 3 and 4) and, in the running start assay, the mutant is observed to facilitate escape from the G44 position when compared with RNA polymerase II in the absence of the elongation factor (Fig. 5A, lanes 19–21 and 27–29; Fig. 5C). Wild type TFIIF strongly stimulates the G45 synthesis rate and decreases the time for the first G45 appearance (Fig. 5D). The kinetic lag in the G45 appearance is less than 0.05 s in the presence of wt TFIIF, but this lag is 0.05–0.1 s in the presence of I176A or in the absence of elongation factor. The lag in first G45 appearance reflects a slow step between formation of one bond and the next that occurs after phosphodiester bond formation and is regulated by TFIIF.<sup>2</sup> In the running start protocol, this slow step is clearly observed in the first appearance of G45 but is not observed in G44 synthesis rates, because this slow step is complete after the 30 or 120 s delay at A43. In the presence of TFIIF, RNA polymerase II negotiates the slow transition between bonds efficiently. In the presence of I176A TFIIF, this transition

becomes much more challenging for RNA polymerase II (Fig. 5A, compare lanes 6–8 with lanes 17–19; Fig. 5D). In the absence of factor, the transition is also slow, further demonstrating regulation of this slow step between bonds by TFIIF.

#### DISCUSSION

A large number of radical amino acid substitutions were introduced in the conserved N-terminal domain of human RAP74 without clear identification of particularly interesting new mutants. It appears that most functional outcomes induced by N-terminal domain mutations are accounted for by damage to the RAP74-RAP30 dimer interface. The  $\beta 4$ - $\beta 5$  loop of RAP74, however, diverges from the dimer core and packs tightly with the  $\alpha 1$  helix in the crystal structure. This is an intriguing association, because  $\alpha 1$  was previously shown to be important for both initiation and elongation (10, 11). Surprisingly, however,  $\beta 4$ - $\beta 5$  and the intervening loop do not appear to participate in  $\alpha 1$  function, because radical damage to the  $\beta 5$ - $\alpha 1$  interaction (<sup>89</sup>KEFR<sup>92</sup>  $\rightarrow$  <sup>89</sup>AAAA<sup>92</sup>) or complete removal of  $\beta 4$ - $\beta 5$  and the intervening loop ( $\Delta 56$ -93) did not have as severe defects in initiation or elongation as critical point mutations in  $\alpha 1$  (L155A, W164A, N172A, I176A, and M177A). Therefore, we conclude that the close association of the  $\beta 4$ - $\beta 5$  loop and  $\alpha 1$  seen in the x-ray crystal structure does not represent the form of  $\alpha 1$  that is active in initiation and elongation of RNA chains.

On the other hand, interaction with the  $\beta 4$ - $\beta 5$  loop could modulate  $\alpha 1$  function. For instance, binding to  $\beta 5$  may shelter the otherwise exposed tryptophan W164 when  $\alpha 1$  is not bound to its transcriptional target (Fig. 2B). We, therefore, suggest a regulatory role for the  $\beta 4$  and  $\beta 5$  interaction with  $\alpha 1$ . We further suggest that  $\alpha 1$  must free itself of contact with the  $\beta 4$ - $\beta 5$  loop while participating in transcription. We note that the  $\Delta 56$ -93 deletion is somewhat defective in supporting initiation, perhaps indicating interaction of the loop with another initiation factor. Because the defect of  $\Delta 56$ -93 in initiation approximates the defect of a radical charge reversal mutant in the loop between  $\beta 4$  and  $\beta 5$  (<sup>80</sup>RRKK<sup>83</sup>  $\rightarrow$  <sup>80</sup>EEEE<sup>83</sup>), elimination of <sup>80</sup>RRKK<sup>83</sup> in  $\Delta 56$ -93 could account for its initiation defect (Fig. 4). It may be that participation of the  $\beta 4$ - $\beta 5$  loop in the proper assembly of the preinitiation complex results in freeing the  $\alpha 1$  helix for its roles during initiation and elongation.

Because substitution with alanine is consistent with maintaining  $\alpha$ -helical structures (20, 21), critical substitutions within  $\alpha 1$  (L155A, W164A, N172A, I176A, and M177A) are unlikely to disrupt local protein secondary structure. Furthermore, these critical substitutions do not appear to affect major interactions with RAP30 or the preinitiation complex (10, 11). We suggest, therefore, that L155, W164, N172, I176, and M177 may be directly involved in contacting a common molecular target in initiation and elongation complexes. Each of these proposed contacts appears to be of critical importance, because L155A, W164A, N172A, I176A, and M177A all have approximately the same defect in initiation and elongation as an eight-amino acid deletion within  $\alpha 1$  ( $\Delta 170$ -177) (10, 11).

The current work fails to identify the molecular target for  $\alpha 1$ , although the  $\beta 4$ - $\beta 5$  loop was implicated by the crystal structure. It has been argued that the target of  $\alpha 1$  must reside in TFIIF, RNA polymerase II, or template DNA, because these are the components in common comparing initiation and elongation complexes, and mutations in  $\alpha 1$  have similar defects in initiation and elongation (10, 11). To date, no contacts to other regions within TFIIF can be identified for  $\alpha 1$  that appear to participate in its transcriptional functions. Furthermore, TFIIF carrying critical lesions in  $\alpha 1$  binds to RNA polymerase II and forms a preinitiation complex of TBP, TFIIB, and RNA polymerase II with no apparent alteration of complex mobility

in gels or defect in assembly even at low TFIIF concentrations. Thus, if  $\alpha 1$  targets RNA polymerase II, this interaction does not appear to be necessary for tight TFIIF-RNA polymerase II binding or complex assembly (10). Perhaps  $\alpha 1$  targets DNA in initiation and elongation complexes, although no clear demonstration of this hypothesis is yet available. It appears from this study and from the crystal structure that  $\alpha 1$  must be freed from tight contact with the  $\beta 4$ - $\beta 5$  loop to swing on a flexible pivot to target a specific region of RNA polymerase II or template DNA. In this regard, it is interesting to compare  $\alpha 1$  helices from different organisms. In every case, PHD secondary structure prediction indicates that these regions are extended  $\alpha$ -helices, although the primary sequences are somewhat divergent (Fig. 1).

These studies have refocused our attention on  $\alpha 1$  of RAP74 as an important functional motif in transcriptional mechanisms. As a preliminary step, we have compared a critical  $\alpha 1$  mutant (I176A) to wt TFIIF in a running start, two-bond transcriptional elongation assay. Our results suggest that TFIIF stimulates escape from paused complexes and also accelerates a slow step that follows phosphodiester bond formation. Further investigation will be required to fully characterize the multiple effects of TFIIF on elongation and the severe defects of  $\alpha 1$  mutants.

## REFERENCES

- Gaiser, F., Tan, S., and Richmond, T. J. (2000) *J. Mol. Biol.* **302**, 1119–1127
- Groft, C. M., Uljon, S. N., Wang, R., and Werner, M. H. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9117–9122
- Kamada, K., De Angelis, J., Roeder, R. G., and Burley, S. K. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3115–3120
- Aso, T., Vasavada, H. A., Kawaguchi, T., Germino, F. J., Ganguly, S., Kitajima, S., Weissman, S. M., and Yasukochi, Y. (1992) *Nature* **355**, 461–464
- Finkelstein, A., Kostrub, C. F., Li, J., Chavez, D. P., Wang, B. Q., Fang, S. M., Greenblatt, J., and Burton, Z. F. (1992) *Nature* **355**, 464–467
- Rost, B., and Sander, C. (1993) *J. Mol. Biol.* **232**, 584–599
- Rost, B., Sander, C., and Schneider, R. (1994) *Comput. Appl. Biosci.* **10**, 53–60
- Rost, B., and Sander, C. (1994) *Proteins* **19**, 55–77
- Rost, B. (1996) *Methods Enzymol.* **266**, 525–539
- Ren, D., Lei, L., and Burton, Z. F. (1999) *Mol. Cell. Biol.* **19**, 7377–7387
- Lei, L., Ren, D., and Burton, Z. F. (1999) *Mol. Cell. Biol.* **19**, 8372–8382
- Wang, B. Q., Kostrub, C. F., Finkelstein, A., and Burton, Z. F. (1993) *Protein Expression Purif.* **4**, 207–214
- Wang, B. Q., Lei, L., and Burton, Z. F. (1994) *Protein Expression Purif.* **5**, 476–485
- Lei, L., Ren, D., Finkelstein, A., and Burton, Z. F. (1998) *Mol. Cell. Biol.* **18**, 2130–2142
- Samkurashvili, I., and Luse, D. S. (1996) *J. Biol. Chem.* **271**, 23495–23505
- Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991) *Mol. Cell. Biol.* **11**, 1195–1206
- Izban, M. G., and Luse, D. S. (1992) *J. Biol. Chem.* **267**, 13647–13655
- Price, D. H., Sluder, A. E., and Greenleaf, A. L. (1989) *Mol. Cell. Biol.* **9**, 1465–1475
- Tan, S., Aso, T., Conaway, R. C., and Conaway, J. W. (1994) *J. Biol. Chem.* **269**, 25684–25691
- Niu, W., Zhou, Y., Dong, Q., Ebright, Y. W., and Ebright, R. H. (1994) *J. Mol. Biol.* **243**, 595–602
- Tang, H., Sun, X., Reinberg, D., and Ebright, R. H. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1119–1124
- Gong, D. W., Horikoshi, M., and Nakatani, Y. (1993) *Nucleic Acids Res.*
- Kephart, D. D., Price, M. P., Burton, Z. F., Finkelstein, A., Greenblatt, J., and Price, D. H. (1993) *Nucleic Acids Res.* **21**, 1319
- Sun, Z. W., and Hampsey, M. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3127–3131
- Henry, N. L., Campbell, A. M., Feaver, W. J., Poon, D., Weil, P. A., and Kornberg, R. D. (1994) *Genes Dev.* **8**, 2868–2878