Functional Domains of Human RAP74 Including a Masked Polymerase Binding Domain*

Bo Qing Wang and Zachary F. Burton†

From the Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824

RAP74, the large subunit of human transcription factor IIIF (TFIIF), has been analyzed by deletion mutagenesis and in vitro assays to map functional domains. Tight binding to the RAP30 subunit involves amino acids between positions 1-172. Amino acids 1-205 are minimally sufficient to stimulate accurate transcription from the adenovirus major late promoter in an extract system, although C-terminal sequences contribute to activity. A partially masked RNA polymerase II binding domain has been mapped to the C-terminal region of the protein (amino acids 363-444). Sequences near the N terminus and within the central portion of RAP74 affect accessibility of this domain. Extending this domain to 363-486 creates a peptide that binds polymerase and DNA and inhibits transcription initiation in vitro from non-promoter DNA sites. This larger C-terminal domain may modify polymerase interaction with template during initiation and/or elongation of RNA chains.

Human general transcription factor (TF)1 IIIF (RNA polymerase II-associated protein (RAP) 30/74, FC, β3 in rat, Factor 1 in Drosophila, and factor γ in yeast) has functions in both initiation and elongation of RNA chains (1-12; reviewed in Refs. 13 and 14). The functions of the general factors will be understood based on their interactions with template, transcript, RNA polymerase II, other general factors, and regulatory factors.

TFIIF is structurally and functionally related to bacterial sigma factors. In humans TFIIF is a hetereric factor, consisting of 28-kDa (RAP30) and 58-kDa (RAP74) subunits, that binds directly to RNA polymerase II. The RAP30 subunit binds RNA polymerase II through a domain that is similar in sequence to the polymerase binding domain of bacterial sigma factors (15, 16). A masked DNA binding domain in the C-terminal region of RAP30 shows additional structural similarity to a DNA binding region of sigma factors involved in contacting the −35 region of bacterial promoters, although the corresponding RAP30 domain is not known to confer promoter recognition specificity to RNA polymerase II (17, 18).

Whether the RAP74 subunit can also bind directly to RNA polymerase II has not been clearly established, although several observations in the literature indicate such a possibility. RAP74 induces a gel mobility shift in a complex containing promoter DNA, TBP, TFIIB, and RNA polymerase II in the absence of RAP30 (19). RAP74 can also stimulate transcription elongation by RNA polymerase II in a reaction utilizing a 3′-dC-tailed DNA template in the absence of other factors (20). At least in the presence of template, therefore, RAP74 interacts with transcription complexes, presumably by binding polymerase.

Binding of RAP30 to RNA polymerase II blocks association between polymerase and nonspecific DNA sites, a function associated with bacterial sigma factors (21, 22). The TFIIF complex additionally dissociates polymerase from DNA sites to which it was previously bound (21, 22), and this is another similarity between TFIIF and sigma factors. The RAP30 subunit by itself does not have this additional capability (21). Dissociation of polymerase from nonspecific DNA sites promotes association with promoters, and in most in vitro systems, TFIIF must be present to promote stable association of polymerase with the preinitiation complex (7, 8). The RAP30 subunit has partial function in polymerase entry (7, 23), but the RAP74 subunit stabilizes assembly and modifies interaction of RAP30 with template (24).

Several laboratories have reported minimal in vitro systems in which some of the general factors become dispensable for accurate initiation from a promoter. Parvin and Sharp (25) reported that TBP, TFIIB, and RNA polymerase II were sufficient for accurate transcription from the adenovirus major late promoter, using a supercoiled template. Similarly, Usheva and Shenk (26) demonstrated that initiator-binding protein YY1, TFIIB, and RNA polymerase II were sufficient for initiation from the P5 promoter of adenovirus-associated virus, using a supercoiled template. Tyrer et al. (19) detected accurate initiation from several promoters in vitro using TBP, TFIIB, RAP30, RNA polymerase II, and supercoiled templates. Superciling obviates the requirement of an ATP-dependent step that involves TFIIE and TFIIF (27-29). TFIIFH includes an ATP-dependent DNA helicase activity that may function to separate DNA template strands during promoter escape by RNA polymerase II (30). Accurate initiation from linear DNA templates additionally requires TFIIF, TFIIE, TFIIFH, and hydrolysis of ATP (27).

The physiological meaning of these minimal systems is not obvious, but it is clear that accurate initiation can occur in the absence of some general factors and by slightly different mechanisms, depending on the factors present in the complex and the state of the template DNA. In several systems the RAP30 subunit of TFIIF demonstrates initiation functions independent of RAP74, particularly in assembly of the preinitiation complex (7). RAP74, on the other hand, has been shown to have some elongation functions in the absence of RAP30, although RAP30 also stimulates elongation in conjunction with RAP74.
RAP30 has not been shown to stimulate elongation in the absence of RAP74 (20, 31).

In experiments from our laboratory, using an extract system depleted of TFIIF by immunoprecipitation with anti-RAP30 antibodies, RAP30 was required for accurate initiation. RAP74 did not stimulate initiation but was required for very early elongation of the transcript (32). In these experiments, transcripts initiated in the absence of added RAP74 were stably associated with template DNA. Some other TFIIF-depleted extracts we have made behave somewhat differently. For these preparations, a small segment from the N-terminal region of RAP74 (amino acids 1–205) is required to prevent release of transcripts. It is not clear from these more recent studies whether RAP74 is completely dispensable for initiation, as indicated by the initial work (32). It is clear, however, that RAP74 sequence between amino acids 172–517 is dispensable for initiation. It is our view, based on these results and those of others, that RAP30 can, in some cases, provide all of the initiation functions of TFIIF. TFIIF is required for promoter escape by RNA polymerase II and, in some cases, to prevent release of newly initiated transcription complexes. RAP30 likely participates in both promoter escape and stabilization of newly initiated complexes since the most important region of RAP74 for these processes is the N-terminal region, which binds RAP30 (Ref. 36 and this paper). These partially distinct functions of TFIIF subunits appear to be influenced by interactions with other general factors and regulators. Since more defined systems may lack both positive and negative regulators, extract systems will continue to be of importance in identifying functions for the general transcription factors.

TFIIF interacts with TFIIIB through the RAP30 subunit, and together these factors cooperate to bring polymerase into the preinitiation complex (7, 8, 33). As discussed above, RNA polymerase II, TFIIF, and either TBP or YY1 can be sufficient for accurate initiation from promoters in vitro, indicating that TFIIB and polymerase might minimally suffice to select transcriptional start sites (25, 26). Some sua7 mutants in the yeast gene encoding TFIIFB are altered for selection of transcriptional start sites (34). Interestingly, the sua71 mutants, in the gene encoding a yeast RAP74 homolog, revert normal start site (34). Interestingly, the gene encoding TFIIB are altered for selection of transcriptional start sites (25, 26). Some other TFIIF-depleted extracts we have made have behaved somewhat differently. For these preparations, a small segment from the N-terminal region of RAP74 (amino acids 1–205) is required to prevent release of transcripts. It is not clear from these more recent studies whether RAP74 is completely dispensable for initiation, as indicated by the initial work (32). It is clear, however, that RAP74 sequence between amino acids 172–517 is dispensable for initiation. It is our view, based on these results and those of others, that RAP30 can, in some cases, provide all of the initiation functions of TFIIF. TFIIF is required for promoter escape by RNA polymerase II and, in some cases, to prevent release of newly initiated transcription complexes. RAP30 likely participates in both promoter escape and stabilization of newly initiated complexes since the most important region of RAP74 for these processes is the N-terminal region, which binds RAP30 (Ref. 36 and this paper). These partially distinct functions of TFIIF subunits appear to be influenced by interactions with other general factors and regulators. Since more defined systems may lack both positive and negative regulators, extract systems will continue to be of importance in identifying functions for the general transcription factors.

TFIIF interacts with TFIIIB through the RAP30 subunit, and together these factors cooperate to bring polymerase into the preinitiation complex (7, 8, 33). As discussed above, RNA polymerase II, TFIIF, and either TBP or YY1 can be sufficient for accurate initiation from promoters in vitro, indicating that TFIIB and polymerase might minimally suffice to select transcriptional start sites (25, 26). Some sua7 mutants in the yeast gene encoding TFIIFB are altered for selection of transcriptional start sites (34). Interestingly, the sua71 mutants, in the gene encoding a yeast RAP74 homolog, revert normal start site (34). Interestingly, the gene encoding TFIIB are altered for selection of transcriptional start sites (25, 26). Some other TFIIF-depleted extracts we have made have behaved somewhat differently. For these preparations, a small segment from the N-terminal region of RAP74 (amino acids 1–205) is required to prevent release of transcripts. It is not clear from these more recent studies whether RAP74 is completely dispensable for initiation, as indicated by the initial work (32). It is clear, however, that RAP74 sequence between amino acids 172–517 is dispensable for initiation. It is our view, based on these results and those of others, that RAP30 can, in some cases, provide all of the initiation functions of TFIIF. TFIIF is required for promoter escape by RNA polymerase II and, in some cases, to prevent release of newly initiated transcription complexes. RAP30 likely participates in both promoter escape and stabilization of newly initiated complexes since the most important region of RAP74 for these processes is the N-terminal region, which binds RAP30 (Ref. 36 and this paper). These partially distinct functions of TFIIF subunits appear to be influenced by interactions with other general factors and regulators. Since more defined systems may lack both positive and negative regulators, extract systems will continue to be of importance in identifying functions for the general transcription factors.

In this report, we have used deletion mutagenesis to map functional domains of human RAP74. We confirm the previous mapping of the RAP30 binding domain (36). The regions of RAP74 that are required for accurate initiation in a depleted extract system are strikingly different from those required in a more defined system (36), and indicate interation of RAP74 with both positive and negative factors in the extract. A partially masked RNA polymerase II binding domain has been identified in RAP74 by deletion of N-terminal and central amino acid sequences. This domain appears to modify template interactions by polymerase.

2 C.-h. Chang and Z. F. Burton, submitted for publication.

3 S. M. Fang and Z. F. Burton, submitted for publication.
polymerase) was blocked with 0.5 ml of SB containing 0.2% BSA by rocking at 4 °C for 1 h. About 500 pmol of RAP74 or RAP47 mutation was mixed with the resin and rocked for an additional hour at 4 °C. The mixtures were then washed five times with 1 ml of SB. The resin was eluted with 40 mM KCl and loaded onto a 16% SDS-PAGE gel. Proteins were visualized by staining with silver nitrate.

Non-promoter Transcription Assays—In vitro transcription assays in which RNA polymerase II initiates from non-promoter DNA sites were performed according to Killean and Greenblatt (21) with some modifications. The DNA template was a supercoiled plasmid DNA, in this case pML, which contains the adenovirus major late promoter subcloned as an XhoI to HindIII fragment (−256 to +196) into pBluescript II KS (+) (Stratagene) between the XhoI sites and HindIII sites of the vector (32). Since only RNA polymerase II and RAP74 are added to the assay, polymerase is unable to recognize the promoter on this template, and initiation occurs from many DNA sites. Similar results were observed with plasmid templates containing no known promoters (data not shown). The reaction mixtures (10 μl) contained 20 mM Tris-HCl (pH 8.0), 3 mM MnCl₂, 0.16 mM EDTA, 16% glycerol, 0.7 mM DTT, 110 mM NaCl, 250 μM ATP, CTP, and GTP, and 25 μM [5,6-3H]UTP (10 Ci/μmol, 1 μCi/μl) reaction, 20 μg/ml template DNA, 50 pmol of RAP74 protein, and 5 pmol of calf thymus RNA polymerase II. The reaction mixtures were incubated at 37 °C for 1 h and spotted on DE-81 filters (Whatman), which were then washed with 0.5 × Na₂HPO₄ and incorporation of [3H]UMP into RNA was determined by scintillation counting.

To test whether RAP74 affects initiation or elongation of RNA chains, polymerase was incubated with template at 37 °C for 15 min, NTPs were added, and elongation was continued for 1 h. RAP47 was added to the reaction at the indicated times, either before or after addition of NTPs.

Gel Mobility Shift Assays—Preparation of the probe DNA encoding the adenovirus major late promoter was amplified by the polymerase chain reaction between coordinates −262 and +30, using pML as the template for amplification. The amplified DNA product was gel-purified using the Qiaex kit (Qiagen) and 5 pmol (960 ng) purified DNA was radiolabeled by incubation with [γ-32P]ATP and T4 polynucleotide kinase. The 5′-labeled DNA was then digested with restriction enzyme HpaII to produce a 72-base pair single-stranded DNA probe extending from coordinates −262 to −191 of the adenovirus major late promoter. This radiolabeled DNA fragment is not known to include any promoter DNA sequences.

Gel mobility shift assays were performed according to Killean and Greenblatt (21) and Conaway and Conaway (22) with some modifications. The reaction mixtures (10 μl) contained 20 mM Tris-HCl (pH 8.0), 24 mM HEPES (pH 7.9), 2.4 mM DTT, 40 mM KCl, 11% glycerol, 0.5 mmg/ml BSA, 25 fmol of labeled DNA probe, 0.3 pmol of calf thymus pol IIa, and RAP74 (0.17, 0.34, 0.51, 0.68, 0.85, and 1 pmol of RAP74-H6). Reaction mixtures were incubated at 28 °C for 20 min and immediately loaded onto a 4% polyacrylamide gel containing 0.09% bisacrylamide, 2.5% glycerol and 0.5 × TBE. Electrophoresis was at 30 mA for 2 h. Dried gels were analyzed by autoradiography.

RESULTS

A set of RAP74 deletion mutants was constructed and analyzed in binding and functional assays (Fig. 1). Wild type human RAP74 is a 58-kDa protein composed of 517 amino acids. For the most part, mutants are named according to the amino acids remaining in the structure. For instance, RAP74(87–517) extends from amino acid 87 to 517 of the wild type protein. Some mutants have been constructed with internal deletions, and these have been named according to the amino acids removed from the sequence. For instance, RAP74(1137–356) includes amino acids 1–136 fused to amino acids 357–517 of the wild type protein. Mutant proteins were constructed with a C-terminal histidine tag to facilitate purification and binding assays. In addition, some mutant proteins have N-terminal extensions. The precise amino acid sequences of mutants are indicated under “Experimental Procedures.”

RAP30 Binding Domain of RAP74—The RAP30 binding domain of human RAP74 was previously mapped to amino acids 62–172 by Yonaha et al., using a two-hybrid genetic system (36). In the experiment shown in Fig. 2, we confirm their result using a direct binding assay. Deletion mutants of RAP47 with C-terminal histidine affinity tag were combined with full-length RAP30 in buffer containing 4 mM urea. Samples were dialyzed into buffer containing 0.5 M KCl and no urea. Previous work in our laboratory had shown that this method was efficient for reconstitution of the RAP30/47 complex (38). Reactions were mixed with Ni²⁺-chelate resin to bind the affinity tag on RAP47 mutants. After washing the resin, bound proteins were eluted with SDS-PAGE sample buffer. Since RAP30 did not carry the affinity tag, in order to bind to the resin, RAP30 had to bind first to a RAP47 fragment. Eluted proteins were analyzed by SDS-PAGE and Coomassie Blue staining (Fig. 2A) and by a Western blot developed with anti-RAP30...
antibodies (Fig. 2B).

By this analysis, amino acids between 1–172 of RAP74 are important for RAP30 binding (lanes 2–7). Further deletion to amino acid 136 abolishes strong RAP30 binding (lane 8). Mutants 74–517 (lane 9) and 87–517 (lane 10) have very weak interaction with RAP30, indicating that sequences between 1 and 74 and between 87 and 172 contribute to binding. Δ136–356 also showed a very weak interaction with RAP30 (lane 16), confirming that at least part of the binding domain is included between amino acids 137 and 172. The histidine affinity tag on RAP74 mutants was necessary for retention of RAP30 on the beads, because RAP74 without the tag did not retain RAP30 (lane 1).

Stimulation of Accurate Transcriptional Activity—Accurate runoff transcription initiated from the Adenovirus major late promoter was assayed in vitro in an extract system depleted of TFII F by immunoprecipitation with anti-RAP30 and anti-RAP74 antibodies (37, 38) (Fig. 3). This system is completely dependent on addition of RAP30 and RAP74 to reconstitute transcription. Depleted extract, RAP30, and template DNA were mixed in the presence of RAP74 or RAP74 deletion mutants. Full-length RAP74 (1–517) appears to have higher activity than any of the deletion mutants, although mutants 1–409, 1–356, and 1–296 also strongly stimulated transcription. Further deletion to 1–205 caused a decrease in stimulation, and deletion to 1–172 abolished stimulation. Apparently, sequences beyond the minimum RAP30 binding domain of amino acids 1–172 are required for accurate transcriptional activity, but the observation that more than half the molecule can be deleted without full loss of activity is startling.

In contrast, the extract system is very sensitive to N-terminal deletions. The 74–517 mutant is completely inactive in reconstituting the depleted extract system. Internal deletion mutants Δ136–258 and Δ137–356 are also inactive for transcription, presumably because of deletion of required sequences between amino acids 172 and 205 and deletion of a portion of the RAP30 binding domain between amino acids 136–172.

A Masked RNA Polymerase II Binding Domain on RAP74—The RAP30 subunit of TFII F includes a polymerase binding domain (15, 16), so the question arises of whether RAP74 might also be able to make functional contact with polymerase. To test for direct interaction two binding assays were employed. RNA polymerase II was combined with histidine-tagged RAP74 mutants and tested for retention on Ni²⁺ affinity beads (Fig. 4A), and these results were confirmed by binding RAP74 mutants to affinity beads on which RNA polymerase II was co-valently immobilized (Fig. 4B).

Full-length RAP74 binds to RNA polymerase II, but not as tightly as some mutants from which N-terminal sequences and sequences from the central portion of the molecule have been removed. In the data shown in Fig. 4A, polymerase binding is most easily scored by detection of the 180-, 145-, 36-, 25-, and 18.5-kDa polymerase subunits (39). A polypeptide of 44 kDa that contaminates this polymerase preparation may be actin (39). This and other contaminating proteins are removed by affinity selection of the polymerase on immobilized RAP74. Interestingly, two polypeptides (28 and 30 kDa), not known to be polymerase subunits, co-purify with Ni²⁺ affinity chromatography with RAP74 mutants, indicating that these peptides are tightly associated with polymerase.

The mutants that bind polymerase most tightly are 358–517, 363–517, 363–510, 363–486, 363–452, and 363–444. Further deletion to 363–409 abolishes binding. Sequences between amino acids 363–444, therefore, appear to be the most important for polymerase binding. A 407–517 mutant has slightly reduced binding, and a 1–409 mutant has weak binding, indicating that portions of the interaction site can be found both N-terminal to amino acid 409 and C-terminal to amino acid 407.

Accessibility of this C-terminal domain appears to be masked by sequences within the central region but made more accessible by N-terminal sequences. The most significant masking of polymerase binding is seen with RAP74(87–517) and (207–517). Full-length RAP74(1–517), on the other hand, binds relatively tightly to polymerase. Since RAP74 (358–517) binds tightly to polymerase, sequences between 207 and 358 appear to be responsible for masking polymerase binding by the C-terminal domain. Sequences from 1 to 87 appear to render the polymerase binding domain more accessible.

RAP74 Inhibits Transcription from Non-promoter DNA Sites—Since RAP74 interacts with polymerase, binding might stimulate or inhibit transcription. For the assays shown in Fig. 5, calf thymus polymerase was mixed with a supercoiled plasmid template in the presence of all four nucleoside triphosphates (NTPs). In the absence of a full complement of accessory factors, polymerase initiates from many sites, so initiation is unlikely to be from a natural promoter sequence. Transcription was measured by incorporation of [3H]JUMP into RNA, which was recovered on DEAE filters and washed free of unincorporated [3H]JTP. RAP74 or RAP74 mutants were added to reactions and tested for their effect on polymerase activity.

Full-length RAP74 and some RAP74 fragments were found
to inhibit polymerase activity 90–95% (Fig. 5A). Some RAP74 mutants confer moderate inhibition (20–70% inhibition), and others do not inhibit polymerase at all. Of the RAP74 C-terminal fragments that bind polymerase tightly, 358–517, 363–517, 363–510, and 363–486 inhibit general transcription strongly (>75%). Mutants 363–452 and 363–444, that bind strongly, did not substantially reduce polymerase activity, indicating that sequences from 452 to 486 are important for inhibition. Mutants 87–517 and 207–517, which are reduced for polymerase binding but include the tight binding site, moderately inhibit polymerase. Mutants Δ136–258 and Δ137–356 inhibited polymerase activity substantially (>75%). These mutants include the polymerase binding region but are missing internal sequences that may mask the polymerase binding site. Moderate inhibition was also seen with mutants 1–409, 1–356, 1–296, 1–205, and 136–258, none of which includes the proposed polymerase binding domain from 363 to 444. In binding assays, however, each of these mutants was observed to

FIG. 3. Accurate transcriptional activity of RAP74 mutants. An extract transcription system was depleted of TFIIF by immunoprecipitation with anti-RAP30 and anti-RAP74 antibodies (37, 38). Transcriptional activity was reconstituted by addition of recombinant RAP30 (5 pmol) and RAP74 or RAP74 mutants. The template contained the Adenovirus major late promoter digested at position +217 relative to transcription initiation. The 217-nucleotide runoff transcript was quantitated using a PhosphorImager, and accurate transcription reported as percentage of the highest value determined in the experiment.
Functional Binding of RAP74 C-terminal Region and Polymerase

Fig. 5. RAP74 inhibits general transcription by RNA polymerase II. A, calf thymus RNA polymerase II was mixed with supercoiled plasmid DNA in the presence or absence of RAP74 or RAP74 deletion mutants. ATP, CTP, GTP, and [3H]UTP were added and transcription continued for 1 h. Incorporation of [3H]UMP into RNA was quantitated on DE-81 filters as described under "Experimental Procedures." Transcription initiation is expected to occur from many sites on the DNA template in this experiment. α-Amanitin was added at 1 μg/ml in a reaction otherwise identical to that shown in column 2. Reactions contained 5 pmol of RNA polymerase II and 50 pmol of RAP74 or RAP74 mutant; (*) indicates that 25 pmol of RAP74 mutant was used in these reactions. B, RAP74 inhibits initiation of RNA chains. Calf thymus RNA polymerase II was added to supercoiled plasmid DNA and incubated for 15 min (addition at t = −15 min). ATP, CTP, GTP, and [3H]UTP were then added (t = 0 min), and RNA synthesis was continued for 1 h. RAP74 was added to reactions at t = −15, 0, 0.5, or 1 min. Once initiation occurs, complexes become resistant to inhibition by RAP74. Values are reported as the average of duplicate determinations. The variation between duplicates observed in experiments in A and B was typically less than 10%, always less than 15%, and qualitatively similar with replicate experiments.

interact weakly with RNA polymerase II (Fig. 4 and data not shown), indicating that a second domain may be located in the region between amino acids 136–258 that may interact with polymerase or DNA to inhibit nonspecific transcription. Mutants containing sequence between 87 and 258 can bind DNA, as indicated by a gel mobility shift assay.⁴

RAP74 appears to inhibit initiation of new chains in the nonspecific transcription assay. This is indicated in the experiment in Fig. 5B, in which polymerase was incubated with template DNA for 15 min before addition of NTPs, and transcription was allowed to continue for 1 h. RAP74 was added at the beginning of the preincubation (−15 min), coincident with NTPs (0 min), 30 s after NTPs, or 1 min after NTPs, as indicated in the figure. Since RAP74 gives the greatest inhibition when added before NTPs, inhibition is most likely exerted at the level of chain initiation. When RAP74 is added 1 min after NTPs, no inhibition was observed, indicating that elongation was not inhibited. Since RNA is stable in the presence of RAP74 (addition at t = +1 min), inhibition is not due to a contaminating RNase in the RAP74 preparation.

Interactions between RNA polymerase II, RAP74, and DNA—RAP74 might inhibit RNA polymerase II by preventing association with DNA, or by altering template contacts more subtly so that non-promoter initiation could not occur. To discriminate between these possibilities, a gel mobility shift assay was done to test the effect of RAP74 on polymerase binding to DNA (Fig. 6). Increasing amounts of RAP74 were combined with polymerase and then added to an end-labeled DNA probe. The mixture was then electrophoresed. In this assay polymerase binds nonspecifically to the DNA (21, 22).

The analysis of this experiment is somewhat complicated because both RAP74 and RNA polymerase II can bind to DNA independently, and both proteins shift the mobility of the DNA to similar positions. The native molecular weight of RAP74 is estimated at 470 kDa by gel filtration (38), and this may explain why the mobility shifts induced by RNA polymerase II (~500 kDa) and by RAP74 are so similar. Comparing lanes 7 and 8 to lane 2, essentially all of the polymerase-DNA complex is supershifted by addition of RAP74, under conditions in which RAP74 does not itself efficiently bind to DNA. The implication of this observation is that RAP74 is bound to RNA polymerase II, rather than both proteins binding independently to the same DNA molecule. This interpretation, of course, is consistent with observations made in the binding experiments shown in Fig. 4, in which RAP74 was shown to bind tightly to polymerase in the absence of DNA. RAP74-induced supershifts of the polymerase-DNA complex are clearly seen in lanes 8 and 10. In lanes 12 and 14, a supershifted species with even slower mobility may represent a DNA probe that has bound to both a RAP74 and a RAP74-polymerase complex. Clearly, RAP74 does not block association of polymerase with DNA, nor does it dissociate complexes once they have formed, since polymerase-DNA complexes are stable when RAP74 is added to the reaction.

RNA polymerase II used for this experiment was in the IIb form, so it has an intact C-terminal domain (CTD) (polymerase was the kind gift of R. Burgess). To demonstrate that polymerase was a component of shifted complexes, monoclonal antibodies directed against the CTD were added to the reactions shown in lanes 15 and 16. All of the polymerase-dependent mobility shifts were disrupted by addition of antibody, while those dependent on RAP74 alone were not affected (lanes 15 and 17). Disruption of DNA binding by anti-CTD antibody indicates that the CTD is proximal to bound template DNA. This observation is consistent with models for CTD functions in initiation and elongation of transcription (40). Results of mobility shift experiments were qualitatively very similar with polymerase primarily in the IIb form, that is with the CTD removed by partial proteolysis. As expected, however, anti-CTD antibodies did not affect polymerase IIb-induced mobility shifts (data not shown).

DISCUSSION

The first report of RAP74 mutagenesis was by Yonaha et al. (36). In those studies a two-hybrid gene reporter system was used to map regions of RAP74 and RAP30 that interact. They determined that subunit-subunit contact required RAP74 amino acids 62–171. Based on a direct binding assay, we show here that RAP74 amino acids between 1–172 are required for tight binding to RAP30 (Fig. 2). Deletion of amino acids from

⁴ B. Q. Wang and Z. F. Burton, unpublished data.
the N terminus of RAP74(1–73) severely impairs binding. The 2-hybrid system may be somewhat more sensitive than the affinity bead procedure to demonstrate weak interactions.

Yonaha et al. also tested mutant proteins for the ability to support accurate transcription in vitro, using a system of biochemically fractionated components and recombinant TBP. They concluded that amino acids from 73 to 205 and between 435 and 517 are required for full transcriptional activity, and amino acids between 356–435 were essential. In contrast, two reports from other laboratories have demonstrated different functional properties of RAP74 sequences. For example, the study of Yonaha et al. (41) showed that amino acids between 1 and 206 could support basal accurate transcription in vitro from the human c-fos promoter, but not transcription activated by the serum response factor. They further showed that 1–356 and 1–286 mutants had decreased activity relative to full-length RAP74, indicating the importance of C-terminal sequences. RAP74 sequence between amino acids 206 and 291 was required for full transcriptional activity.

In our study, RAP74 was deleted to 1–205 without full loss of activity, but deletion of N-terminal sequences between 1 and 74 eliminated activity, as observed by Kephart et al.

Accurate transcription involves both initiation and elongation of RNA chains. Several laboratories have presented evidence for RAP74-independent initiation, although most of these reports involve use of supercoiled templates (19, 25–27). Our laboratory has presented evidence that initiation can occur from a linear template in an extract system depleted of TFII F, in the absence of added RAP74 (32). This conclusion, however, has been somewhat controversial, as others using different systems have failed to verify this result. Tan et al. (31) demonstrated that RAP74 was necessary for accurate initiation from a linear template using a transcription system consisting of highly purified and recombinant components. The highly purified system and the TFII F-depleted extract system appear to differ in this respect.

Additionally, our laboratory has shown that RAP74 prevents dissociation of newly initiated major late promoter transcripts. Sequences within both the N-terminal and C-terminal regions of RAP74 contribute to the stability of transcription intermediates. Sequences between amino acids 409 and 517 are required to prevent release of about 50% of complexes formed in extracts. Sequences between 136 and 205 are essential for preservation of the remaining complexes. We find, for instance, that the 1–172 mutant, which appears nearly inactive in runoff transcription (Fig. 3), can support accurate initiation almost as efficiently as full-length RAP74. Newly initiated complexes formed in the presence of the 1–172 mutant, however, are very unstable and dissociate from template. RAP74 may be dispensable for accurate initiation but, in some cases, necessary for the stability of the newly initiated complex. In other systems, additional factors may stabilize complexes in the absence of RAP74, and in this case, the requirement for RAP74 in promoter escape is observed (32). It is our view that in extract systems the most essential functions of RAP74 are for elongation rather than initiation. In early elongation, RAP74 is required to drive polymerase out of a pause close to the promoter. RAP74 is also required to prevent transcript release, although factors present in some extract systems can replace RAP74 for this function (32).

Recent cloning of RAP74 counterparts from Drosophila melanogaster (11) and Saccharomyces cerevisiae (35, 42) allows for comparison of sequences with the expectation that the most important functional domains may also be the most highly conserved during evolution. Functional domains of human RAP74 are shown in Fig. 7A. Multiple sequence alignments are shown that indicate conservation between RAP74 and related proteins in Fig. 7B. These alignments are based on the method of hydrophobic cluster analysis (HCA) (43), which is used to compare weakly similar sequences using visual pattern recognition to indicate the most likely alignment. In HCA plots, the sequence is duplicated and displayed as a two-dimensional representation of an a helix. Hydrophobic amino acids are circled and prolines (stars), glycines (diamonds), serines (boxes with dots), and threonines (boxes) are indicated (Fig. 7C). The most highly conserved sequences in this evolutionary family are similar to the functional domains mapped in this work.

The N-terminal domain between amino acid positions 66 and 187 of human RAP74 is the most highly conserved portion of the molecule. This correlates with our mapping of the RAP30 binding domain between amino acids 1 and 172 (Fig. 2) and the minimal domain for accurate initiation in an extract system.
Fig. 7. Evolutionary conservation of RAP74 functional domains. A, functional domains of human RAP74. Black bars indicate sequences aligned in B. B, multiple sequence alignments comparing human RAP74 (h) (47, 48), Xenopus RAP74 (x) (49), Drosophila Factor 5a (df5a) (11), and S. cerevisiae Ssu71p/Tfg1p (ySsu71) (35, 42). Conserved amino acids are indicated in bold type or by underlining. Potential casein kinase II sites are indicated by larger type “S” or “s” in region II. Stop codons are indicated by *. C, HCA plots (43) are shown for regions I and III of the RAP74 evolutionary family. HCA plots were the basis for the alignments shown in B.
between amino acids 1 and 205 (Fig. 3). We also have unpublished data that indicates a DNA interaction site between amino acids 87 and 258 of RAP74. Since D. melanogaster factor Sa can substitute for human RAP74 in the accurate transcription assay in the extract system (20), many specific amino acid changes are tolerated within this region without full loss of function.

Since accurate transcription assays were done in an extract system, factors that stimulate and inhibit transcription are expected to be present that may not be components of more purified and defined systems. As a result, domains of RAP74 may be dispensable in the extract system that are required for transcription in more defined systems. One explanation for such an observation is that an additional factor in the extract complements or replaces the function of an otherwise necessary RAP74 domain. On the other hand, sequences that are required for function in the extract may be dispensable in more defined systems, for instance, because an inhibitory factor present in the extract requires a RAP74 domain to counteract its effect. If this inhibitor is missing from the more defined system, this domain may become dispensable for transcription.

In a more defined transcription system, amino acids 73–435 were minimally required for transcription (36). Comparison of their results with ours demonstrates that the N terminus of RAP74 is required for transcription in the extract system but not in a more purified system. As we show here, the N terminus of RAP74 is important for RAP30 binding, so a factor present in extracts but missing in more purified systems may make this subunit contact more important for assembly of transcription complexes. Conversely, the C-terminal domain of RAP74 is required for transcription in the purified system but dispensable in the extract system. A positive transcription factor might replace the function of the C-terminal domain of RAP74 or might obviate its requirement by otherwise stabilizing the transcription complex. Since all regulators of transcription have not yet been identified, characterized, and isolated for in vitro use, comparison of purified and extract systems will continue to be of importance to indicate these functions.

From data not shown in this report, we have identified at least two sites at which casein kinase II phosphorylates the central domain of RAP74 in vitro. This central region of RAP74 has the appearance of a regulatory hinge separating N- and C-terminal domains, and perhaps phosphorylation is important in regulating or coordinating the function of these domains. RAP74 is known to be highly phosphorylated in vivo (1). Alkaline phosphatase-treated human TFIIF has decreased affinity for RNA polymerase II and reduced transcriptional activity in a defined transcription system (44). At least two likely casein kinase II sites are present within this sequence (Fig. 7B; region II), and the region that includes these sites is conserved in human, Xenopus, and Drosophila proteins, although most of the surrounding central region is divergent between vertebrates and flies (11).

A partially masked RNA polymerase II binding domain has been located between amino acids 363 and 444 on human RAP74 (Fig. 4). Extending this domain to 363–486 makes it a potent inhibitor of transcription initiation by RNA polymerase II in vitro from non-promoter DNA sites (Fig. 5). Although the precise mechanism of non-promoter initiation from a supercoiled template is not known, these experiments show that binding to this region of RAP74 influences a specific catalytic function of polymerase. General transcription inhibition, therefore, is a measure of the specificity of the interaction between the C-terminal region of RAP74 and RNA polymerase II. The C-terminal domain between 363 and 517 binds to DNA independent of polymerase in a gel mobility shift assay, so DNA interactions are likely a component of C-terminal domain function.

Accurate transcription assays in both extract and purified systems indicate the importance of the C-terminal domain in RAP74 function. In our extract system, using the adenosine major late promoter, RAP74(1–517) has a higher transcriptional activity than RAP74(1–409) (Fig. 3), indicating that this C-terminal domain influences accurate transcription. Using a purified system, Yonaha et al. (36) showed that C-terminal sequences of RAP74 were essential for transcription. In their studies, RAP74(1–435) had decreased activity compared to 1–517. Further deletion to 1–356 abolished activity, demonstrating the importance of the sequence between 356 and 435 in the purified system. Comparing their results to ours, the minimally sufficient sequence for the purified system between 1 and 435 includes most or all of the RNA polymerase II binding domain, but is missing sequences required for inhibition of transcriptional initiation from non-promoter DNA sites (Figs. 4, 5, and 7). Further deletion to 1–356 removes the polymerase binding domain and inactivates the protein for transcription in the purified system. When this mutant is tested in the crude extract system, however, it stimulates transcription, as do more radical deletions to 1–205 (Fig. 3). In experiments utilizing the c-fos promoter, RAP74(1–206) was shown to support accurate transcription in vitro in a system using reconstituted fractions and recombinant components (41), consistent with the results presented here. As mentioned above, deletion to 1–172 supports accurate initiation but not efficient elongation.

The C-terminal region of RAP74 has recently been shown to be important to stimulate a protein phosphatase that dephosphorylates the RNA polymerase II CTD (45). The function of this phosphatase may be to regulate elongation by reducing CTD phosphorylation and/or to promote polymerase recycling after termination, since polymerase enters the preinitiation complex most efficiently in the dephosphorylated state (40). Interestingly, the central region of RAP74 masks phosphatase stimulation, and the presence of the N-terminal region decreases masking, as seen in the general transcription inhibition assay described above. Most likely, stimulation of the CTD phosphatase is mediated through the interaction of the C-terminal region of RAP74 with polymerase.

Since binding of RAP74 inhibits transcription from non-promoter DNA sites, this domain appears to modify template contacts by polymerase. A domain of a transcription factor that interacts with RNA polymerase II and DNA might be expected to regulate contacts between polymerase and template at various stages of the transcription cycle. For instance, the C-terminal domain may help release polymerase from non-specific DNA sites to facilitate termination and polymerase recycling to a promoter. RAP74 is implicated in this function, because TFIIF catalyzes polymerase release from non-specific DNA sites but RAP30 by itself does not (21, 22). Such a role for RAP74 is consistent with its function in stimulating the CTD phosphatase, as discussed above. However, RAP74 by itself is incapable of catalyzing polymerase release (Fig. 6). Apparently, both subunits play a role in this process, and although the C-terminal domain of RAP74 may participate in this function, this conclusion cannot be drawn from the available data. RAP74 has been shown to stimulate elongation by RNA polymerase II in the absence of other transcription factors (20). Potentially, the C-terminal domain of RAP74 could be involved in stimulating polymerization by altering contacts with template. As mentioned above, the C-terminal region contributes to the stability of elongation complexes.

The amino acid sequence between positions 363 and 517 is
conserved between human RAP74, Xenopus RAP74, and to a lesser extent Drosophila Factor 5a and yeast Tfglp/SSu71p (Fig. 7, B and C). Inspection of the human RAP74 polymerase-binding sequence (363–444) does not reveal a domain with significant hydrophobic structure. This sequence is somewhat serine, threonine, proline, and glycine-rich, but the relationship of these characteristics and polymerase binding is unknown. The sequence from 450 to 517 shows conservation between yeast, Drosophila, and vertebrates, and has significant hydrophobic structure (Fig. 7, B and C). Some of this conserved region (444–486) appears to be required for inhibition of non-promoter initiation by RNA polymerase II (Fig. 5), but sequence from 487 to 517, although conserved, is dispensable for these in vitro assays.

These studies are consistent with the model that RAP74 has N- and C-terminal domains separated by a flexible, highly charged hinge domain. The N-terminal domain binds to RAP30 and has functions in assembly of the preinitiation complex, which are particularly noticeable in highly purified systems. The C-terminal domain interacts with RNA polymerase II and appears to modify polymerase contacts with template and with a CTD phosphatase. The central region may regulate C-terminal domain function, as indicated by masking of RNA polymerase II-binding (Fig. 4), inhibition of non-promoter transcription (Fig. 5), and stimulation of CTD phosphatase (45). Consistent with regulatory functions for the highly charged central hinge, this is the site for in vitro phosphorylation by casein kinase II. Interactions between the N-terminal and C-terminal regions may occur, because N-terminal sequences reduce masking by central sequences.

Since RAP74 can bind to RNA polymerase II independently from RAP30, RAP74 can enter preinitiation complexes or remain in elongation complexes in the absence of the RAP30 subunit. Independent function of TFIIF subunits, therefore, may be important in regulated initiation or elongation. For instance, RAP30 can stimulate initiation independent of RAP74 (19, 32). On some promoters, therefore, RAP74 could be supplied to the complex after initiation to commence elongation. RAP74 only weakly stimulates elongation independent of which are particularly noticeable in highly purified systems. The C-terminal domain interacts with RNA polymerase II and has functions in assembly of the preinitiation complex, to the phosphorylated region of RAP74, it will be interesting to determine whether phosphorylation influences regulation by serum response factor. The DNA binding and dimerization domain of serum response factor interacts with the C-terminal region of RAP74 (amino acids 437–517), so this region may receive signals from activators. Since TFII F is involved both in initiation and elongation, this raises the possibility that control can be exerted during multiple stages of the transcription cycle.

Acknowledgments—We thank laboratory members C.-h. Chang for RAP30/74-depleted HeLa extract, S. M. Fang for sequence analysis, and S. Reimers for calf thymus RNA polymerase II in the IIB form. We gratefully acknowledge R. Burgess, C. Wooddel, D. Hager, and N. Thompson for the kind gifts of lla form RNA polymerase II and anti-CTD antibodies.