ABSTRACT The C terminus of mammalian transcription factor RAP30 has been found to be a cryptic DNA-binding domain strikingly similar to the C-terminal DNA-binding domain present in conserved region 4 of members of the σ70 family of bacterial σ factors. This RAP30 domain shares strongest sequence similarity with the DNA-binding domain present in region 4 of Bacillus subtilis sporulation-specific σK. Like the region 4 DNA-binding activity of Escherichia coli σ70, the RAP30 C-terminal DNA-binding activity is masked in intact RAP30 but is readily detectable when the RAP30 C terminus is expressed as a fusion protein. Consistent with a role for RAP30 DNA-binding activity in transcription, mutations that abolish DNA binding also abolish transcription. Therefore, RAP30 may function at least in part through the action of an evolutionarily ancient DNA-binding domain that first appeared prior to the divergence of bacteria and eukaryotes.

Transcription factor TFIIF (also referred to as factor 5 from Drosophila, βy from rat, and RAP30/74 or FC from human cells) is a heterodimer composed of ~30 kDa (RAP30) and ~70 kDa (RAP74) polypeptides (1-5) and has been shown to promote selective binding of RNA polymerase II to the promoters of eukaryotic protein-coding genes (7-10). Previous studies have established that promoter recognition by RNA polymerase II is a multistage process requiring assembly of a nucleoprotein recognition site for polymerase at the promoter, followed by binding of polymerase at this site (11, 12). In the first stage, the TATA factor binds stably to the core promoter to form an "Initial complex," which serves as at least part of the recognition site for polymerase on the DNA. In the second stage, the general initiation factors TFIIB and TFIIF function in concert to promote selective binding of polymerase to the Initial complex.

Results of previous studies revealed that TFIIF shares both structural and functional properties with Escherichia coli σ70 and led to speculation that it may be a eukaryotic counterpart of bacterial σ factors. First, TFIIF appears to function by a mechanism analogous to that utilized by bacterial σ factors. Like E. coli σ70, which expediates binding of E. coli RNA polymerase to promoters by increasing the affinity of polymerase for promoter sequences and reducing the affinity of the enzyme for nonpromoter DNA (13), TFIIF has been shown not only to promote binding of RNA polymerase II to promoters bound by the TATA factor (7-10) but also to suppress formation of nonproductive binary complexes of polymerase and DNA (14, 15).

Second, TFIIF possesses an RNA polymerase II-binding domain similar to the highly conserved RNA polymerase-binding domains present in region 2 of bacterial σ factors. Greenblatt and coworkers identified a RAP30 region exhibiting weak but significant sequence similarity to the RNA polymerase-binding domains of E. coli σ70 and Bacillus subtilis σ0 (16) and, more importantly, obtained evidence that TFIIF is capable of binding to both mammalian and E. coli RNA polymerases through this RAP30 region (17).

Finally, our laboratory recently noted that the C terminus of RAP30 shares sequence similarity with the highly conserved DNA-binding domain present in C-terminal region 4 of bacterial σ factors (18). This ~80 amino acid DNA-binding domain is shared by nearly all σ70 family members and is proposed to interact specifically with the ~35 promoter element to initiate formation of the "closed" complex in the first step of binding of E. coli RNA polymerase to its promoter (19-22). Multiple sequence alignment of σ70 family members revealed that this σ domain consists of two highly conserved regions: an upstream α-helical region designated 4.1, which in some σ factors can be modeled as an amphipathic α-helix, and a downstream region designated 4.2, which shares sequence similarity with the helix-turn-helix DNA-binding domain found in homeodomain proteins and in phage 434 Cro repressor. The C terminus of RAP30 (18) was found to exhibit the strongest similarity to C-terminal regions 4.1 and 4.2 of B. subtilis sporulation-specific σK (spoIVBSpoIIC) (23-26). This finding, together with results from restriction-site protection (7) and phenanthroline-copper complex DNA-cleavage inhibition (8) experiments that are consistent with the possibility that TFIIF could stabilize the preinitiation complex by binding directly to DNA between the TATA box and transcriptional start site, motivated us to investigate the possibility that the C terminus of RAP30 might also be a DNA-binding domain.

A distinguishing feature of the DNA-binding activities present in conserved region 4 of bacterial σ factors is that they are masked in the intact proteins (19-22, 27). Unlike the helix-turn-helix DNA-binding domains of the homeodomain proteins and Cro repressor, for example, the C termini of intact σ factors do not bind DNA in the absence of RNA polymerase. Thus, although genetic evidence had strongly implied that σ factors contain DNA-binding domains that interact with promoter sequences, evidence for a direct, physical interaction between σ region 4 and promoters was lacking until recently, when Dombroski and coworkers showed that glutathione S-transferase (GST) σ70 fusion proteins containing the C terminus of σ70 are capable of binding DNA (27).

Like E. coli σ70, neither TFIIF nor intact RAP30 exhibits detectable DNA-binding activity (9, 10, 14, 15). To test the possibility that the RAP30 C terminus is capable of binding DNA, we used a modification of the strategy of Dombroski et al. (27). Here we describe the results of these experiments, which show that the RAP30 C terminus is a DNA-binding domain with properties similar to those of the C-terminal DNA-binding domain present in conserved region 4 of members of the σ70 family of bacterial σ factors.

Abbreviations: MBP, maltose-binding protein; Ad2, adenovirus type 2; AdMLP, Ad2 major late promoter.

*To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
MATERIALS AND METHODS

Oligonucleotides. Oligonucleotides were as follows: Nde I-1, CACGCAAAAAA-TATGGCGGAGCGC; BamHI-C, ACTGTTTCACACGGATCCCTGCTGC; Stu I-162, GTGCTTAACGCGGTACTCAATTTG; BamHI-244, TCAAGTCGCTGGTCCATGATTTGAATG; BamHI-239, CTTCTTTGGATATCGCTACTGCGTTTC; BamHI-234, CCTGATCTGATCCTACAAGCGGTT; BamHI-219, TACTGCTGAGTCATCAGTCCACATATC; BamHI-219, 0.5 M Tris-HCl (pH 7.9), 20 mM Hepes-NaOH (pH 7.9), 60 mM KCl, 0.2 mM dithiothreitol, 0.5% of bovine serum albumin per ml, and 7.5% of (vol/vol) glycerol. Reaction mixtures were incubated for 20 min at 28°C and applied to 4% polyacrylamide gels containing 3% glycerol and 0.8% of TBE (90 mM Tris/64.6 mM borate/2.5 mM EDTA, pH 8.3). Electrophoresis was carried out until the bromophenol blue in a marker lane reached the bottom of the gel. Gel DNA binding was quantitated by densitometry of autoradiograms with an LKB Ultradens XL laser densitometer or by measurement of radioactivity in specific protein–DNA complexes with an AMBIS radioanalytic scanner. A4 measurements were carried out as described by Carey (29).

Construction of RAP30 Fusion Proteins with C-Terminal Deletion Mutants. DNA fragments encoding RAP30-C-terminal deletion mutants were synthesized by PCR with pET11c-RAP30 (18) as template, digested with Nde I and BamHI, and cloned into the Nde I and BamHI sites of pET-11c. The forward PCR primer for construction of all mutants was Nde I-1. The reverse primers for construction of RAP30-224, RAP30-239, RAP30-244, RAP30-219, and RAP30-209 were BamHI-244, BamHI-239, BamHI-234, BamHI-219, and BamHI-209, respectively.

Electrophoretic Mobility-Shift Assays. Electrophoretic mobility-shift assays were carried out in 15-μl reaction mixtures containing 2 ng of a 5'-32P-labeled double-stranded oligonucleotide containing adenovirus type 2 (Ad2) major late promoter (AdMLP) sequences from Ad2 positions -30 to +10 (28), 20 mM Tris-HCl (pH 7.9), 20 mM Hepes-NaOH (pH 7.9), 60 mM KCl, 0.2 mM dithiothreitol, 0.5% of bovine serum albumin per ml, and 7.5% of (vol/vol) glycerol. Reaction mixtures were incubated for 20 min at 28°C and applied to 4% polyacrylamide gels containing 3% glycerol and 0.8% TBE (90 mM Tris/64.6 mM borate/2.5 mM EDTA, pH 8.3). Electrophoresis was carried out until the bromophenol blue in a marker lane reached the bottom of the gel. Gel DNA binding was quantitated by densitometry of autoradiograms with an LKB Ultradens XL laser densitometer or by measurement of radioactivity in specific protein–DNA complexes with an AMBIS radioanalytic scanner. A4 measurements were carried out as described by Carey (29).

RESULTS

The RAP30 C Terminus Binds DNA. Because neither TFIIIC nor intact RAP30 exhibits detectable DNA-binding activity (9, 10, 14, 15), we used a modification of the strategy of Dombroski et al. (27, 30) to ask whether the RAP30 C terminus is capable of binding DNA. A MBP-RAP30 fusion protein designated RAP30-244 (Fig. 1A), which contains the RAP30 region similar to region 4 of B. subtilis r6, was constructed and purified to >95% homogeneity by amylase affinity and Mono Q chromatography. MRAP-162-249 was found to bind an oligonucleotide containing the core region of AdMLP in electrophoretic mobility-shift assays (Fig. 1B); MBP alone did not exhibit detectable DNA binding activity (data not shown; also see Fig. 3C, lanes 16 and 17). Based on the results of similar experiments carried out under conditions where MRAP-162-249 was present in large molar excess over AdMLP DNA, we estimate that MRAP-162-249 binds this DNA fragment with a A4 ≤ 50 nM.

The C-terminal DNA-binding domain of bacterial σ factors is proposed to promote stable binding of bacterial RNA polymerase to promoters through sequence-specific interactions with the -35 promoter element. If the RAP30 C-terminal DNA-binding domain promotes stable binding of RNA polymerase II to promoters, it might be expected to stabilize the preinitiation complex by binding avidly but not necessarily sequence-specifically to promoter DNA, based on substantial evidence indicating that the TATA factor is primarily responsible for promoter recognition. To determine the specificity of DNA binding by MRAP-162-249, equilibrium competition binding experiments were performed. Binding of MRAP-162-249 to the AdMLP oligonucleotide did not appear to be sequence specific (Fig. 2). However, MRAP-162-249 did exhibit a marked preference for binding to double-stranded DNA, suggesting that DNA binding is not merely a consequence of nonspecific ionic interactions between RAP30 C-terminus and DNA. Double-stranded AdMLP DNA (lanes designated Ad+ strand) is a more efficient competitor than either of the AdMLP single strands (lanes designated Ad- strand). Double-stranded plasmid DNA pBluescript II KS(+) poly(dG:dC), and
poly(dA-dT) competed for binding at least as efficiently as double-stranded AdMLP DNA, whereas the single-stranded DNA homopolymers poly(dT) and poly(dA) (Fig. 2B, lanes 4–11) and poly(dG) and poly(dC) (data not shown) as well as the RNA homopolymers poly(C) and poly(A) (lanes 12–19) were significantly less efficient competitors. It is noteworthy that although poly(dT) and poly(A) competed less efficiently than any of the double-stranded DNAs tested, they competed more efficiently than other single-stranded polynucleotides.

**Fig. 1.** DNA-binding activity of the RAP30 C terminus. (A) RAP30 domain structure, showing regions that are proposed to interact with RAP74 (31) and RNA polymerase II (16, 17) and that are similar to region 4 of B. subtilis σK (18). Pol II, RNA polymerase II; MRAP-(162–249), MBP-RAP30 fusion protein containing amino acids 162–249 of RAP30. (B) Electrophoretic mobility-shift assay of DNA binding by MRAP-(162–249). Binding reactions containing 2 ng of 5'-32P-labeled AdMLP [Ad2(−50 to +10)] DNA and the indicated concentrations of purified MRAP-(169–249) were carried out as described in text. The electrophoretic mobilities of free and bound AdMLP DNA are indicated. ori, Origin.

**Fig. 2.** DNA-binding specificity of the RAP30 C terminus. Electrophoretic mobility-shift assays with 2 ng of 5'-32P-labeled AdMLP DNA, 200 ng of purified MRAP-(169–249), and the indicated molar excess of competitor polynucleotide were carried out as described in text. We note that, in the gel of Fig. 2B, quantitation with AMBIS radioanalytic scanning indicates that the total 5'-32P-labeled AdMLP DNA recovered in lanes 4–7 between the position of the free probe and the top of the gel is essentially the same (within 10% error), suggesting that protein–DNA complexes in lanes 6 and 7 are distributed throughout the lanes and not lost. Lanes: Ad, AdMLP [Ad2(−50 to +10)] DNA fragment; pBSKS+, supercoiled pBluescript II KS(+) plasmid; poly dT, DNA homopolymer poly(dT); poly dA, DNA homopolymer poly(dA); poly C, RNA homopolymer poly(C); poly A, RNA homopolymer poly(A); poly dGdC, double-stranded poly(dG-dC); poly dAdT, double-stranded poly(dA-dT); ori, Origin.

Close Correlation Between the Transcription and DNA-Binding Activities of RAP30 and MRAP30-(162–249) C-Terminus Deletion Mutants. To delimit RAP30 sequences required for DNA binding, a series of deletion mutants of MRAP30-(162–249) was constructed (Fig. 3A) and purified (Fig. 3B). Nearly the entire RAP30 C terminus, including sequences similar to both regions 4.1 and 4.2 of B. subtilis σK, was required for maximal DNA binding (Fig. 3C and Table 1). Neither MRAP-(196–249), which lacks sequences similar to
region 4.1, nor MRAP-(162–209), which lacks sequences similar to region 4.2, bound DNA. The mutant MRAP-(162–244), which lacks the 5 C-terminal amino acids of RAP30, bound DNA as well as MRAP-(162–249) did; however, mutants lacking 10 or more amino acids from the C terminus of RAP30 did not bind DNA efficiently.

If the RAP30 C-terminal DNA-binding domain is important for transcription, deletion of 10 or more amino acids from the C terminus of wild-type RAP30 should dramatically reduce its transcriptional activity. RAP30 mutant RAPΔ244, which lacks the 5 C-terminal amino acids, was as active in transcription as wild-type RAP30 (Fig. 3D, lanes 1 and 2, and Table 1). In contrast, the activity of mutants lacking 10 (RAPΔ239) or more (RAPΔ234, RAPΔ219, and RAPΔ209) C-terminal amino acids of wild-type RAP30 was severely compromised. Thus, the transcriptional activity of RAP30 C-terminal deletion mutants closely parallels the DNA-binding activity of MBP-RAP30 fusion protein C-terminal deletion mutants.

**DISCUSSION**

Members of the $\sigma^{70}$ family of bacterial $\sigma$ factors compose a homologous family of transcription factors (19–22). Multiple sequence alignment of $\sigma^{70}$ family members has identified four highly conserved regions. Among these regions, the two most highly conserved are regions 2 and 4, which are believed to be largely responsible for promoter recognition. Besides the core RNA polymerase-binding domain, conserved region 2 of $\sigma^{70}$ has been shown to contain a DNA-binding domain capable of interacting specifically with the −10 promoter element (27). Conserved region 4, which is located downstream of region 2 in the C terminus of $\sigma^{70}$ family members, has been shown for $\sigma^{70}$ to contain a DNA-binding domain that interacts specifically with the −35 promoter element (27). In addition, glutathione $S$-transferase fusion proteins that include regions 2 and 4 of $\sigma^{K}$, $\sigma^{32}$, and *E. coli* $\sigma^{E}$ have been shown to bind promoter DNA (30), consistent with the proposal that regions 2 and 4 of these $\sigma$ factors function similarly to the DNA-binding domains of $\sigma^{70}$. The DNA-binding activities of $\sigma^{70}$, $\sigma^{K}$, and $\sigma^{32}$ are masked by N-terminal amino acids in the intact factors but are readily detectable when regions 2–4 are expressed as glutathione $S$-transferase fusion proteins (27, 30).

RNA polymerase II general transcription factor TFIIF is a heterodimer composed of ~30-kDa (RAP30) and ~70-kDa...
Table 1. Transcription and DNA-binding activities of RAP30 and MRAP-(162-249) fusion protein and its C-terminal deletion mutants

<table>
<thead>
<tr>
<th>C-terminal deletion, aa</th>
<th>DNA binding</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>None 162-249</td>
<td>++ +</td>
<td>RAP30 ++ +</td>
</tr>
<tr>
<td>5 162-244</td>
<td>++ +</td>
<td>RAP30A244 ++ +</td>
</tr>
<tr>
<td>10 162-239</td>
<td>++ +</td>
<td>RAP30A239 +</td>
</tr>
<tr>
<td>15 162-234</td>
<td>++ +</td>
<td>RAP30A234 +</td>
</tr>
<tr>
<td>30 162-219</td>
<td>++ +</td>
<td>RAP30A219 +</td>
</tr>
<tr>
<td>40 162-209</td>
<td>++ +</td>
<td>RAP30A209 +</td>
</tr>
</tbody>
</table>

The data in this table were obtained by densitometry of appropriate exposures of autoradiograms from the electrophoretic mobility shift and transcription assays shown in Fig. 3. DNA-binding activity of MBP-RAP30 fusion proteins was assayed by electrophoretic mobility shift assay as described in text with 2 µg of each protein. +++, Less than 5% of the total, Ad2-(-50 to +10) fragment (AdMLP) electrophoresed as free DNA, and >95% of the Ad2-(-50 to +10) fragment was detected in the discrete DNA-protein complex labeled “bound” in Fig. 3C: +, 50-60% of the total Ad2-(-50 to +10) fragment electrophoresed as free DNA, and <5% of the Ad2-(-50 to +10) fragment was detected in the discrete DNA-protein complex labeled “bound;” and -, >95% of the Ad2-(-50 to +10) fragment electrophoresed as free DNA. The transcript activity of wild-type RAP30 and its deletion mutants was assayed as described in text with 3 ng of each protein. +++, More than 95% of RAP30 transcription activity; +, 5-10% of RAP30 transcription activity; and -, <1% of RAP30 transcription activity. aa, Amino acids.

(RAP74) polypeptides (1–6). On the basis of a variety of evidence indicating (i) that TFIIIF promotes interactions of the RNA polymerase II transcriptional apparatus with promoter sequences between the TATA box and transcriptional start site (7, 8) and (ii) that the RAP30 C terminus shares statistically significant sequence similarity with the DNA-binding domain present in C-terminal region 4 of members of the σ family of bacterial σ factors (18), we investigated the possibility that the RAP30 C terminus might be a DNA-binding domain. In this regard, we show that, although neither TFIIIF nor intact RAP30 exhibits detectable DNA-binding activity (9, 10, 14, 15), an MBP-fusion protein containing the 88 C-terminal amino acids of RAP30 is capable of binding DNA. Consistent with a role for RAP30 DNA binding in transcription, we observe a close correlation between the transcription and DNA-binding activities of C-terminal deletion mutants of RAP30 and MBP-RAP30 fusion proteins.

Whereas the DNA-binding domains of E. coli σ20 bind preferentially to promoter sequences, the MBP-RAP30 fusion protein appears to bind with comparable affinity to both promoter and nonpromoter sequences. A lack of sequence-specific DNA binding by the RAP30 C terminus is not unexpected in the light of substantial evidence indicating that the TATA factor (TFIIF or TBP) is the primary determinant of promoter specificity in assembly of the RNA polymerase II preinitiation complex (11, 12). Indeed, results of restriction-site protection (7) and phenanthroline–copper complex DNA-cleavage inhibition (8) experiments suggest that TFIIIF promotes formation of protein–DNA contacts in promoter sequences that are located between the TATA box and transcriptional start site and that are not highly conserved among RNA polymerase II promoters. On the basis of these results and the recent findings of Greenblatt and colleagues, who observed that RAP30 in the preinitiation complex can be cross-linked to this region of the AdMLP (32), we believe it is reasonable to propose that TFIIIF could function to stabilize binding of the RNA polymerase II transcriptional apparatus to promoters through nonspecific binding of the RAP30 C terminus to promoter sequences between the TATA box and transcriptional start site.

In summary, the C terminus of RAP30 has been found to be a cryptic DNA-binding domain similar to region 4 of bacterial σ factors. Taken together with previous results suggesting that RAP30 sequences upstream of this C-terminal DNA-binding domain function as an RNA polymerase II-binding domain similar to the core RNA polymerase-binding domain in σ conserved region 2 (16, 17), the findings presented here reveal an underlying RAP30 architecture similar to that of σ family members. The coincidence of these unique DNA-binding domains in RAP30 and bacterial σ factors suggests that they may play an evolutionarily conserved role in messenger RNA synthesis.

We thank R. Kornberg, R. Losick, P. Silverman, J. Helmann, D. Robison, G. Zurawski, and S. Zurawski for helpful discussions. We also thank K. Jackson of the Molecular Biology Resource Center of the Oklahoma Center for Molecular Medicine for oligonucleotide synthesis. This work was supported by Grant GM41628 from the National Institutes of Health and by Contract HR0-011/3916 from the Oklahoma Center for the Advancement of Science and Technology.