THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 38, Issue of September 19, pp. 23921–23929, 1997
Printed in U.S.A.

Characterization of UEF-4, a DNA-binding Protein Required for Transcriptional Synergism between Two AP-1 Sites in the Human Urokinase Enhancer*

(Received for publication, March 28, 1997, and in revised form June 13, 1997)

Dario De Cesare‡, Martina Palazzolo‡, Jens Berthelsen‡§, and Francesco Blasi‡§¶

From the ‡Dipartimento di Genetica e Biologia dei Microrganismi, University of Milan, and §DIBIT, H. S. Raffaele Scientific Institute, via Olgettina 58, 20132 Milan, Italy

The enhancer of the inducible urokinase gene depends on three essential but not sufficient transactivating elements, an upstream PEA3/AP-1A and a downstream AP-1B site. Enhancer activity also requires the interposed 74-base pair-long cooperation mediator (COM) region that allows transcriptional synergism between the transactivating sites. The 5′-half of COM (uCOM) forms four retarded complexes with HeLa or Hep-G2 nuclear proteins (UEF-1–4). We have identified the binding sequence for UEF-4 and generated uCOM elements uniquely mutated in the UEF-4-binding site or uniquely binding UEF-4. Introduction of these and other mutations in the context of the urokinase enhancer showed that all uCOM sites are important for enhancer activity but that UEF-4 and UEF-1 plus UEF-2/3 can substitute for each other, suggesting functional redundancy of urokinase enhancer factors. UEF-4 was purified from HeLa nuclear extract by affinity chromatography and shown to contain two polypeptides of 105 and 65 kDa, respectively, of which at least the former was endowed with DNA binding activity.

Enhancer sequences represent the ultimate target of signal transduction pathways that lead to induction of specific sets of genes. Although transcription can be activated by a large variety of stimuli and despite the presence in many enhancers of the same or of similar protein-binding motifs, only a subset of genes is switched on in response to a specific signal, implying that very fine mechanisms regulate the specificity of enhancer function.

Urokinase-type plasminogen activator (uPA)[1] is a serine protease important in fibrinolysis and cell recruitment and hence in a variety of processes requiring cell migration, *e.g.* inflammation and cancer (1–5). Therefore, uPA synthesis is induced in a variety of pathological or experimental conditions in many different cells. In culture, induction or modulation of uPA synthesis was first observed with phorbol ester PMA (6) and then reproduced in many cell lines with a large number of growth factors, hormones, and differentiation factors (7). On the other hand, in the embryo, uPA is expressed constitutively in the extraembryonic trophoblast (8, 9) and in the adult in the kidney, lung, and scattered fibroblastic cells (10).

The 5′-flanking regulatory region of the human uPA gene has been studied in some detail. The promoter shows the presence of a typical TATA-box, is rich in Sp1 sites, and is strongly activated by an enhancer located about 2,000 base pairs upstream of the transcription start site (11). The minimal uPA enhancer region contains an upstream combined PEA3/AP-1A (octameric) site and a downstream heptameric AP-1B site (see Fig. 1). All three such sites are important for induction of uPA gene transcription by a variety of extracellular stimuli such as PMA, epidermal growth factor, okadaic acid, and cytoskeleton disruption; the inactivation of only one of the three sites results in the loss of enhancer function (12–15). Synergism between PEA3/AP-1A and AP-1B depends on the integrity of the 74-base pair region, called COM (cooperation mediator), that separates these sites (13). Synergism requires a series of protein-binding sites, clustered in a specific bipartite uCOM-dCOM arrangement (Fig. 1) (1, 13, 14). In fact, even in the presence of intact PEA3/AP-1A and AP-1B elements, disruption of all COM protein-binding sites leads to inactivation of the enhancer function (13). However, COM has no direct transactivation function, and its mechanism of action is still mostly unknown. It has been shown that uCOM and dCOM each contributes about 50% of the activity of the whole region and that COM action is position-dependent but orientation-independent (1). Furthermore, when the 17-base pair uCOM-DNA is incubated with a HeLa nuclear extract, four DNA-protein complexes (urokinase enhancer factors; UEFs) are formed, UEF-1, UEF-2/3, and UEF-4 (13). The binding site for UEF-2/3 has been identified (16), and its sequence is shown in Fig. 1. The dCOM region binds two factors that have not yet been analyzed.

Sequences homologous to the uPA COM regions are present and shown to be active also in promoters other than uPA, like interleukin-3, LD78, stromelysin, and other AP-1-regulated genes (16–20). Moreover, purified UEF-2/3 has been shown to bind not only to the uPA uCOM but also to the homologous region of the interleukin-3 promoter (16).

As a further step in characterizing the COM region and the COM-binding factors, we have identified the UEF-4-binding site, studied its function in the uPA enhancer, and purified UEF-4. The results show that the UEF-4 site is very similar to a negative regulatory region in the AP-1-dependent LD78/MIP-1α chemokine, a PMA-regulated gene, and is conserved in the regulatory regions of a variety of other genes. Moreover, the UEF-4 and the UEF-1 plus UEF-2/3 sites are all necessary for enhancer activity and can substitute for each other. Finally, we

* This work was supported by grants from the Associazione Italiana Ricerche sul Cancro and Consiglio Nazionale delle Ricerche (PF ACRO, Applicazioni Cliniche della Ricerca Oncologica). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) X12641.

¶ To whom correspondence should be addressed: DIBIT H.S. Raffaele, via Olgettina 58, 20132 Milan, Italy. Tel.: 39 2 2643 4832; Fax: 39 2 2643 4844.

[1] The abbreviations used are: uPA, urokinase-type plasminogen activator; PMA, phorbol 12-myristate 13-acetate; COM, cooperation mediator; uCOM, 5′-half of COM; dCOM, 3′-half of COM; UEF, urokinase enhancer factor; CAT, chloramphenicol acetyltransferase; EMSA, electrophoretic mobility shift assay; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis.

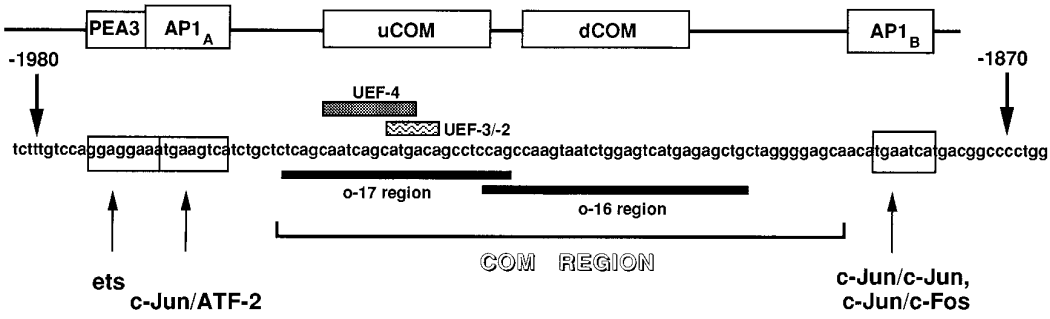

FIG. 1. **Protein-binding sites in human uPA enhancer.** At the *top*, the various functional regions are depicted. At the *bottom*, the transcription factors binding to the PEA3/AP-1A and AP-1B sites are shown. In the *middle*, the *filled bars* in the COM region (o-17 and o-16 regions) show the sequences utilized for bandshift analysis of the two areas. The *bars above* the uCOM sequence identify specific protein recognition sequences in uCOM. The UEF-3/2 recognition sequence has been described before (16), while the UEF-4 recognition sequence is the subject of this paper.

have purified UEF-4 protein and begun its characterization. The purified factor binds both the uPA UEF-4 and the LD78/MIP-1α sequences and is made up of 105- and 60-kDa polypeptides, the former endowed with DNA binding activity.

MATERIALS AND METHODS

Cell Culture and Transfection Analysis

HepG2 cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum. Approximately 1–2 $\times$ 10^6 cells were electroporated with 3 μg of the reporter DNA construct plus 27 μg of carrier Bluescript plasmid (21) and 0.5 μg of CMVβ-GAL plasmid as internal control, for a total of 30.5 μg of DNA. Cells were electroporated in 0.5 ml of complete medium, 250 V and 960 microfarads, using a Gene Pulser apparatus (Bio-Rad). Cells were plated in Dulbecco's modified Eagle's medium containing 10% fetal calf serum and allowed to attach overnight. For PMA induction, the medium was changed to Dulbecco's modified Eagle's medium containing 0.5% fetal calf serum and 100 ng ml^{-1} PMA. Control cells received equivalent amounts of Me_2SO, the solvent for PMA. Cell extracts were divided in two aliquots, of which one was analyzed by a β-galactosidase assay (21). The remaining aliquot was heated at 65 °C and assayed for CAT activity by diffusion of reaction products into scintillation fluid (22).

Electrophoretic Mobility Shift Assays (EMSAs)

Preparation of HeLa cell nuclear extracts was carried out according to Ref. 23, with minor modifications (16). The NaCl concentration of samples resuspended in buffer C was adjusted to 0.4 M with 5 M NaCl. The protease inhibitor aprotinin (1 mg ml^{-1}) was added to buffers C and D. The protein concentration of nuclear extracts was determined by a Bio-Rad protein assay. ^{32}P-Labeled oligonucleotides (T4-polynucleotide ligase) were used as probes. Gel retardation reactions were carried out in a 20-μl volume containing 20,000–30,000 cpm probe and 2 μg of poly(dI-dC). Nuclear extracts were incubated in 50 mM Tris-HCl, pH 7.9, 12.5 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 20% glycerol with poly(dI-dC) and competitor DNA as indicated for 10 min at 20 °C. For EMSA analysis of the UEF-4 purified protein 50 μg of bovine serum albumin and 0.1 μg of poly(dI-dC) were used in binding reactions. Probe was added, and the incubation continued for 10 min. Samples were then run on 5% native polyacrylamide gels (30:1 in 0.25 $\times$ TBE; 1 $\times$ TBE: 89 mM Tris, 90 mM boric acid, and 1 mM EDTA).

UV Cross-linking Analysis

Cross-linking analysis was carried out as described before (24). EMSA reactions were scaled up by increasing 5-fold the amount of nuclear extract and bromodeoxyuridine-substituted probes used in binding reactions, and the reaction products were separated as described above. The wet gel was UV-irradiated at 364 nm for 30 min and autoradiographed at 4 °C. The bands corresponding to DNA-protein complexes were cut out and equilibrated in 100 mM Tris-HCl, pH 6.8, 50 mM DTT, and 2% SDS before loading on a 10% SDS-polyacrylamide gel. Molecular weight markers were obtained from Amersham Corp.

Methylation Interference Analysis

Methylation interference with DMS was performed essentially as described (25), with minor modifications, under experimental conditions that preferentially reveal contacts at sequences containing G residues in the major groove of the DNA double helix. Briefly, either the top or bottom strand of o-17Δ2 oligonucleotide was labeled by T4 polynucleotide kinase and then annealed with the unlabeled complementary strands. Probes were methylated by dimethyl sulfate, used in combination with nuclear extracts in 5-fold scaled-up binding reactions, and DNA-protein complexes were separated on native gel electrophoresis (see above). Bands corresponding to free and retarded probe were cut out. The DNA was recovered by electroelution, phenol-extracted, ethanol-precipitated, and subjected to piperidine cleavage. The products were analyzed by 15% polyacrylamide gel electrophoresis in 8 M urea.

Plasmid Constructions

DNA manipulations were carried out by standard techniques (21), and plasmid structures were verified by DNA sequencing. All constructs were generated by cloning polymerase chain reaction-produced fragments into pBLCAT2 (26) previously cut by *Sal*I and *Bam*HI. Specific point mutations were obtained by polymerase chain reaction using primers that introduced the desired mutation. Following, the template and the oligonucleotide used as direct primer in polymerase chain reaction for each construct are shown (the mutated nucleotides are indicated in boldface lowercase letters). The reverse primer was the same for all constructs: 5′-TGGGCGGGCCGGATCCTCT-3′.

WTCAT—The −1977/−1858 fragment was polymerase chain reaction-amplified from the plasmid containing the whole uPA promoter (−2345/+30) (13) by using specific primers (27). The WTCAT template was used for constructing DCAT. The direct primer was 5′-CAGGTCGACTCTAGAGGAAATGAAGTCATCTGCTCTCAGCAATCAGC**t**T-**c**ACAGCCTCCAGC-3′.

FIECAT—In this case, the template was the plasmid FIEuPA (13), with the same primers used for the WTCAT.

ECAT—The template was the plasmid EuPA (13), with the same primers used for the WTCAT. This plasmid was used as a template for constructing DECAT, and the direct primer was the same as for the DCAT plasmid (Ref. 8).

Δ2CAT—The template was the plasmid WTCAT, and the direct primer was CAGGTCGACTCTAGAGGAAATGAAGTCATCTGCTCTCAGCAATCAGCATGACCCTCCAGC.

Δ2DCAT—The template was the plasmid Δ2CAT, and the direct primer was CAGGTCGACTCTAGAGGAAATGAAGTCATCTGCTCTCAGCAATCAGCTTCACCCTCCAGC).

Δ2ECAT—The template was ECAT (see above); the direct primer was the same as for Δ2DCAT (see above).

G2CAT—The template was WTCAT2, and the direct primer was CAGGTCGACTCTAGAGGAAATGAAGTCATCTGCTCTCAGCAATCATCATGACAG.

The DCAT plasmid has been previously described (1).

Preparation of Nuclear Extracts and Purification of UEF-4

For purification of UEF-4, nuclear extracts (319 mg of total proteins) in buffer C (23) were subjected to ammonium sulfate precipitation at 55% saturation. The precipitate was resuspended in one-quarter of the original volume of TK100 buffer (25 mM Tris-HCl, pH 8.0, 100 mM KCl, 200 mM EDTA, 20% glycerol, 1 mM DTT, 1 mM $Na_2S_2O_5$, 0.5 mM phenylmethylsulfonyl fluoride) and dialyzed against the same buffer. The material was loaded on a Q-Sepharose HP column (55 ml, Pharmacia Biotech Inc.), which was washed with the same buffer. The flow-through, containing 95% of UEF-4 binding activity and only 18% of

TABLE I
Sequence of oligonucleotides

Sequences begin from the 5′ end, and only the coding strand is shown. The sequence in boldface type identifies the UEF2/3-binding site (**TGACAG**) or remaining parts thereof. Lowercase letters identify substituted nucleotides with respect to the o-17 sequence. Underlined nucleotides identify substituted nucleotides with respect to the o-17Δ2 sequence. A hyphen identifies a deleted base.

CAGCAATCAGCA**TGACAG**CCTCCAGC	o-17wt (uCOM)
TGCTCTCAGCAATCAGCt**Tc****ACAG**CCTCCAGC	o-17D
TGCTCTCAGCAATCAGCt**TGACAG**CCTCCAGC	o-17D1
TGCTCTCAGCAATCAGC**AT**c**ACAG**CCTCCAGC	o-17D2
TGCTCTCAGCAATCAGC**Ag****GACAG**CCTCCAGC	o-17K
CAGCAATCAGCA**TGACA-**CCTCCAGC	o-17Δ1
CAGCAATCAGCA**TGAC--**CCTCCAGC	o-17Δ2
CAGCAATCAGCt**Tc****AC--**CCTCCAGC	o-17Δ2D
CAGCAATAAGCA**TGAC--**CCTCCAGC	o-17Δ2G1
CAGCAATAATCA**TGAC--**CCTCCAGC	o-17Δ2G2
CAGCAATAAGAA**TGAC--**CCTCCAGC	o-17Δ2G3
CAGCAATCAGCA**TGA---**CCTCCAGC	o-17Δ3
CAGCAATCAtCA**TGA---**CCTCCAGC	o-17Δ3G2

the total protein (5-fold purification) was dialyzed against H2K150 buffer (25 mM HEPES, pH 7.9, 150 mM KCl, 200 mM EDTA, 1 mM $MgCl_2$, 20% glycerol, 1 mM DTT, 1 mM $Na_2S_2O_5$, 0.5 mM phenylmethylsulfonyl fluoride) and fractionated through heparin-Sepharose CL-6B (HiLoad 26/10 column, 40 ml, Pharmacia). The column was washed with H2K150 buffer and eluted with 90 ml of a 150–660 mM KCl gradient in H2 buffer (25 mM HEPES, pH 7.9, 200 mM EDTA, 1 mM $MgCl_2$, 20% glycerol, 1 mM DTT, 1 mM $Na_2S_2O_5$, 0.5 mM phenylmethylsulfonyl fluoride). Fractions with the highest UEF-4 activity (screened by EMSA) were pooled and dialyzed against H2K150. The pool was passed over a column containing the mutated UEF-4-binding site DNA (o-17Δ2D; see Table I for sequence) to eliminate nonspecific DNA-binding proteins; the column was prepared by coupling a streptavidin-agarose resin (3.2 ml, Pierce) to the 5′-biotinylated (top strand) double-stranded o-17Δ2D oligonucleotide. The flow-through, containing 96–98% of the loaded UEF-4 activity, was finally passed over a o-17Δ2-streptavidin-agarose column. The o-17Δ2-streptavidin-agarose column was prepared by coupling 1.6 ml of streptavidin-agarose to the biotinylated o-17Δ2 oligonucleotide (UEF-4-binding site). The column was washed with H2K150 buffer and eluted with 15 ml of a 200–575 mM KCl gradient in H2 buffer. Fractions were screened by EMSA and analyzed by SDS-polyacrylamide gel electrophoresis, followed by silver staining. Samples for electrophoresis (1 μl) were mixed with an equal volume of 2 × loading buffer (20 mM Tris-HCl, pH 8.0, 5% SDS, 2 mM EDTA, 10% β-mercaptoethanol, 0.02% bromphenol blue), heated at 100 °C for 5 min, and resolved on 10–15% gradient gels by using the Phast System (Pharmacia). Pharmacia low molecular weight markers were used as protein standards.

RESULTS

The UEF-binding Sites Overlap—The uCOM oligonucleotide (o-17, see Table I for sequence) forms with HeLa nuclear extracts four retarded complexes (UEF1–4) (13) in electrophoretic mobility shift assays (EMSAs), of which UEF-2/3 bind specifically the TGACAG sequence (16). A substitution of two nucleotides, one of which lies within the TGACAG sequence (o-17D, Table I), prevents the formation of all four complexes (1), indicating that these factors might bind to overlapping sequences or that the four bands (or some of the bands) represent multimers of the same protein(s). Since we have previously characterized the UEF2/3-binding site (16), we used a UEF-2/3-specific affinity resin (U3BE-agarose, see "Materials and Methods") to deplete a nuclear extract of this activity and tested the depleted extract for UEF-1 and UEF-4 binding activities. The U3BE resin has been previously shown to purify to homogeneity the UEF2/3 complexes (16). The U3BE-treated extracts were used in EMSA with a labeled uCOM oligomer as a probe (o-17). These extracts showed a specific loss of UEF-2/3, but not of UEF-1 and UEF-4 (Fig. 2). Control DNA affinity resin containing a mutated TGACAG sequence (U3BEmut, see "Materials and Methods") failed to deprive nuclear extracts of UEF-2/3 factors, demonstrating the specificity of the observed

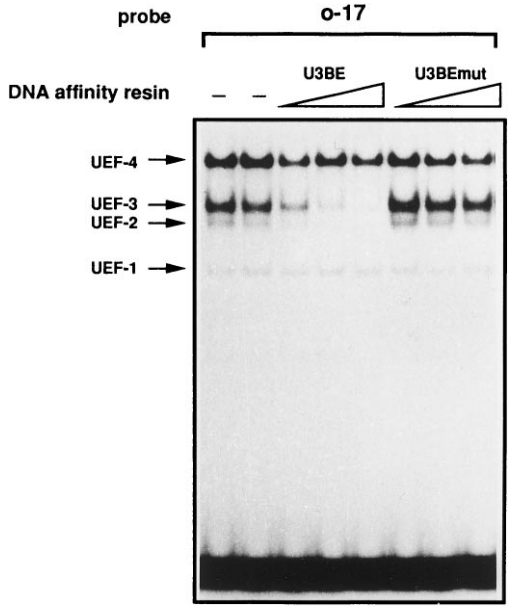

FIG. 2. **UEF-1 and UEF-4 bind uCOM independently of UEF2/3.** EMSA was carried out using ^{32}P-labeled o-17 oligonucleotide and 10 μg of nuclear HeLa cells extracts in the absence (− *lanes*) or in the presence of the U3BE or U3BEmut affinity resin (see "Materials and Methods" for sequences). Different amounts of resin were used (5, 10, or 15 μl, respectively). 2.5 fmol of probe was added in each case. The sequence of o-17 is reported in Table I.

depletion (Fig. 2). We conclude that UEF-1 and UEF-4 binding activities are independent of the presence of UEF-2/3 and hence may be due to different proteins. Moreover, the data of Fig. 2 also indicate that UEF-1 and UEF-4 have a DNA sequence specificity distinct from that of UEF-2/3.

Identification of the Binding Sequence for UEF-4—To analyze the DNA binding specificity of UEF-4, oligomers with a mutated uCOM were employed as competitors in binding assays. We began the analysis by deleting one, two, and three nucleotides, respectively, at the 3′-end of the TGACAG consensus in the o-17 (uCOM) sequence, generating the o-17Δ1, o-17Δ2, and o-17Δ3 oligonucleotides (see Table I for sequences). As shown in Fig. 3 *A*, the o-17Δ1 oligonucleotide competed as efficiently as the wild type o-17 for the formation of all UEF complexes. However, mutation o-17Δ2, while still able to compete for UEF-4, no longer competed for UEF-1 and UEF-2/3. On the other hand, o-17Δ3 competed for UEF-1 and UEF-4, but no longer for UEF-2/3. Oligonucleotide o-17D, shown for comparison, was unable to compete for any of the complexes. These results were confirmed by direct binding studies; labeled o-17Δ2 oligonucleotide gave rise to a single retarded complex, which co-migrated with the UEF-4 complex (Fig. 3, *B*). This band was competed for by unlabeled wild type o-17 oligonucleotide to the same extent as the UEF-4 complex formed by the wild type o-17 probe. Thus, the complex formed by o-17Δ2 behaved like UEF-4 not only in electrophoretic mobility but also in binding specificity. Labeled o-17Δ3 oligonucleotide gave rise to two retarded complexes exhibiting the same relative mobility and the same binding specificity of UEF-1 and UEF-4 (data not shown).

To better define the UEF-4-binding sequence, we used methylation interference analysis with o-17Δ2, which only binds UEF-4 with the same affinity of o-17 (see Fig. 3*B*). Partially dimethyl sulfate-modified o-17Δ2 probes, labeled on either strand, were used in scaled-up binding reactions. Fig. 4 shows a representative experiment. In the UEF-4-bound DNA, two G residues on the bottom strand and one on the top strand became totally unreactive to piperidine treatment (*filled circles*).

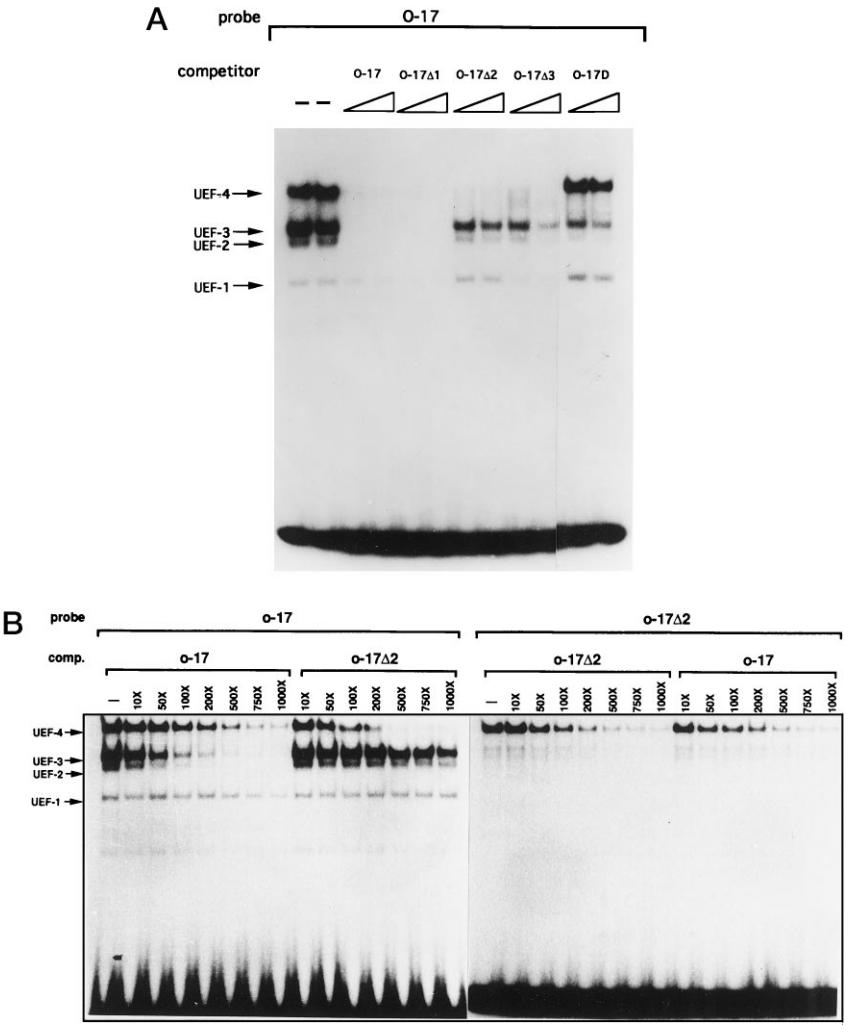

FIG. 3. **Mutant o-17Δ2 oligonucleotide binds only UEF-4.** EMSA analysis of o-17 and o-17Δ2 oligonucleotides (see Table I) with 5 μg of HeLa nuclear extracts and 2.5 fmol of labeled probe. *A*, deletions in the o-17 oligonucleotide affect its ability to compete for UEF factors binding. Competitors were used in 200- and 600-fold excess. *B*, comparison of direct binding and competing activity of o-17 and o-17Δ2 oligonucleotides. The excess of unlabeled competitors is indicated *above* each *lane*.

In addition, a weak protection was observed on one G residue on the top strand (*open circle*). These data indicate that UEF-4 binds DNA mostly outside of the UEF-2/3-binding site but that one weak contact also occurs with the first guanine of the TGACAG sequence. Thus, the UEF-2/3-binding site appears to partially overlap with the UEF-4-binding site.

To support these findings, the three G residues totally protected in the methylation interference assay in the o-17Δ2 oligonucleotide were individually mutated to T residues. The ability of the unlabeled mutated o-17Δ2G1, -G2, and -G3 oligonucleotides to compete for UEF-4 formation was tested by EMSA (data not presented) and showed that o-17Δ2G1, o-17Δ2G2, and o-17Δ2G3 (see Table I for sequences) were significantly less effective than the o-17Δ2 oligomer in inhibiting UEF-4 complex formation, confirming the results obtained by methylation interference analysis. As a negative control, mutation D was inserted in the context of the o-17Δ2 sequence (o-17Δ2D; see Table I). Densitometric quantitation of the data indicated that an at least 4-fold higher concentration of mutated oligomers was needed to achieve 50% inhibition with respect to o-17Δ2 (Fig. 5).

Taken together, the results of Figs. 4 and 5 strongly support the conclusion that the G residues, whose methylation prevented binding, were involved in contacts between DNA and UEF-4. However, the methylation interference analysis was carried out in a mutated context of the uCOM region, the o-17Δ2 sequence, which is only permissive for the binding of UEF-4. Therefore, mutation G2 was introduced in the o-17 wild type sequence. The mutated uCOM oligomer (o-17G2, Table I) was assayed both in binding and in competition studies. As shown in Fig. 6*A*, when o-17G2 was used as a competitor with o-17 and o-17Δ2 probes, it uniquely failed to compete for the formation of UEF-4 complex. When used as a probe, the o-17G2 mutation still bound UEF-1 and UEF-2/3, but not UEF-4. In addition, labeled o-17G2 was incapable to form a UEF-4 complex with HeLa nuclear proteins, while normally forming the other complexes (Fig. 6*B*); in this experiment, the different intensity of the UEF bands shifted by the wild type and the o-17G2 oligonucleotides was due to different specific activities of the probes and not to weaker affinity. When G1 and G3 mutations were inserted into the o-17 wild type sequence, similar results were obtained (data not shown). We conclude that the three G residues identified by methylation interference do not significantly interact with UEF factors other than UEF-4. Thus, the sequence CAGC, immediately 5′ of TGACAG, is specific for UEF-4 and has no overlap with other UEF-binding sites.

To define the 5′-boundary of the UEF-4 site, we constructed 5′-deleted oligonucleotides and tested their binding and competing activity. As shown in Fig. 7 (*A* and *B*), the shortened oligonucleotides U4a and U4b bound to and competed with UEF-4 as efficiently as the o-17Δ2 sequence. On the other hand, U4c and U4d showed a reduced or an absent UEF-4 binding activity, respectively, setting a 5′-limit for the UEF-4-binding site (5′-AATC . . .).

The presence of a weak protected G residue in methylation interference suggests that the binding sites for UEF-2/3 and UEF-4 overlap. In fact, substitution of A and G in the CATGAC

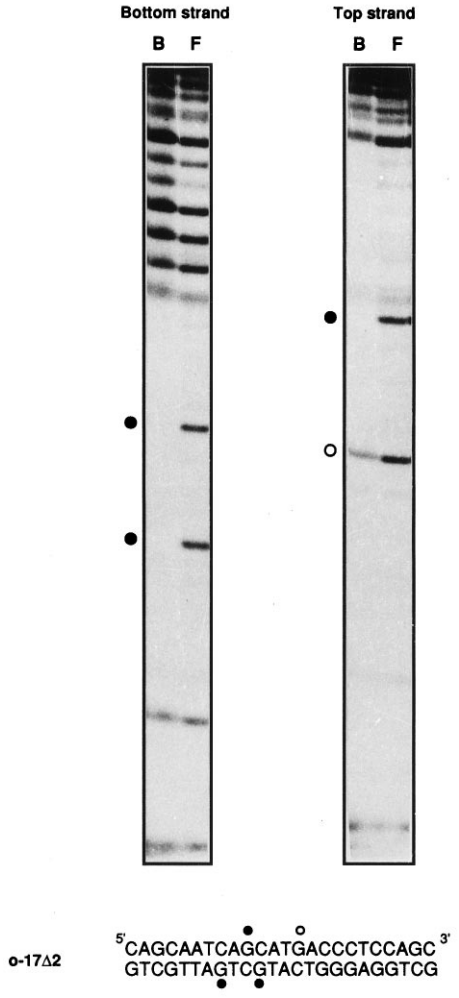

FIG. 4. **Methylation interference analysis of the UEF-4 complex.** UEF-4-bound or -free ^{32}P-labeled o-17Δ2 were isolated from a larger scale EMSA and treated for methylation interference as indicated under "Materials and Methods." *Bottom strand* indicates the antisense, and *top strand* indicates the sense strand of the oligonucleotide as present in the uPA enhancer sequence. *B* and *F* refer to UEF-4-bound and -free o-17Δ2 oligonucleotides. *Filled circles* show G residues whose methylation is strongly interfered with by UEF-4 binding. *Empty circles* show residues with weak interference. The sequence of the oligonucleotide is reported at the *bottom* along with the effect of methylation.

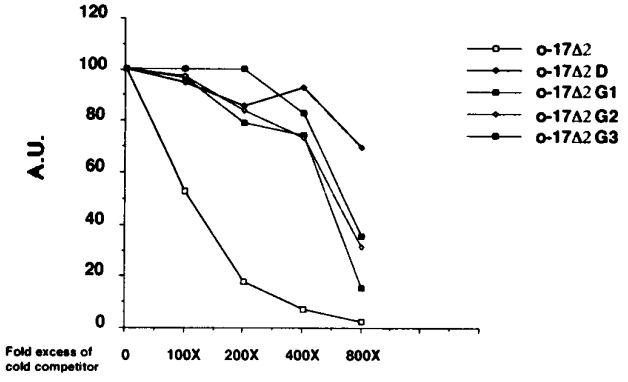

FIG. 5. **Role of the protected G residues in UEF-4 binding.** EMSA was carried out with 5 μg of HeLa nuclear extract and 2.5 fmol of ^{32}P-labeled o-17Δ2 probe. Competitors were used at 100-, 200-, 400-, and 800-fold excess. Quantitation was obtained through densitometric analysis. *A.U.*, arbitrary units. The sequences of the oligonucleotides are shown in Table I.

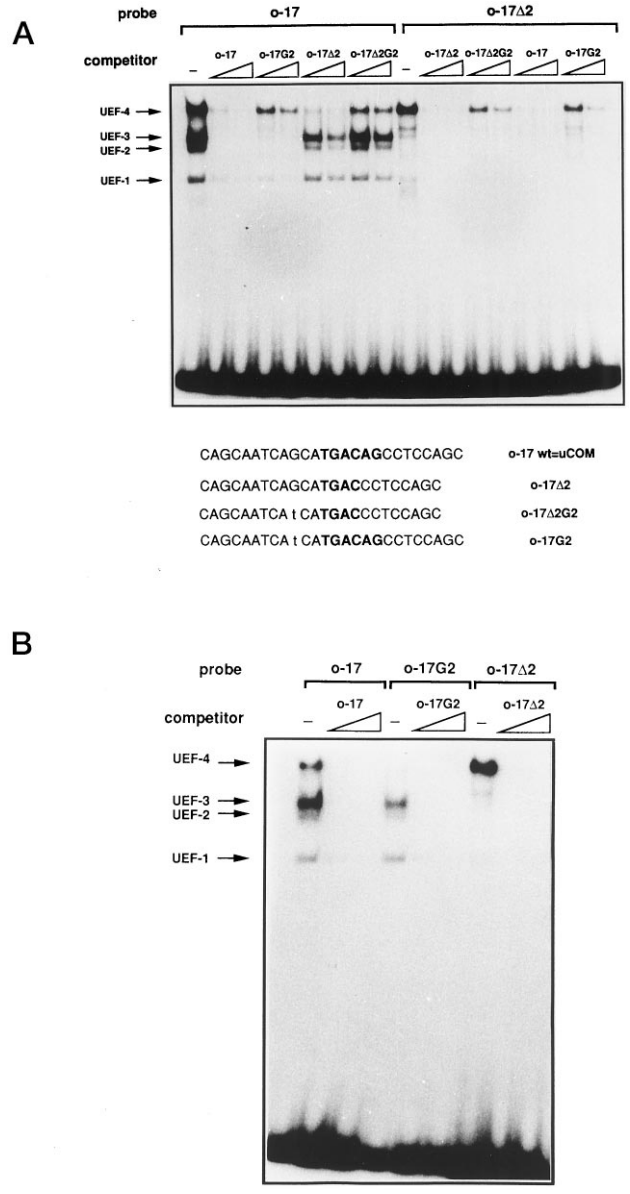

FIG. 6. **The G residues involved in UEF-4 contacts are important for binding also in the context of the wild type uCOM sequence.** EMSA was carried out with 5 μg of HeLa nuclear extract and 2.5 fmol of probe. *A*, ability of different mutated oligonucleotides to compete for binding. The DNA sequence of the employed oligonucleotides is shown at the *bottom*. *B*, direct binding of the o-17G2 mutant oligonucleotide. Competitors were used at 200- and 800-fold excess. Probes o-17 and o-17Δ2 are used as controls.

sequence of o-17 (oligonucleotide o-17D) also destroys UEF-4 binding (see above). To obtain a better refinement of the 3′-border, we individually mutated the TGA nucleotides in the o-17 sequence, and tested their binding ability. As shown in Fig. 7*C*, mutations D1 (A to T) and K (T to G) destroyed UEF-4 binding activity, while mutation D2 (G to C) had a weak effect only. On the other hand, all three mutations interfered with the binding of UEF-1, and mutation D2 also interfered with the binding of UEF2/3. These data extend the 3′-border of the UEF-4 site to at least the T of the TGACAG sequence and show that the binding sites for all factors indeed overlap within the ATGA sequence. Thus, the sequence 5′-AATCAGCAT(G)-3′ is the binding site for UEF-4.

In Vivo Analysis of the Function of the UEF-4 Site—Having identified mutations that in the context of the otherwise wild type uCOM uniquely prevent formation of UEF-4, we have

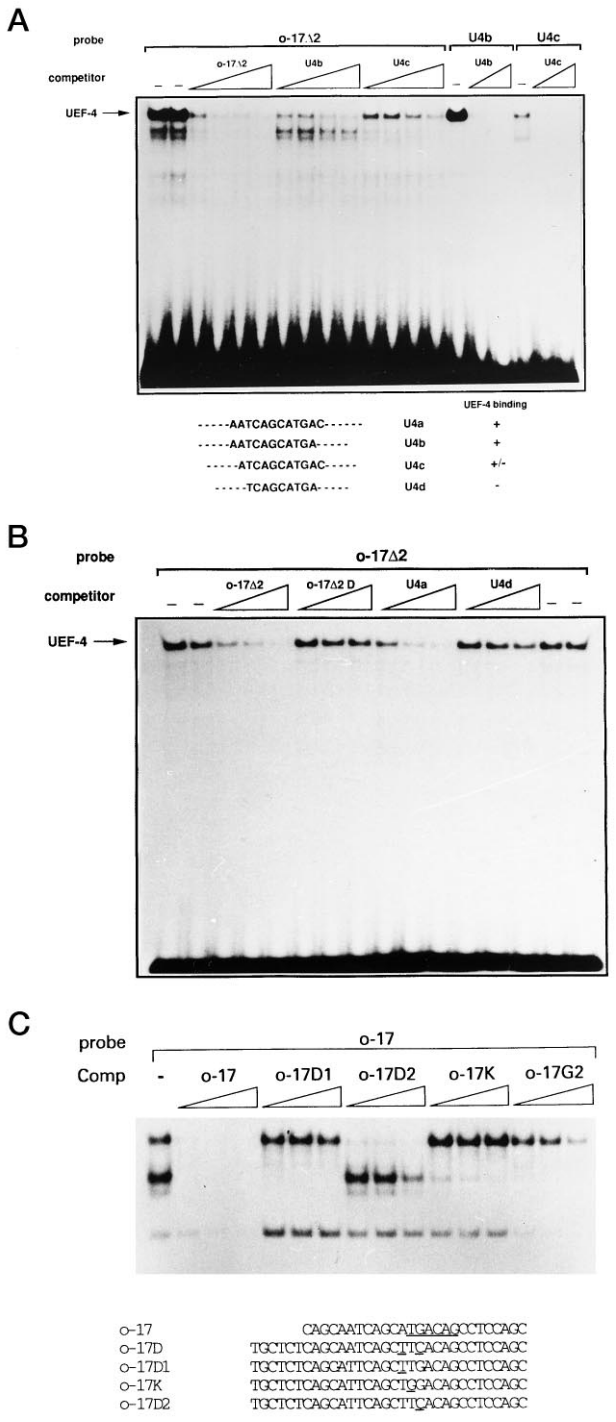

FIG. 7. **Refining the 5′- and 3′-ends of the UEF-4 recognition sequence.** Conditions of EMSA as in Fig. 5. *A*, competitors employed at 100-, 200-, 400-, and 800-fold excess. The sequence of the oligonucleotides employed is shown at the *bottom*. *B*, competitors employed in 200-, 400-, and 600-fold excess. *C*, competitors (*Comp*) used at 200-, 400-, and 600-fold excess. The sequences of oligonucleotides used to define the 3′-end of the UEF-4 site are shown *below*.

employed them to investigate its specific function in the uPA enhancer. The mutations were introduced in the context of the otherwise wild type uPA enhancer, which was cloned in the pBLCAT2 vector upstream of a thymidine kinase promoter, and assayed by transient transfection analysis in HepG2 cells. All constructs had intact PEA3/AP1A and AP1B sites. CAT activity of HepG2 cells was measured under both basal and PMA-induced conditions, since the enhancer contributes to transcription under both conditions (1). Transfection efficiency, evaluated by co-transfection with a *lacZ* expression vector, allowed standardization of CAT activity values (see "Materials and Methods"). The results are summarized in Fig. 8. In the case of the intact uPA enhancer (*WTCAT*), the basal CAT activity was given a value of 1.0 (see the actual CAT activity in the legend). In this case, PMA elicited a 10-fold increase of transcription. Induction was abolished in the presence of mutations in the PEA3, AP1A, or AP1B site (not shown, but see Refs. 1 and 13). Other controls included two different mutations that destroy all COM-binding sites (FIECAT and DECAT) and had a drastic effect on both basal activity, which was reduced to 0.4 (the level of the enhancerless pBLCAT2), and on PMA induction, which was essentially blocked, in agreement with previous data (1, 13). This confirms that COM integrity is essential for enhancer activity. Moreover, we also confirmed that dCOM and uCOM each contributed by about 50% to the PMA-induced enhancer activity, as shown by the two mutants E and D (constructs DCAT and ECAT), in which the protein binding activity of dCOM and uCOM, respectively, is destroyed. The role of UEF-4 in uCOM activity was therefore tested with construct G2CAT, in which only the binding site for UEF-4 was mutated. No effect on basal activity and a minor reduction of PMA inducibility were noticed. Thus, the presence of UEF-1 and UEF-2/3 sites was sufficient to ensure almost full uCOM activity. When we analyzed the activity of a construct in which the binding of UEF-1 and UEF-2/3, but not of UEF-4, had been destroyed (construct Δ2CAT), we found a slightly higher basal level and full PMA inducibility (about 10-fold). Thus, UEF-4 could substitute for UEF-1 and UEF-2/3. However, when also dCOM was mutated (construct Δ2ECAT), basal transcription was reduced to the level of the enhancerless pBLCAT2 construct (to 0.4), but it was still partially inducible by PMA with an induction ratio of about 3, comparable with that observed for construct ECAT, in which the enhancer is also missing dCOM activity. These results indicate that UEF-4 can fully substitute for UEF-1 and UEF-2/3, in both basal and PMA-stimulated transcription, but only in the presence of an intact dCOM. Apparently, UEF proteins bound to uCOM must cooperate with proteins bound to dCOM to ensure full enhancer activity, and thus COM function may be the result of interactions between the two subregions.

Characterization of UEF-4—We purified the UEF-4 DNA binding activity through a variety of techniques ending with DNA affinity chromatography on the o-17Δ2 oligonucleotide (see "Materials and Methods"). The purified factor, analyzed on a 10–15% gradient SDS-polyacrylamide gel electrophoresis, showed two main bands of about 105 and 65 kDa, respectively (Fig. 9*A*, *lane labeled 18*). The two bands represent at least 90% of the total proteins as showed by further diluting fraction 18 (data not shown). The DNA binding specificity of purified UEF-4 was tested by EMSA; the complex formed by labeled o-17Δ2 oligonucleotide with an aliquot of the purified fraction had a mobility identical to the complex formed with the crude nuclear extract and was competed for by unlabeled o-17Δ2, but not by o-17Δ2D (Fig. 9*B*), as expected for UEF-4. In addition, UEF-4 also bound an oligonucleotide reproducing the ICK-1 element of the LD78/MIP-1α promoter (see Fig. 11 for sequence) (data not shown).

To test whether both polypeptides were involved in DNA binding, we performed UV cross-linking analysis with the UEF-4-binding site, using both purified UEF-4 and crude HeLa nuclear extract and a bromine-substituted o-17Δ2 probe. The complex shifted by the labeled bromine-substituted o-17Δ2 probe on native polyacrylamide gel was UV-irradiated *in situ*, excised, and subjected to SDS-PAGE analysis under reducing conditions (see "Materials and Methods"). The results (Fig. 10)

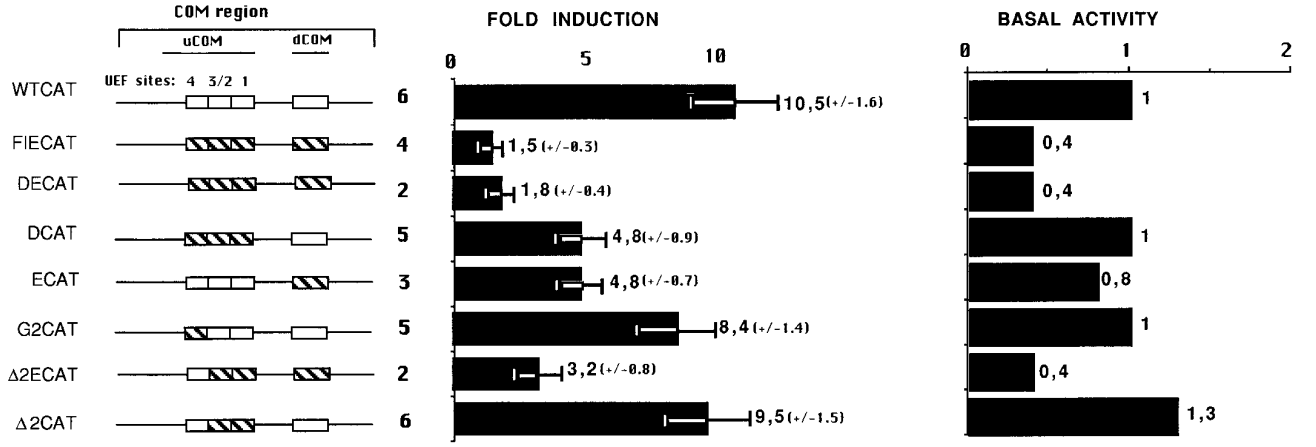

FIG. 8. **Effect of COM mutations on uPA enhancer activity in HepG2 cells.** HepG2 cells were transfected with pBLCAT2 constructs modified as described under "Materials and Methods," and CAT activity was measured from both untreated and PMA-treated cells. CAT values are normalized for transfection efficiency (see "Materials and Methods"). The value of 1 is assigned to the activity measured in untreated HepG2 cells with the wild type enhancer construct (*WTCAT*). This corresponds to 0.5% conversion of [^{3}H]acetyl-CoA. *Fold induction* represents the ratio between CAT activity of PMA-treated *versus* untreated cells. The data represent the average of different experiments (± S.D.). It should be noted that FIECAT construct has the same induction ratio and basal activity as the enhancerless pBLCAT2 vector (not shown). In the *left part* of the *figure*, the *stripes* identify mutationally inactivated sites.

showed in both cases a unique 120–130-kDa cross-linked adduct. The migration of the adduct was in good agreement with that of the 105-kDa species identified by SDS-PAGE analysis of the purified fraction, considering the 15–18-kDa contribution of the oligonucleotide. The data therefore indicate that the 105-kDa polypeptide is involved in direct DNA binding. Whether or not the 65-kDa polypeptide is also involved in DNA binding remains to be established.

DISCUSSION

The −2.3 kilobase enhancer mediates transcriptional activation of the uPA gene by PMA, epidermal growth factor, okadaic acid, granulocyte-macrophage colony-stimulating factor, and cytoskeletal rearrangements (12, 13, 15, 28). It is formed by a complex and novel arrangement of protein-binding sites (Fig. 1). The combined PEA3/AP-1A and AP-1B sites, bordering the 5′- and 3′-ends of the enhancer, are all essential for both transcriptional induction and basal activity of the uPA gene. AP-1A is an octameric sequence that in PMA-induced HepG2 cells binds the c-Jun/ATF-2 dimer; the AP-1B is a heptameric motif, similar to the collagenase PMA-responsive element, which is recognized by c-Jun/c-Jun and c-Jun/c-Fos dimers (27). In granulocyte-macrophage colony-stimulating factor-stimulated macrophages, the PEA-3 site appears to bind the Ets-2 protein (28). In several cells, COM-mediated synergic action of these sites is required to activate transcription (13, 14). Remarkably, PEA3 and AP-1 sites, that in several enhancers are sufficient to activate transcription (29–31), do not do so in the uPA gene if the interposed 74-base pair COM region is mutated (13). Thus, the activity of the uPA enhancer requires cooperation between transactivating sites. COM-mediated transactivation does not appear to be restricted only to the PEA3/AP-1A and AP-1B sites. Indeed, COM allows cooperation between the PEA3/AP-1A site and a glucocorticoid-responsive element (13). Thus, COM-mediating synergic transcriptional activation may be at least in part independent of the nature of the transcription factors involved.

In agreement with the view that COM activity is mediated by its binding proteins, disruption of all individual protein-binding sites of COM abolished the enhancer activity (Fig. 8). On the other hand, although essential, COM has no direct transactivation capacity nor facilitates binding of transactivator proteins to their specific sites (1). The bipartite COM structure is reproduced at the functional level. Indeed, while uCOM and dCOM can substitute for each other in basal transcription, each contributes by about 50% to the COM-dependent PMA induction (Ref. 1; see also Fig. 8).

In this paper we have characterized the UEF-4-binding site in uCOM. First, the proteins binding to the UEF-4 and UEF-1 sites were shown to be distinct from UEF-2/3, since their binding activity was not influenced by the depletion of UEF-2/3 from HeLa nuclear extracts (Fig. 2). This was further supported by the isolation of distinct DNA sequences binding only UEF-4 (o-17Δ2) or uniquely affecting UEF-4 binding (o-17Δ2). The independent binding of the various factors was also confirmed by binding studies performed with limiting amounts of uCOM that failed to show supershifted complexes (not shown). Thus, simultaneous binding of UEF proteins is unlikely to occur in the context of uCOM. Since a two-nucleotide substitution in o-17 separated by a single base pair, mutation D, impaired the binding of all UEF proteins, the UEF-binding sites probably overlap. The mutagenic analysis of uCOM identified the sequence AATCAGCATG as the UEF-4-binding site. This sequence overlaps by two nucleotides the UEF-2/3-binding site, TGACAG (Fig. 11*B*), and mutation of the overlapping sequence destroys binding of all UEF factors (D mutation).

The functional role of UEF-4 was analyzed *in vivo* by transfecting into HepG2 cells (under basal or PMA-induced conditions) a thymidine kinase promoter-driven CAT reporter plasmid carrying a wild type or mutated uPA enhancer. The data show that the UEF-4 and UEF-1 plus UEF-2/3 sites of uCOM could functionally replace each other. Indeed, a mutation of the UEF-4 site (mutation o-17Δ2) had no effect on basal activity and caused only a minor reduction of PMA-induced transcription (Fig. 8). Likewise, a uCOM mutated for UEF-1 and UEF-2/3 binding, but still carrying an intact UEF-4-binding site (o-17Δ2 mutation), displayed full basal and PMA-induced activities (Fig. 8). The requirement for at least one of these protein-binding sites was shown by the effect of mutation D, which destroys all three binding sites (UEF-1, UEF-2/3, and UEF-4) and impaired the activity of the uPA enhancer behaving like a uCOM deletion. The results were different, however, when dCOM was functionally absent. The double mutation o-17Δ2E, lacking UEF-1/2/3-binding sites in uCOM and all of the dCOM-binding sites, displayed a very low basal level activity comparable with that of an enhancerless constructs but was still 4-fold induced by PMA (Fig. 8). Therefore, UEF-4, as well

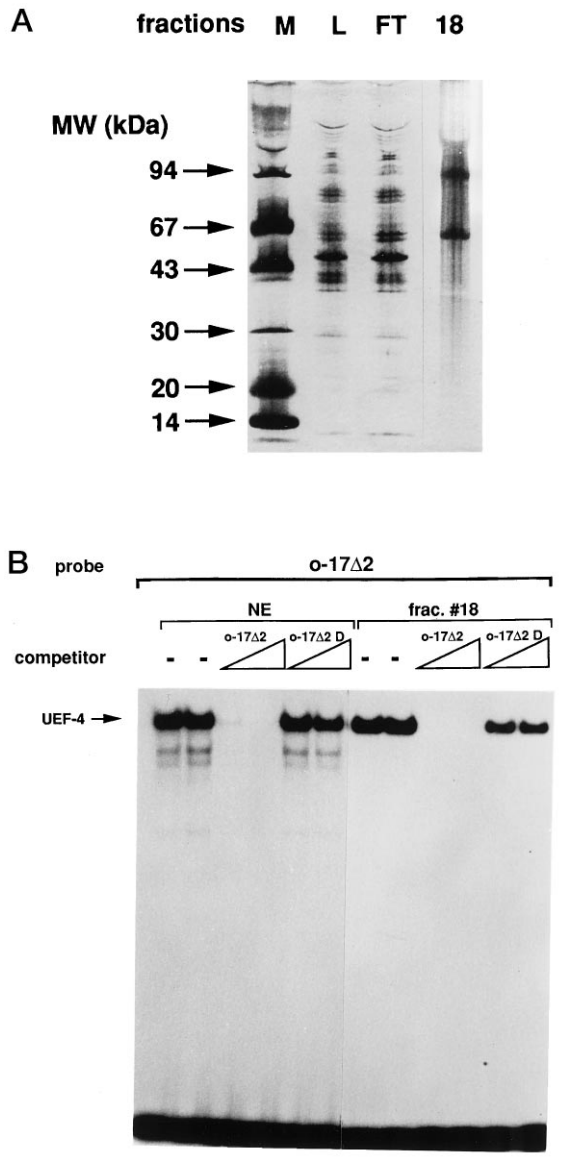

FIG. 9. **Electrophoretic and DNA binding properties of purified UEF-4.** *A*, silver staining of the SDS-PAGE analysis of purified UEF-4 (fraction 18 from the DNA affinity column; see "Materials and Methods"). *M*, molecular weight markers (size indicated on the *left*). *L*, an aliquot of the proteins loaded onto the DNA-affinity column; *FT*, flow-through; *18*, the analysis of 1 μl of fraction 18 of this column. *B*, EMSA of 5 μg of whole nuclear extract (*NE*) of HeLa cells or of 1 μl of fraction 18 eluted from the DNA affinity column. In both cases, 2.5 fmol of probe and 200- and 600-fold excess unlabeled competitor were employed.

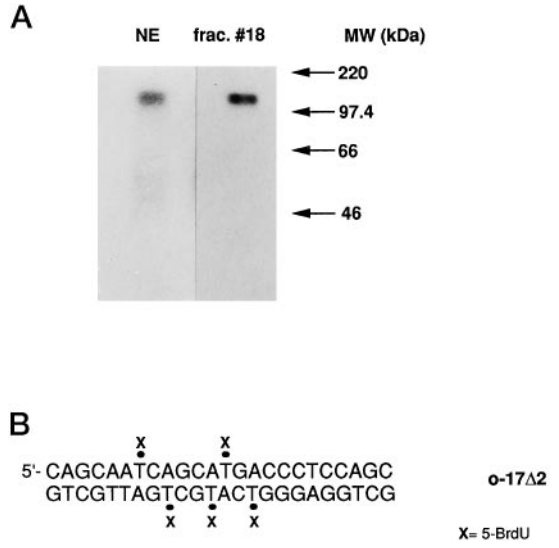

FIG. 10. **The 110-kDa polypeptide has DNA binding activity.** *A*, SDS-PAGE (10% acrylamide) analysis of the products of the cross-linking reaction between bromine-substituted, ^{32}P-labeled o-17Δ2 oligonucleotide and 15 μg of nuclear extract (*NE*) or 6 μl of purified UEF-4 (fraction 18; *frac. #18*). The UEF-4 complex was separated by EMSA, UV-cross-linked *in situ*, and analyzed by SDS-PAGE (see "Materials and Methods"). The position of the molecular weight markers is shown on the *right*. *B*, sequence of the o-17Δ2 oligonucleotide and position of the 5-bromodeoxyuridine substitutions.

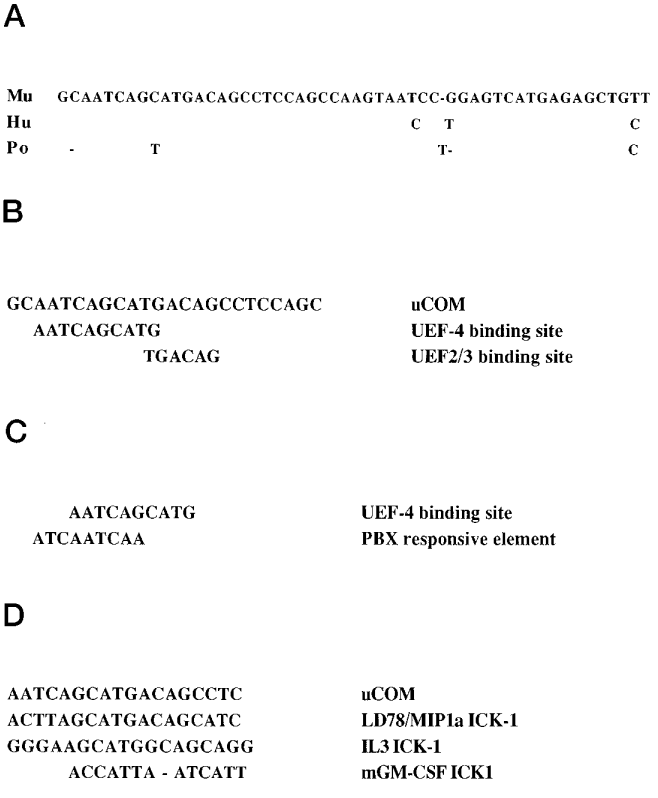

FIG. 11. **Evolutionary conservation and sequence homologies throughout the COM region.** *Panel A* shows the DNA sequence of the mouse (*Mu*) uPA-COM region. The sequences of the human (*Hu*) and porcine (*Po*) uPA enhancers are indicated *below* (only the differences are shown). *Dashes* indicate a gap. *Panel B* shows the extent of overlap between UEF-2/3- and UEF-4-binding sites in uCOM. *Panel C* highlights the sequence homology between UEF-4 and the PBX recognition element (15). *Panel D* shows the homologies between the uCOM and the ICK-1 sequence of several inducible cytokine and protease genes (taken from the following sources: LD78/MIP-1α ICK-1 (22), interleukin-3 (27), and granulocyte-macrophage colony-stimulating factor (18).

as the UEF-1 plus UEF-2/3-binding site, is an important requirement in the function of COM in mediating PMA induction of transcription.

Unraveling the molecular basis of COM function requires better understanding of the proteins involved and their structural and functional characterization. *In vitro* binding studies showed that uCOM can bind three distinct UEF factors differing in electrophoretic mobility and binding specificity. In fact a DNA affinity resin, containing the TGACAG exanucleotide flanked by unrelated sequences, was able to deprive HeLa nuclear extracts of UEF-2/3 binding activities. We have now isolated to near homogeneity the UEF-4 factor exploiting the information obtained on its DNA sequence specificity. The most purified UEF-4 preparation (Fig. 9) contains two protein bands of 105 and 65 kDa and binds DNA with the expected DNA sequence specificity. UV cross-linking analysis (Fig. 10) shows

that at least the 105-kDa species is endowed with DNA binding activity. However, we cannot yet exclude the possibility that the 65-kDa protein also participates in the binding.

Interestingly, the 5′-half of the UEF-4-binding site is identical to the 5′-half of the site recognized by PBX (see Fig. 11*C*), a 40–55-kDa homeodomain protein regulating the DNA binding activity of Hox (32). It may be possible, therefore, that the 5′-half-site of the UEF-4 recognition sequence binds a PBX family member. However, neither the 65- nor the 105-kDa polypeptide of purified UEF-4 cross-reacts with specific anti-PBX antibodies (data not shown).

The DNA sequence of COM is highly conserved among human, porcine, and murine species (Fig. 11*A*). In addition, COM-homologous sequences are present in the ICK-1 elements of proteases like stromelysin and of chemokine and cytokine genes, like LD78 (the murine homolog of human MIP-1α chemokine), interleukin-3, and granulocyte-macrophage colony-stimulating factor, all PMA-inducible genes (16–20, 33) (Fig. 11*D*). Taken together, these considerations suggest that COM-binding proteins play a general role also in transcription units other than uPA. Of particular interest is the homology between uCOM and the ICK-1 sequence of the *LD-78/MIP1*α gene (18), which extends over 15 of the 18 bases, including the UEF-2/3 and part of the UEF-4-binding sites. We have now shown that this sequence indeed binds the purified UEF-4 factor. We previously showed that UEF-2/3 also bound the negatively regulating ICK-1 sequence of the interleukin-3 gene (16). COM, like the ICK-1 elements, represents a transcription-inhibitory sequence that, at least in the uPA enhancer, prevents AP-1 cooperation. In fact, the COM requirement can be relieved by its deletion, which allows the two AP-1 sites to cooperate (1). The characterization of the proteins binding to COM will therefore lead to a better understanding of the transcription regulation of a set of proteases and cytokine genes.

REFERENCES

1. De Cesare, D., Palazzolo, M., and Blasi, F. (1996) *Oncogene* **13,** 2551–2562
2. Blasi, F., Conese, M., Møller, L. B., Pedersen, N., Cavallaro, U., Cubellis, M. V., Fazioli, F., Hernandez Marrero, L., Limongi, P., Muñoz-Canoves, P., Resnati, M., Riittinen, L., Sidenius, N., Soravia, E., Soria, M. R., Stoppelli, M. P., Talarico, D., Teesalu, T., and Valcamonica, S. (1994) *Fibrinolysis* **8,** Suppl. 1, 182–188
3. Carmeliet, P., Schoonjans, L., Kieckens, L., Ream, B., Degen, J., Bronson, R., De Vos, R., van den Oord, J. J., Collen, D., and Mulligan, R. C. (1994) *Nature* **368,** 419–424
4. Danø, K., Behrendt, N., Brünner, N., Ellis, V., Ploug, M., and Pyke, C. (1994) *Fibrinolysis* **8,** Suppl. 1, 189–203
5. Gyetko, M. R., Chen, G.-H., McDonald, R. A., Goodman, R., Huffnagle, G. B., Wilkinson, C. C., Fuller, J. A., and Toews, G. B. (1996) *J. Clin. Invest.* **97,** 1818–1826
6. Wigler, M., and Weinstein, I. B. (1976) *Nature* **259,** 232–233
7. Besser, D., Verde, P., Nagamine, Y., and Blasi, F. (1996) *Fibrinolysis* **10,** 215–237
8. Sappino, A.-P., Huarte, J., Belin, D., and Vassalli, J.-D. (1989) *J. Cell Biol.* **109,** 2471–2479
9. Teesalu, T., Blasi, F., and Talarico, D. (1996) *Mech. Dev.* **56,** 103–116
10. Larsson, L. I., Skriver, L., Nielsen, L. S., Grøndahl-Hansen, J., Kristensen, P., and Danø, K. (1984) *J. Cell Biol.* **98,** 894–903
11. Verde, P., Boast, S., Franze', A. M., Robbiati, F., and Blasi, F. (1988) *Nucleic Acids Res.* **16,** 10699–10716
12. Lee, J. S., von der Ahe, D., Kiefer, B., and Nagamine, Y. (1993) *Nucleic Acids Res.* **21,** 3365–3372
13. Nerlov, C., De Cesare, D., Pergola, F., Caracciolo, A., Blasi, F., Johnsen, M., and Verde, P. (1992) *EMBO J.* **11,** 4573–4582
14. Nerlov, C., Rørth, P., Blasi, F., and Johnsen, M. (1991) *Oncogene* **6,** 1583–1592
15. Rørth, P., Nerlov, C., Blasi, F., and Johnsen, M. (1990) *Nucleic Acids Res.* **18,** 5009–5017
16. Berthelsen, J., Vandekerkhove, J., and Blasi, F. (1996) *J. Biol. Chem.* **271,** 3822–3830
17. Mathey-Prevot, B., Andrews, N. C., Murphy, H. S., Kreissman, S. G., and Nathan, D. G. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 5046–5050
18. Nomiyama, H., Hieshima, K., Hirokawa, K., Hattori, T., Takatsuki, K., and Miura, R. (1993) *Mol. Cell. Biol.* **13,** 2787–2801
19. Shoemaker, S. G., Hromas, R., and Kaushansky, K. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9650–9654
20. Sirum-Connolly, K., and Brinckerhoff, C. E. (1991) *Nucleic Acids Res.* **19,** 335–341
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Neumann, J. R., Morency, C. A., and Russian, K. O. (1987) *BioTechniques* **5,** 444–446
23. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489
24. Hansen, S. K., Nerlov, C., Zabel, U., Verde, P., Johnsen, M., Baeuerle, P. A., and Blasi, F. (1992) *EMBO J.* **11,** 205–213
25. Siebenlist, U., and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 122–126
26. Luckow, B., and Schutz, G. (1987) *Nucleic Acids Res.* **15,** 5490
27. De Cesare, D., Vallone, D., Caracciolo, A., Sassone-Corsi, P., Nerlov, C., and Verde, P. (1995) *Oncogene* **11,** 365–376
28. Stacey, K. J., Fowles, L. S., Colman, M. S., Ostrowski, M. C., and Hume, D. A. (1995) *Mol. Cell. Biol.* **15,** 3430–3441
29. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P., and Karin, M. (1987) *Cell* **49,** 729–739
30. Lee, W., Mitchell, P., and Tjian, R. (1987) *Cell* **49,** 741–752
31. Wasylyk, B., Wasylyk, C., Flores, P., Begue, A., Leprince, D., and Stehelin, D. (1990) *Nature* **346,** 191–193
32. Lu, Q., Knoepfler, P. S., Scheele, J., Wright, D. D., and Kamps, M. P. (1995) *Mol. Cell. Biol.* **15,** 3786–3795
33. Miyatake, S., Otsuka, T., Yokota, T., Lee, F., and Arai, K. (1985) *EMBO J.* **4,** 2561–2568

Characterization of UEF-4, a DNA-binding Protein Required for Transcriptional Synergism between Two AP-1 Sites in the Human Urokinase Enhancer

Dario De Cesare, Martina Palazzolo, Jens Berthelsen and Francesco Blasi

J. Biol. Chem. 1997, 272:23921-23929.
doi: 10.1074/jbc.272.38.23921