

Purification and Characterization of UEF3, A Novel Factor Involved in the Regulation of the Urokinase and Other AP-1 Controlled Promoters*

(Received for publication, August 18, 1995, and in revised form, November 28, 1995)

Jens Berthelsen^{‡§}, Joel Vandekerckhove[¶], and Francesco Blasi^{‡§||}

From the [‡]Department of Genetics and Biology of Microorganisms, University of Milano, Via Celoria 26, 20135 Milano, Italy, the [¶]DIBIT (Department of Biological and Technological Research), S. Raffaele Scientific Institute, the [§]Department of Molecular Cell Biology, University of Copenhagen, Øster Farimagsgade 2A, 1353 Copenhagen, Denmark, and the ^{||}Laboratorium voor Genetica, Rijksuniversiteit Gent, B-9000 Gent, Belgium

Basal as well as induced transcription from the human urokinase-type plasminogen activator gene requires an enhancer containing two elements, a combined PEA3/AP-1 and a consensus AP-1 site. The integrity of each of these binding sites as well as their cooperation is required for activating transcription. The two elements are separated by a 74-base pair cooperation mediating (COM) region required for the cooperation between the transactivating sites. The COM region contains binding sites for four different unidentified urokinase-type plasminogen activator enhancer factors (UEF 1 to 4), all four required for correct COM activity. We have purified UEF3 from HeLa nuclear extracts by several chromatographic steps including DNA affinity purification. Purification and UV cross-linking data showed that UEF3 is a complex of three polypeptides (p40, p50, and p64). Amino acid sequence from one peptide of p64 was obtained, which showed no homology to other known proteins. Both crude and purified UEF3 specifically bound to the sequence TGACAG as shown by electrophoretic mobility shifts and methylation interference studies. DNA-binding specificity of purified UEF3 was identical to that of NIP, a non-characterized factor binding and regulating multiple AP-1-regulated promoters like stromelysin and interleukin-3. Thus UEF3 appears to be a general DNA-binding factor involved in modulating the transcriptional response of AP-1 containing promoters.

Urokinase plasminogen activator (uPA)¹ is a serine protease and a key enzyme in the proteolytic cascade activating plasminogen and thereby leading to degradation of the extracellular matrix (Blasi and Verde, 1990). The uPA gene is expressed constitutively in the ectoplacental cone during embryo implantation, in kidney and lung in the adult (Larsson *et al.*, 1984;

Sappino *et al.*, 1989). Expression of the uPA gene is induced in inflammatory and tumoral pathological conditions, and appears to be regulated by a very large number of growth factors and cytokines (Blasi and Verde, 1990). Transcriptional regulation of the human uPA gene relies upon an enhancer element positioned 2 kilobase pairs upstream of the transcriptional start site (Cassady *et al.*, 1991; Nerlov *et al.*, 1991; Rørth *et al.*, 1990; Verde *et al.*, 1988). Two elements within the enhancer have been shown to be essential for its activity: a combined PEA3/AP-1 site similar to that of the collagenase gene, and a 74-base pair downstream consensus AP-1 site (Gutman *et al.*, 1990; Nerlov *et al.*, 1991; Rørth *et al.*, 1990) (see Fig. 1A). These sites are important for both basal level and induced enhancer activity (Cassady *et al.*, 1991; Nerlov *et al.*, 1991; Rørth *et al.*, 1990) and appear to cooperate to activate transcription.

A cooperation mediating (COM) element is positioned between the combined PEA3/AP-1 and the AP-1 site (Nerlov *et al.*, 1992). The COM element does not appear to have transactivation-mediating activity by itself, but is important for synergistic activation of the PEA3/AP-1 and AP-1 sites. DNase footprinting and site-directed mutagenesis have identified two areas of the COM region (u-COM and d-COM) important for the function of the enhancer. Four different so far unidentified nuclear factors, uPA enhancer factors (UEF) 1 to 4, appear to bind to these regions. Mutations within the COM element, affecting the binding of the different UEFs impair both inducibility (Nerlov *et al.*, 1992) as well as basal level activity² of the enhancer.

Promoters of the interleukin-3, LD78, and stromelysin genes (Mathey-Prevot *et al.*, 1990; Sirum-Connolly and Brinckerhoff, 1991; Nomiya *et al.*, 1993) are also regulated via AP-1 sites, and appear to have sequence homologies to the uPA enhancer not only in the transactivating elements, but also in the COM region (Fig. 1B). In particular, a NIP site is present, close to both PEA3 and AP-1 sites in the stromelysin promoter (Sirum-Connolly and Brinckerhoff, 1991) and to AP-1 and Elf-1 sites in the interleukin-3 promoter (Gottschalk *et al.*, 1993; Mathey-Prevot *et al.*, 1990). The activity of the NIP sites in these promoters appears to be that of modifying the efficiency of elements responding to the transcription-inducing signals. In this view, and because of the sequence homologies, one or more UEF factor of the uPA enhancer might correspond to the NIP-binding proteins. In this paper we have purified and preliminarily characterized UEF3 and shown that it recognizes a TG(A/G)CAG sequence, common to both the uPA COM region and the NIP element.

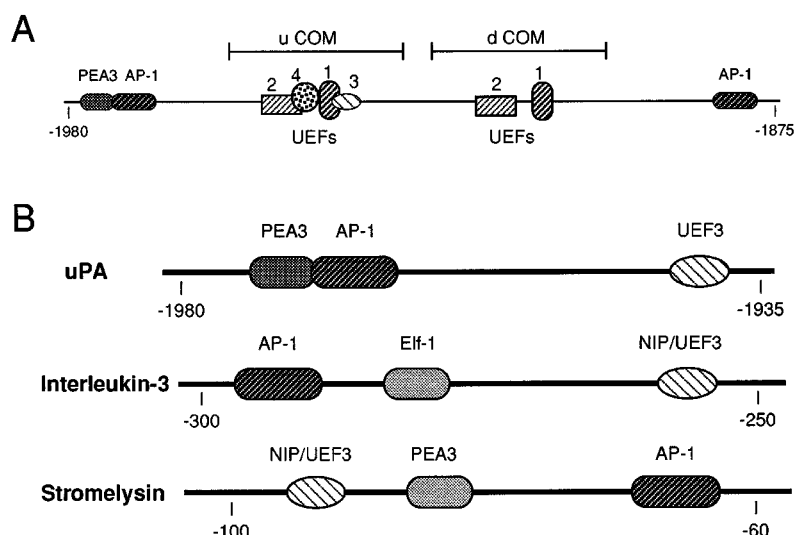
* This work was supported by grants of the Italian Association for Cancer Research (AIRC), Consiglio Nazionale delle Ricerche (PF ACRO), the AIDS Grant of Italian Ministry of Health, the Danish Cancer Society, and the Human Capital and Mobility Program of the European Union. The costs of publication 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.

|| To whom correspondence should be addressed: DIBIT H.S. Raffaele, via Olgettina 60, 20132 Milan, Italy. Tel.: 39-2-2643-4832; Fax: 39-2-2643-4844.

¹ The abbreviations used are: uPA, urokinase plasminogen activator; DTT, dithiothreitol; COM, cooperation mediating region; PMSF, phenylmethylsulfonyl fluoride; UEF, uPA enhancer factor(s); EMSA, electrophoretic mobility shift assay; PAGE, polyacrylamide gel electrophoresis; IL-3, interleukin 3; UC, upper complex, LC, lower complex.

² M. Palazzolo and D. De Cesare, personal communication.

FIG. 1. Structure of the uPA, IL-3, and stromelysin promoters. A, organization of the uPA enhancer, showing the upstream combined PEA3-AP-1 site, the downstream AP-1 site, and the interspersing COM element divided into the functional regions, u-COM and d-COM. Binding positions for the different UEFs within the u-COM and d-COM are indicated. B, comparison of overall structure similarity between the uPA enhancer PEA3-AP-1 and UEF3 region with regions from the interleukin-3 and stromelysin promoters containing an inducible AP-1 site.



MATERIALS AND METHODS

Oligonucleotides—Oligonucleotides were synthesized on a Applied Biosystems 381 DNA Synthesizer. Introduction of 5-bromodeoxyuridine (Milligen, Millipore) into the oligonucleotides was done by using the X-port of the synthesizer. The sequence of the oligonucleotides used in this study is shown in Table I.

Preparation of Nuclear Extracts— 10^9 commercially obtained nuclei (4C, Mons, Belgium) were resuspended in 5 ml of buffer C (20 mM HEPES, pH 7.9, 25% glycerol (v/v), 0.42 M NaCl, 1.5 mM $MgCl_2$, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$), and extracted essentially as described (Dignam *et al.*, 1983). Briefly, just prior to homogenization, the concentration of NaCl was adjusted to 0.42 M. The homogenate were rotated for 45 min on ice and centrifuged at 38,000 rpm for 60 min in a Beckman Ti70 rotor at 4 °C. The resulting supernatant was processed in either of two ways: 1) for all binding studies the extract was dialyzed against 50 volumes of buffer H2K150 (25 mM HEPES, pH 7.9, 20% glycerol (v/v), 150 mM KCl, 1 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$) for 8–12 h. The dialysate was centrifuged at 25,000 $\times g$ 20 min in a Sorvall SS34 rotor. 2. When used for purification, the extract was dialyzed twice against 15 volumes of buffer H1K125 (25 mM HEPES, pH 7.2, 20% glycerol (v/v), 125 mM KCl, 1 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$) for a total of 10 h.

Protein Concentration Determination—Determination of proteins contents in extracts and fractions were done by the Bio-Rad Protein Assay (Bio-Rad), also known as the Bradford assay, according to the manufacturer's protocol. As protein concentration standards, appropriate dilutions of γ -globulin (Bio-Rad) were used.

Electrophoretic Mobility Shift Assays (EMSA)—For most binding studies, 6 μ g of HeLa nuclear extract was incubated in 9 μ l of buffer H2K150 and 2 μ g of poly(dI-dC) (Pharmacia) on ice for 10 min. 10^4 cpm labeled probe (~ 1 –2 fmol) and, if necessary, unlabeled competitor in a total of 10 μ l were added to the extract, incubated at room temperature for 20 min, and loaded on a 5% native polyacrylamide gel in 0.5 \times TBE, dried, and exposed to Agfa curix film with intensifying screens at -80 °C for 1–24 h. When EMSA was used as an assay to follow purification, the assay was performed as above with the exception that the amounts of protein and poly(dI-dC) used varied (see figure legends) depending on the purity of the protein extract.

DNA Affinity Columns—Oligonucleotides were coupled to CNBr-Sepharose by a method modified from (Kadonaga and Tjian, 1986). Single stranded oligonucleotides were purified by NAP-5 column (Pharmacia), and 800 μ g of each strand labeled with polynucleotide kinase (BioLabs, 250 units) in 200 μ l of kinase buffer supplemented with 3 mM ATP, 20 μ Ci of [γ - 32 P]ATP 2 h at 37 °C. The complementary strands were mixed and annealed by heating to 90 °C and slowly decreasing the temperature to 37 °C over a period of 5 h using a Omnigene tempcycler (Hybaid). The oligonucleotide was then precipitated twice with ethanol and resuspended in 200 μ l of oligonucleotide-ligation buffer (160 mM Tris, pH 7.6, 10 mM $MgCl_2$, 15 mM DTT, 0.2 mg/ml bovine serum albumin, 1 mM spermidine, 4 mM ATP), added to 2000 units of T4 ligase (New England Biolabs) and ligated at 20 °C for 12 h. The majority of the ligation products were in the size between 75 and 1000 base pairs corresponding to 3–50 concatamers. The ligated products were ex-

tracted once with phenol and once with phenol/chloroform, transferred to a dialysis tube and dialyzed 3 times against 500 ml of TEN100 for a total of 4 h. The dialysate was precipitated, resuspended in 200 μ l of TEN100 and passed through a 8 ml of G-50 Sephadex column to remove single-stranded and unligated double-stranded oligonucleotides from the ligated products. Ligated DNA was ethanol-precipitated and resuspended in 200 μ l of coupling buffer (10 mM potassium phosphate buffer, pH 8.0) and mixed with activated CNBr-Sepharose (Pharmacia), 500 μ g of DNA/g of matrix. The coupling was allowed to proceed for 16 h, the resin washed, blocked with 1 M ethanolamine, pH 8.0, for 5 h, and the coupling efficiency measured. Efficiencies were in the range of 60–100 μ g of DNA/ml of matrix giving a concentration of binding sites of 3.4–5.7 μ M.

Purification of UEF3—Nuclear extracts from 2.5×10^{10} nuclei corresponding to approximately 1300 mg of protein in 140 ml of buffer C, prepared as described above, were loaded on a S-Sepharose XK26 column (Pharmacia) equilibrated in H1K125 (25 mM HEPES, pH 7.2, 20% glycerol (v/v), 1 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$, 125 mM KCl) at a flow rate of 3 ml/min. The flow-through was collected, the column washed with 120 ml of H1K125 and the bound proteins were eluted with a 330 ml of 125–650 mM KCl gradient with a flow rate of 2.5 ml/min and fractions of 10 ml were collected. Fractions were assayed for UEF3 activity by EMSA. Active fractions from two sequential S-Sepharose columns were pooled and dialyzed against 20 volumes of H2K200 (25 mM HEPES, pH 7.9, 20% glycerol, v/v, 1 mM $MgCl_2$, 0.2 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 1 mM $Na_2S_2O_5$ with 200 mM KCl) for 6 h, centrifuged, and loaded onto a 60-ml heparin-Sepharose (Pharmacia) column at 1.5 ml/min. The column was washed with 120 ml of H2K200 and step eluted with H2K450 buffer (same buffer with 450 mM KCl) collecting fractions of 10 ml.

Active fractions from two sequential heparin-Sepharose columns were pooled, dialyzed twice against 2500 ml of H2K150 for a total of 12 h, then cleared by centrifugation. The extract was passed through a 6-ml mutant O-1 Sepharose column at 15 ml/h. Column flow-through with an absorbance higher than 0.025 absorbance units was collected. The column was washed with 18 ml of H2K150 and step eluted with H2K1000 (same buffer but with 1000 mM KCl). Fractions of 1 ml were collected. The flow-through from the mutant column was supplemented with 25 mg of poly(dI-dC) and cleared by centrifugation. The load was applied to a 3-ml O-1 Sepharose column at a flow rate of 15 ml/min, after which the column was washed with 20 ml of H2K150 and 40 ml of H2K200 and eluted with an 18 ml of H2K200 to 1000 gradient, collecting fractions of 1 ml. Fractions were screened both by EMSA and SDS-PAGE. By both criteria, fractions were selected for the final step. Fractions from the O-1 Sepharose chromatography were pooled and dialyzed against 500 ml of H1N100 (H1 buffer with 100 mM NaCl) for 4 h. The dialysate was spun and loaded onto a 1.6/5 Mono-S column (Pharmacia SMART system) at 100 μ l/min. The column was washed with 2.5 ml of H1K100 at 100 μ l/min and eluted with 1 ml of H1K100 to 1000 gradient at a flow rate of 25 μ l/min. Fractions of 50 μ l were collected.

Protein Sequencing—The UEF3 components were separated by SDS-PAGE and electrotransferred onto a polyvinylidene difluoride membrane (Bauw *et al.*, 1989). Proteins were visualized on the blot by Amido

Black staining and on membrane digested with trypsin (Bauw *et al.*, 1989). The peptides released in the digestion buffer were separated by reverse phase high performance liquid chromatography, using a 4.6 × 250-mm C-4 column (Vydac 218TP54). A linear gradient of acetonitrile in 0.1% trifluoroacetic acid, rising from 5 to 70% in 60 min, was used to elute peptides. The column eluate was monitored at 214 nm and peptides collected manually. Some peptide peaks were selected for automated sequence analysis. Peptides were sequenced with a 470A type sequencer (Applied Biosystems Inc.) equipped with a 120A model phenylthiohydantoin-derivative analyzer.

Gel Electrophoresis—SDS-PAGE used for purification fractions was done using the Phast system (Pharmacia). Protein samples were added to a equal volume 2 × SDS sample buffer (20 mM Tris, pH 8.0, 4% SDS, 200 mM DTT), heated at 100 °C for 5 min, added one-sixth volume of 15% iodoacetamide (Heukeshoven and Dernick, 1988), and allowed to react 15 min in the dark. The samples were loaded on a 10–15% gradient polyacrylamide Phast gel (Pharmacia). As protein size markers, high and low molecular weight protein standards (Bio-Rad) were used. Gels were silver stained by a standard procedure for the Phast system (Pharmacia) or by a sensitive procedure (Heukeshoven and Dernick, 1988). This sensitive method is approximately 10 times more sensitive than the standard procedure.

UV Cross-linking—In a modified EMSA reaction, purified UEF3 was added to 0.1 µg of poly(dI-dC), 50 µg of bovine serum albumin, 50,000 cpm of 5-bromodeoxyuridine-substituted O-1, and H2K150 to 20 µl. The gel was run as described above. After separation, the wet gel was exposed to UV irradiation at 306 nm for 30 min and then blotted to DE-81 paper (Whatman) in 0.25 × TBE in a semidry blotter at 1 mA/cm². The blot was exposed to film for 2 h and bands corresponding to UC and LC were cut out and eluted with 200 µl of elution buffer (10 mM Tris, pH 8, 1 mM EDTA, 1.5 mM NaCl, 0.2% SDS, 0.2% Nonidet P-40) supplemented with 30 µg of bovine serum albumin at 37 °C for 4 h. The eluted protein-DNA adducts were recovered by acetone precipitation and the pellet resuspended in 15 µl of SDS-loading buffer (Sambrook *et al.*, 1989) and run by 10% SDS-PAGE together with ¹⁴C-labeled markers (Amersham).

Methylation Interference—Methylation and purification of oligonucleotides

cleotides for methylation interference experiments were performed as described (Siebenlist and Gilbert, 1980). Methylated oligonucleotides were labeled and purified as described above, with the exception that the double-stranded oligonucleotides were only labeled on one strand.

EMSA was performed with purified UEF3 and the partially methylated oligonucleotides as probe. The wet gel was blotted to DE-81 paper as described for UV cross-linking and bands corresponding to UC, LC, and free probe were cut out and eluted with 1.5 M NaCl and recovered by ethanol precipitation. The oligonucleotides were resuspended, extracted twice with phenol/chloroform, precipitated, washed, and dried followed by cleavage with piperidine. The samples were dried and resuspended in formamide loading buffer. 10⁴ cpm of each sample were loaded on a 20% acrylamide gel containing 8 M urea and 0.5 × TBE. The gel was run at 40 W (1 W/cm) for 90 min, dried, and exposed as described previously.

RESULTS

The UEF3 Binding Site—The sequences of the oligonucleotides used in this study are reported in Table I. EMSA showed that the u-COM region (represented by oligonucleotide O-17) formed four retarded bands with a nuclear extract of HepG2 cells, UEF 1 to 4 (Nerlov *et al.*, 1992). By substitution of two bases in the O-17 oligonucleotide, the binding site for UEF3 was suggested to include the TGACAG sequence (Nerlov *et al.*, 1992). This sequence appears elsewhere in the regulatory region of the uPA gene, at position –1482 (Verde *et al.*, 1988). In order to characterize the UEF3 binding sequence, and to search for an oligonucleotide capable of binding only UEF3, we tested whether the oligonucleotide O-1, corresponding to the uPA sequence –1470 to –1500 and thus including the TGACAG at –1482, was capable of binding HeLa nuclear factors. As shown in Fig. 2, using O-17 as probe in EMSA, retarded bands corresponding to the different UEFs were observed, confirming previous data. When O-1 was used as a probe with HeLa nuclear extract, two specific complexes were formed (Fig. 2, *first lane*), designated upper complex (UC) and lower complex (LC). Both complexes were readily competed by 50- and 500-fold excess unlabeled O-1 (*lanes 2 and 3*) as well as O-17 (*lanes 6 and 7*). Two of the four retarded bands observed with the O-17 probe co-migrated with the UC and LC of the O-1 probe (*lane 10*). While unlabeled O-17 was able to compete for binding of all the UEFs, unlabeled O-1 only competed for the two complexes co-migrating with UC and LC (*lanes 11 and 12 versus 15 and*

TABLE I
Sequence of oligonucleotides used in this study

O-17	CAGCAATCAGCATGACAGCCTCCAGC
O-17ml CTCG
O-1	CACCTGAGAGTGACAGAAGGAAGGCAGGGAG
O-1mT.C.....
NIP IL-3	ATGGGAAGCATGGCAGCAGGTGAGAGGATC
NIP Strom	TTTCATCCAAATGGCAGCAGGAC

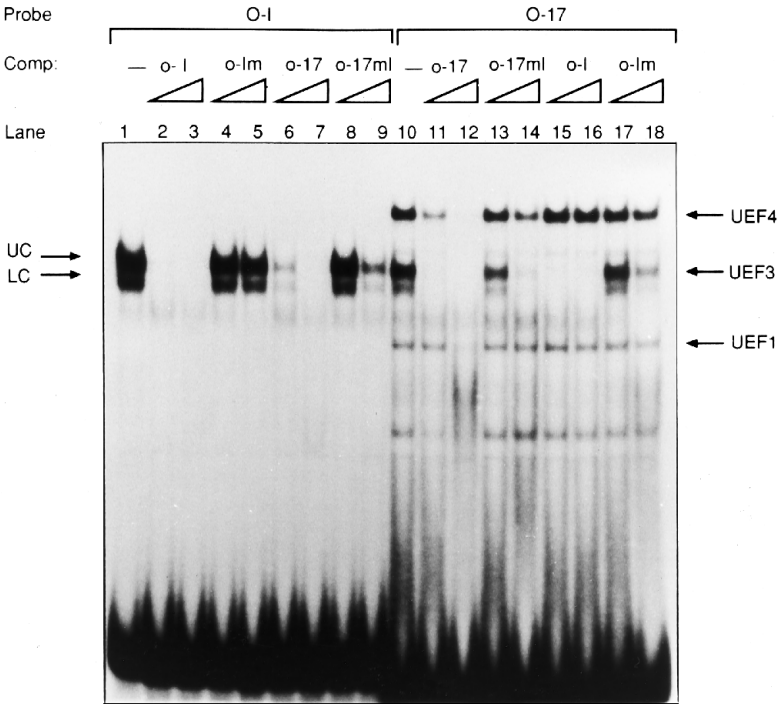


FIG. 2. UEF3 binds to identical sequences contained within the O-1 and O-17 oligonucleotides. Electrophoretic mobility shift assay with HeLa nuclear extracts and radiolabeled O-1 (*lanes 1–9*) or O-17 (*lanes 10–18*) probes. Unlabeled competitor oligonucleotides were used at 50- and 500-fold excess, as indicated. The two bands corresponding to UEF3, UC, and LC, are indicated together with UEF-1 and -4. Binding of the previously identified UEF2 was not observed with the extract used in this study.

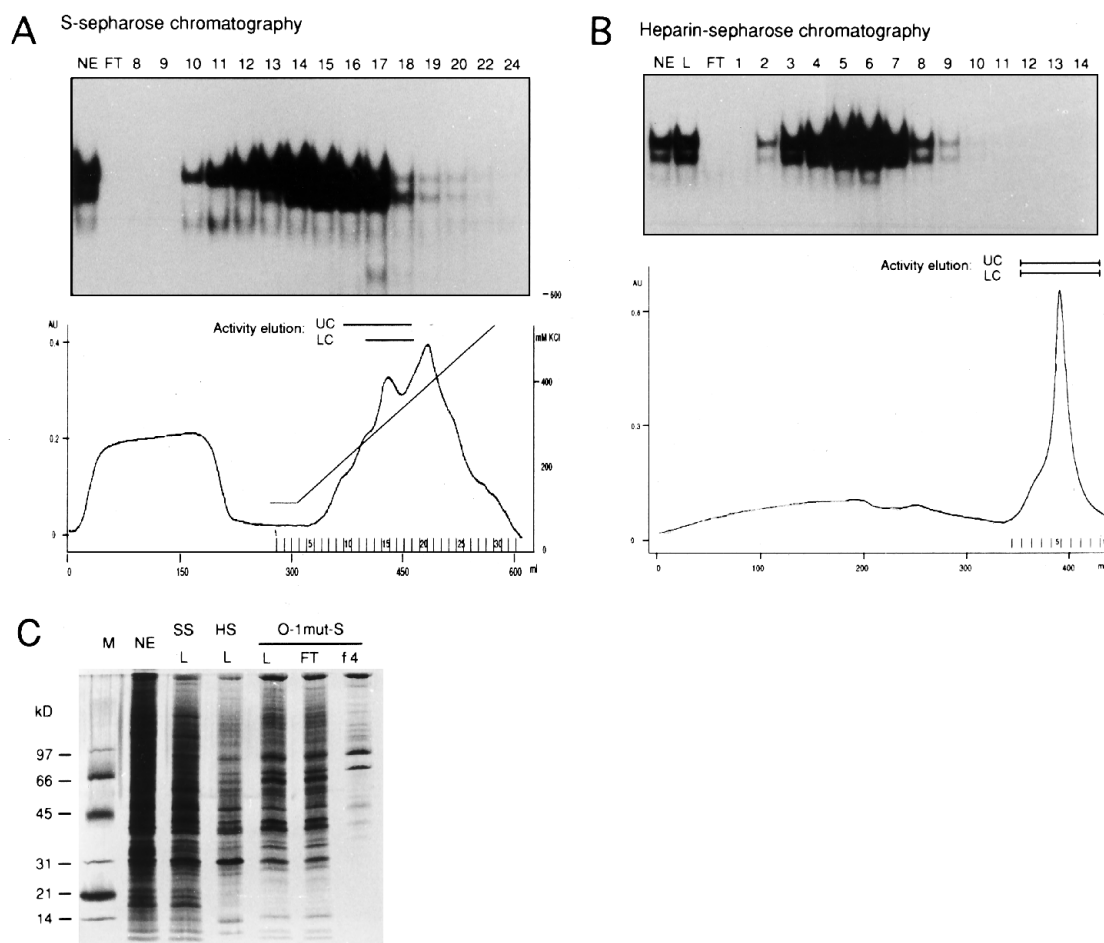


FIG. 3. Initial fractionation of HeLa nuclear extracts. A, S-Sepharose chromatography. Lower panel, optical density profile showing the column flow-through (0–300 ml) and elution (300–600 ml) together with the theoretical salt gradient and collected fractions. AU, absorbance units at 280 nm. Top panel, EMSA assay on 1 μ l of each fraction, in the presence of 2 μ g of poly(dI-dC) and labeled O-1 oligonucleotide. NE, nuclear extract; FT, flow-through; numbers indicate fraction number. B, heparin-Sepharose fractionation profile. Lower panel, absorbance profile of the column flow-through (0–340 ml) and step elution (340–450 ml); AU, absorbance units at 280 nm. Top panel, EMSA analysis, as above. L, load; FT, flow-through; numbers indicate fraction numbers. C, SDS-PAGE analysis of fraction from the initial fractionation steps after mutant DNA Sepharose. 0.5 μ l of each fraction was used. NE, nuclear extracts; SS L, S-Sepharose load; HS L, heparin-Sepharose load; O-1 mut L, load of the O-1mut chromatography; FT, flow-through of the same; f4, fraction 4 of the same. 8 fractions of 1 ml were collected from the elution of the O-1mut Sepharose. Fraction 4 was the peak fraction containing the majority of the proteins. Molecular weight markers are indicated. The gel was silver stained.

16). These data suggest that two of the bands observed with O-17 correspond to those bound by O-1, UC and LC. In order to support this conclusion, two point mutations, partially overlapping the TGACAG sequence, were introduced into O-1 (O-1m) and tested as competitor in EMSA with either O-1 or O-17 probes. As shown in Fig. 2, O-1m was unable to compete for binding of UC and LC with O-1 (lanes 4 and 5) and severely impaired with O-17 (lanes 17 and 18) probe. A mutant O-17 (O-17ml) substituted in the TGACAG sequence was similarly impaired in competing for UC and LC (lanes 8 and 9 and 13 and 14). These data overall show that O-1 and O-17 bind a common factor, UEF3. Thus, these data provided us with an oligonucleotide (O-1) binding only one of a subset of the different UEFs, which could be used for sequence-specific affinity purification and binding characterization experiments.

Purification of UEF3—To investigate the nature and protein composition of UEF3 we set up a protein purification scheme to purify UEF3 from HeLa cell nuclear extracts. The final scheme presented here is a preparative scale purification that eventually allowed us to obtain microsequence information (detailed under “Materials and Methods”). As a first step, cation exchange chromatography on a S-Sepharose HP column was used. The majority of UEF3 activity was found to elute from

230 to 370 mM KCl with LC present only from 270 to 370 mM KCl (Fig. 3A). Enriched fractions from two sequential S-Sepharose HP runs were pooled and loaded onto a heparin-Sepharose column. Approximately 50% of the S-Sepharose eluate bound to the heparin column, and UEF3 eluted over the entire elution peak (Fig. 3B). Enriched fractions from two heparin-Sepharose columns were then passed through a DNA column made with the mutant O-1m sequence to eliminate nonspecific DNA binding activity. EMSA on the eluted fractions showed no UEF3 binding to the mutant columns (not shown). Fig. 3C shows SDS-PAGE analysis of eluates from S-, heparin- and O-1m-Sepharose columns. Several protein bands could be seen eluting from the O-1m column at high salt concentrations (lane f4). In our experience, the O-1m column was an important purification step, since the proteins binding this resin would otherwise interfere strongly with the following DNA affinity chromatography. The removal of interfering factors can be seen from the purification data as the flow-through was enriched in specific binding activity, as compared to the load (Table II).

The flow-through from the mutant O-1m column was supplemented with poly(dI-dC), loaded on the O-1 Sepharose column and eluted with a 200 to 680 mM KCl gradient (Fig. 4). The fractions were screened by EMSA (Fig. 4A) and SDS-PAGE

TABLE II
Summary of the purification

Chromatographic step	Volume ^a	Total protein ^b	Specific activity ^c	Yield	Purification
	ml	mg	units/mg	%	-fold
NE from 10 ¹¹ cells	670	4556	279	100	1.0
S-Sepharose	360	792	1,318	82	4.7
Heparin-Sepharose	135	432	2,250	76	8.1
mut O-1 Sepharose	160	416	2,384	78	8.6
O-1 Sepharose	11	0.11 ^d	680,000	6	2,447
Mono-S	0.25	0.02 ^d	137,375	2.2	4,943

^a Volume is given after active fractions were pooled and dialyzed.

^b Protein concentration was measured by the Bradford assay.

^c Activity was measured in binding units. 1 unit is equivalent to 1 fmol of shifted probe by bandshift. Activity was measured by counting retarded bands in a β -counter.

^d Protein concentration was estimated from silver or Coomassie-stained SDS-PAGE.

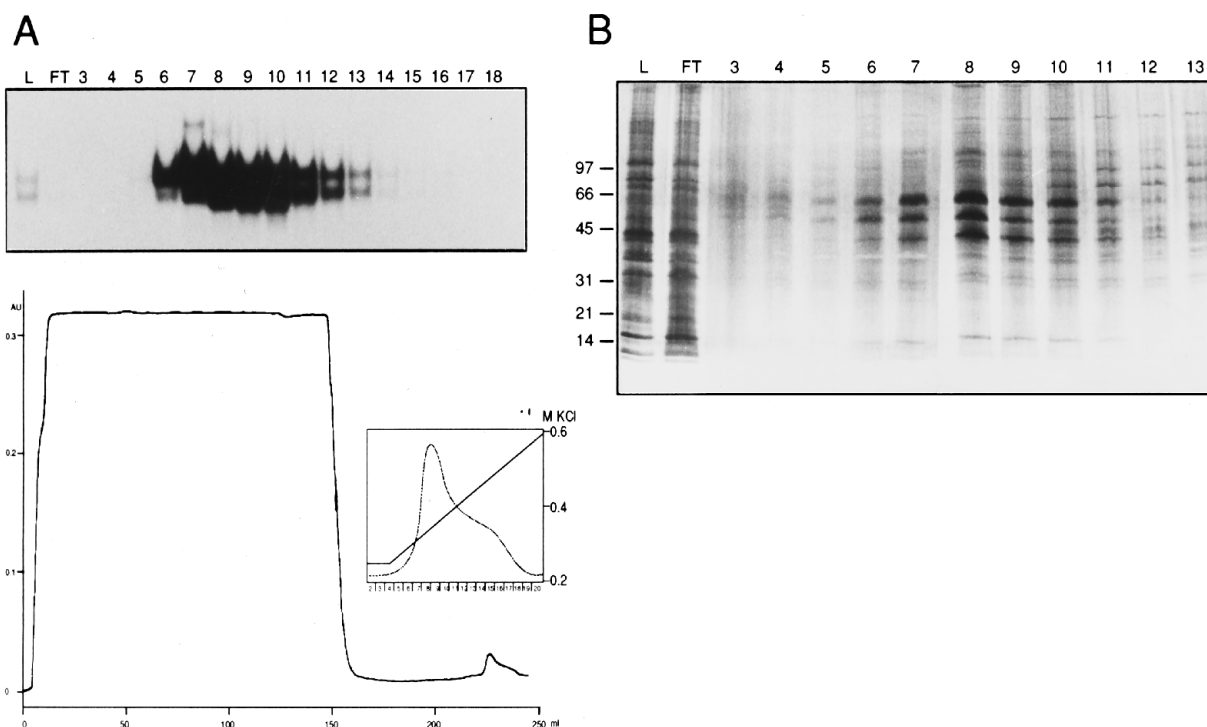


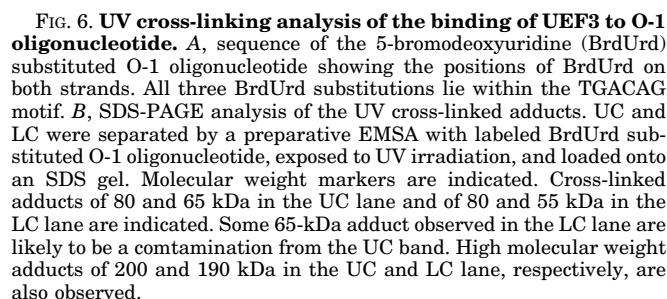
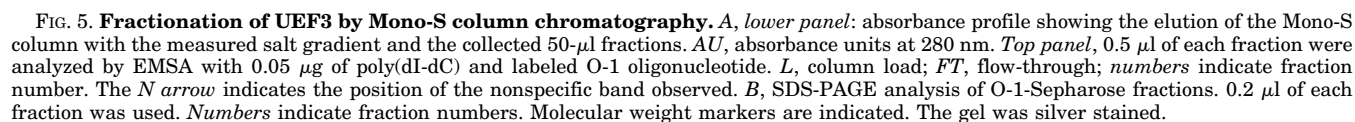
FIG. 4. Fractionation of UEF3 by O-1 DNA-Sepharose chromatography. A, lower panel: O-1-Sepharose absorbance profile showing the column flow-through (0–150 ml), the 200 mM KCl wash (150–225 ml), and the elution (225–250 ml). Inset, blow-up of elution profile from 225 to 245 ml showing the absorbance of the collected 1-ml fractions together with the theoretical salt gradient. AU, absorbance units at 280 nm. Top panel, 1 μ l of each fraction were analyzed by EMSA with 0.5 μ g of poly(dI-dC) and labeled O-1 oligonucleotide. L, column load; FT, flow-through; numbers indicate fraction number. B, SDS-PAGE analysis of O-1-Sepharose fractions. 1 μ l of each fraction was used. L, load; FT, flow-through; numbers, fraction numbers. Molecular weight markers are indicated. The gel was silver stained.

analysis (Fig. 4B). UEF3 activity was eluted from 270 to 450 mM KCl, correlating with the observed absorbance peak. Active fractions displayed major bands of 40, 50, and 64 kDa on SDS-PAGE. The enriched 64-, 50-, and 40-kDa bands will from now on be referred to as p64, p50, and p40. Based on EMSA and SDS-PAGE, fractions 5–12 were selected for the last chromatography step. These fractions were pooled, dialyzed, and applied onto the PC 1.6/5 Mono-S column for the SMART system (Fig. 5A). The column was eluted with a 100–1000 mM KCl gradient. Screening of the fractions by EMSA showed that UEF3 eluted as a small peak in fractions 11–14 from 220 to 350 mM KCl. SDS-PAGE analysis, Fig. 5B, showed that fractions 11–13 which had the highest activity also had the highest level of p64, p50, and p40. The less active adjacent peak, fractions 14 and 15, contained other polypeptides including a 116-kDa band. The presence of this protein band correlated with a nonspecific binding activity seen by EMSA (arrow in Fig. 5A) which did not co-migrate with UC and LC.

The result of this last chromatographic step indicates that

p64, p50, and p40 all participate in the binding activity of UEF3. Table II summarizes the purification. This procedure gave a calculated yield of more than 1.7% as measured by UEF3 binding activity in EMSA and a purification of almost 5,000-fold. The major loss of UEF3 activity occurred after fractionation by O-1-Sepharose chromatography. The last Mono-S step gave another 5-fold loss. However, it is our experience that UEF3 loses activity upon purification and as such, the overall yield may be higher than the calculated 1.7%.

UV Cross-linking—In order to test for the DNA binding activity of the three protein bands observed in the purified preparation of UEF3, we employed UV cross-linking analysis (Chodosh *et al.*, 1986) in which fractions from the Mono-S column were cross-linked to O-1 oligonucleotide in which the thymidine residues overlapping the TGACAG motif were substituted by 5-bromodeoxyuridine (Fig. 6A). In gel cross-linking was performed after EMSA separation of UC and LC. The presence of 5-bromodeoxyuridine residues did not interfere with binding of UEF3 as determined by comparison with an



migrating at 65 kDa was seen, together with two weaker bands at 80 and 200 kDa. From LC, a strong 55 kDa and two weaker 80- and 190-kDa adducts were observed. Estimating a mass contribution of 15 kDa for the 31-mer oligonucleotide itself, the 55-, 65-, and 80-kDa DNA-protein adducts appear to correspond to p40, p50, and p64. In the SDS-PAGE gel of Fig. 5*B*, however, we did not observe any bands that would correspond to the high molecular weight cross-linked adducts of 200 and 190 kDa. The nature of these bands has not been further investigated.

In conclusion, we have found evidence for a direct DNA interaction of the three proteins present in purified UEF3. The 55-kDa cross-linked adduct may correspond to p40, the 65-kDa adduct to p50, and the 80-kDa adduct to p64. Furthermore, the cross-linking data also give information on the composition of UC and LC. In fact, Fig. 6 shows that LC contains p40 and p64 while UC contains p50 and p64. As such, UC and LC would be formed by p40 or, respectively, p50 with p64 as a common factor. The suggested composition of UC and LC also seems to be supported by the comparison of EMSA and SDS-PAGE analysis as seen from Fig. 5, *A* and *B*. Fraction 11 from the Mono-S column contained mainly UC activity and was mainly composed of p64 and p50. Similarly, LC-enriched fraction 13 displayed mainly p40 and p64.

Methylation Interference—To further map the binding site for UEF3 and investigate any differential DNA binding of UC and LC, we performed methylation interference (Siebenlist and Gilbert, 1980) using O-1 as a probe. Under the conditions of methylation used here, the methylated residues are mainly the N-7 position of G and, to a lesser extent, the N-3 position of A. After methylation, the O-1 oligonucleotide was 5'-end labeled on either the top or the bottom strand and used as a probe in EMSA with purified UEF3. The oligonucleotide retained in the UC, LC, and free probe (taken from a lane run without UEF3) were extracted, cleaved at the methylated sites, and run on a sequencing type gel (Fig. 7). Methylation of several G and A residues on the top and bottom strands strongly interfered with

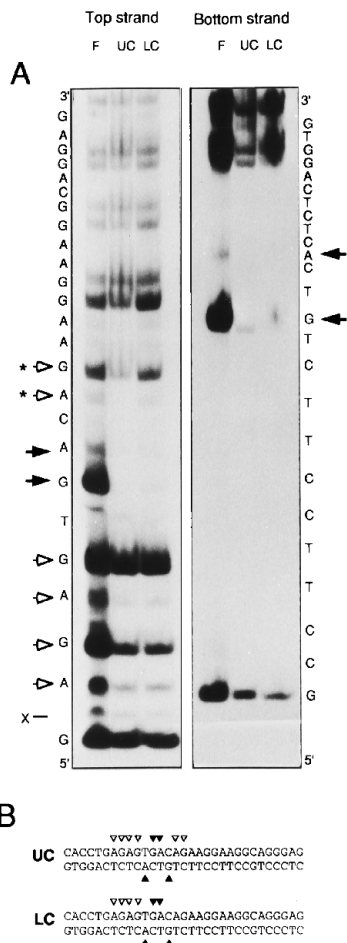


FIG. 7. Methylation interference analysis of UEF3 binding to O1 oligonucleotide. A, methylation interference reaction. Methylated O-1 oligonucleotide labeled on either the top strand or the bottom strand was used with purified UEF3 in a preparative EMSA. The DNA from recovered UC or LC bands was cleaved and separated on a denaturing sequencing-type gel along with unbound (F) probe. The sequence of the oligonucleotide corresponding to the observed bands is shown. Filled arrows show the bases where strong interference was observed; open arrows, weak interference; open arrows with asterisk, weak interference specific for UC only. Band marked X is an artificial band co-migrating with the dye-front. B, summary of the methylation interference on UC and LC. Closed and open triangles indicates strong and weak interference, respectively.

binding (marked with a *thick arrow*). This interference was equal for both UC and LC and lay within the proposed TGA-CAG binding motif (methylation position TGACAG/CTGTCA). Methylation of the last two residues of the TGACAG motif (TGACAG) weakly interfered with binding of only the UC (Fig. 7, *asterisk*). These residues are, however, important for binding, as an oligonucleotide where these two residues were mutated failed to form either UC and LC with purified UEF3 (data not shown). In addition, methylation of 4 residues directly upstream of the TGACAG motif weakly interfered with binding of both UC and LC (*open arrows*); of these, only the first two are conserved between the O-1 and O-17 oligonucleotides. These data taken together indicate that TGACAG is the core binding motif for UEF3, and that the same nucleotides within this motif are important for both UC and LC binding.

Binding of UEF3 to the NIP Element—The core UEF3 binding motif TGACAG has a five out of six nucleotide identity with the NIP binding site, important for AP-1 activation of the promoters of the stromelysin and interleukin-3 genes (Mathey-Prevot *et al.*, 1990; Sirum-Connolly and Brinckerhoff, 1991).

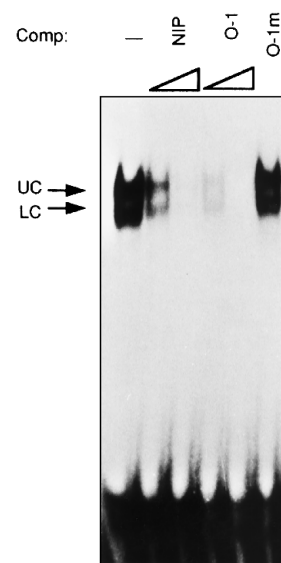


FIG. 8. UEF3 binds the NIP element with the same specificity of u-COM. Purified UEF3 was incubated with labeled IL-3 NIP element (sequence in Table I) and 0.1 μ g of poly(dI-dC) and separated by EMSA. Binding of UC and LC to the NIP element is indicated. Competition was carried out with a 50- and 500-fold excess of unlabeled NIP and O-1, and with a 500-fold excess of O-1m oligonucleotides, respectively.

Crude extracts from HepG2 cells were previously shown to contain an activity capable of binding the NIP element from the stromelysin gene in EMSA (Nerlov *et al.*, 1991). Because of the sequence homology between NIP and the UEF3-binding sequence (Table I), we tested purified UEF3 from HeLa cells for binding to a NIP probe using the stromelysin NIP element and EMSA. The purified protein did indeed bind the NIP oligonucleotide as shown by competition as well as direct binding assays. Fig. 8 shows that the binding was specific as it could be competed by 50-fold excess unlabeled NIP or O-1 oligonucleotides, but not by a 500-fold excess of O-1m.

Microsequencing of UEF3/p64—In order to obtain partial amino acid sequence information, fractions 11–13 containing highly enriched factor (see Fig. 5) were pooled, precipitated, and run in three slots of a mini SDS-PAGE gel (Matsudaira and Burgess, 1978). Proteins were electroblotted onto a polyvinylidene difluoride membrane (Bauw *et al.*, 1989) and detected by Amido Black staining. The p64 band, containing approximately 10 μ g of protein was subjected to *on membrane* trypsin digestion and the peptides were separated by narrow bore (0.2 \times 25 cm) reverse phase (C4) column chromatography. Fig. 9 shows the corresponding high performance liquid chromatography traces; the peptides indicated with an *arrow* or with *dots* were sequenced. In three cases, we obtained sequences derived from porcine trypsin. Peptide number 14 (*arrow*) showed a unique sequence as deduced from the yields of the phenylthiohydantoin residues recovered after each cycle (Table III). The deduced sequence is NH₂-XWLFQHHGHPYPTDEK. The first amino acid could not be unambiguously identified because we measured both serine and methionine after the first Edman cycle. Since serine is a frequent contaminant at the start of Edman degradation, we tentatively assume that the less abundant methionine residue is the real first residue in this peptide. The sequence of this peptide was searched in the GenBank and EMBL nucleotide data bases as well as the PIR and Swissprot protein data bases using the GCG package and FASTA and TFASTA programs as well as the BLAST-netserver Blastn, Blastp, and Tblastn programs. These searches did not identify any homologies between the 17-amino acid sequence of the p64

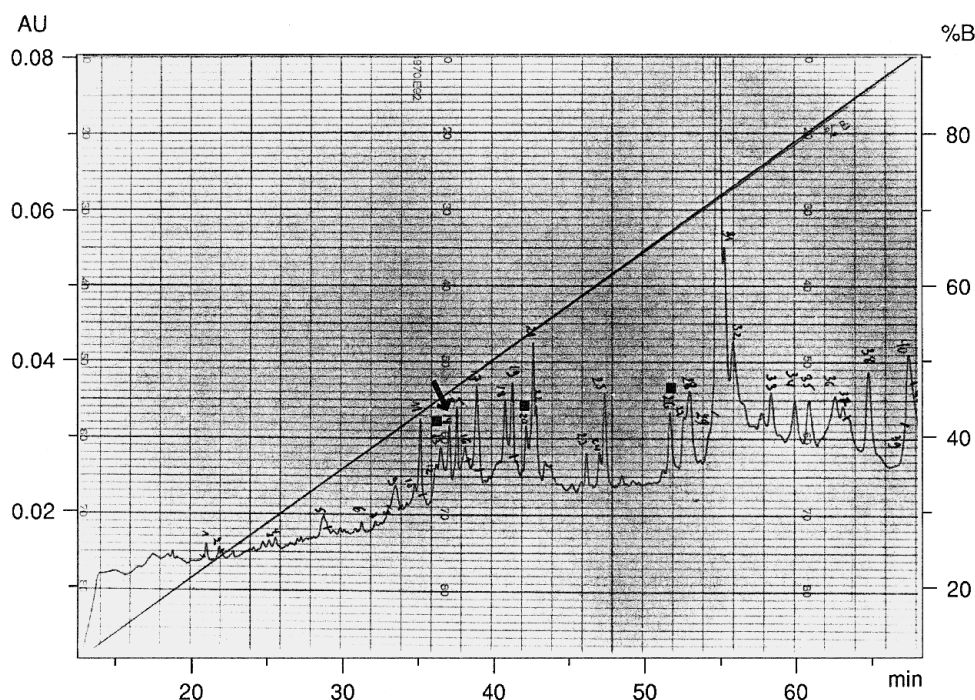


FIG. 9. Reverse phase (C4) column chromatography analysis of the p64 band of UEF3. Vydac 218TP54 column (internal diameter 4.6 mm, 250 mm height) was eluted at a flow of 1 ml/min with a linear gradient of solution A (0.1% trifluoroacetic acid) and buffer B (0.1% trifluoroacetic acid in 70% acetonitrile) (see "Materials and Methods"). The percentage of buffer B is indicated by the straight line. The left ordinate shows the optical density at 214 nm (AU, absorbance units), the abscissa the elution time in minutes. The arrow indicates peak 14 which gave the sequence shown in Table III. The other peaks identified with filled squares show other peptides used for microsequencing (see "Results").

TABLE III
Sequence analysis of one peptide from UEF3/p64
Peptide 14 from the HPLC run of Fig. 9.

Cycle No.	Amino acid residues and yield (pmol)
1	Ser, 17.8 ^a Met, 7.2
2	Trp, 7.8 ^b
3	Leu, 17.5
4	Phe, 14.4
5	Gln, 12.8
6	His, 5.0
7	Ile, 12.0
8	Gly, 14.0
9	His, 4.8
10	Pro, 12.1
11	Tyr, 8.8
12	Pro, 9.9
13	Thr, 5.0
14	Glu, 7.2
15	Asp, 7.2
16	Glu, 6.6
17	Lys, 3.1

^a The assignment of the first residue has some doubt due to a rather high background. Met is preferred as the amino-terminal amino acid of this peptide, since high contamination with serine is often seen at the start of the Edman degradation.

^b Phenylthiohydantoin-Trp and diphenylurea co-elute in the phenylthiohydantoin analysis. The given yield of phenylthiohydantoin-Trp is corrected for subtracting diphenylurea by extrapolation.

peptide and any known sequence, showing that the p64 part of UEF3 is a novel protein.

DISCUSSION

A 106-base pair conserved enhancer is located about 2 kilobase pairs upstream of the transcription start site in the human, mouse, and porcine urokinase genes (see Fig. 1A). This enhancer has been shown to mediate both the high basal transcription activity of the human, as well as the induction by phorbol 12-myristate 13-acetate and epidermal growth factor in the mouse and the human uPA genes (Verde *et al.*, 1988;

Rørth *et al.*, 1990). The activity of the uPA enhancer is dependent on the integrity of each of two protein binding sites: the upstream combined PEA3-AP-1 site and the downstream consensus AP-1 site (Rørth *et al.*, 1990; Nerlov *et al.*, 1991). These two sites are separated by the 74-base pair COM sequence, made up of two regions endowed with specific protein binding activity which overall bind 4 different factors (UEF 1–4). Three mutations in the COM element affect the inducibility of the uPA gene, and the combined mutations totally destroy enhancer activity (Nerlov *et al.*, 1992). A further characterization of the different UEFs is therefore important to understand the function of the COM element. In this paper we have addressed the nature of UEF3, since this factor seems also to bind similar regulatory sites in other promoters. In three promoters, uPA, stromelysin, and IL-3, the UEF3 binding site maps close to a consensus AP-1 site and to a site binding *Ets*-family proteins (Fig. 1B) (Gottschalk *et al.*, 1993; Mathey-Prevot *et al.*, 1990; Nerlov *et al.*, 1991; Rørth *et al.*, 1990; Sirum-Connolly and Brinckerhoff, 1991), suggesting similar regulatory features.

In the O-17 oligonucleotide spanning the u-COM element, the UEF3 binding site overlaps with binding sites for UEF1, -2, and -4.³ We have, however, identified an oligonucleotide (O-1) in another region of the uPA promoter (–1500 to –1470), that binds solely UEF3. By EMSA, UEF3 forms two complexes, UC and LC, which are common to O-1 and O-17. Competition and mutagenesis studies suggest that both complexes constitute the same DNA binding activity, UEF3 (Fig. 2).

We have purified UEF3 from HeLa cell nuclei to near homogeneity by different standard chromatography steps, including DNA affinity. The purified UEF3 has the same DNA-binding properties of the HeLa nuclear extract as it binds the same sequence, forms two complexes UC and LC, and is affected by the same mutations. The purification shows that UEF3 is formed by three polypeptides: p40, p50, and p64. All three

³ D. De Cesare, personal communication.

polypeptides are able to interact with DNA, as shown by the UV cross-linking experiment, with p64 giving the weakest binding. Both UV cross-linking data and the Mono-S chromatography elution profile suggest that UC was formed by p64 and p50 subunits, LC by p64 and p40. The two complexes may thus have the p64 subunit in common. This implies that the purified UEF3 should contain p64 in a 2:1:1 ratio with respect to p40 and p50, which is actually seen by silver-stained SDS-PAGE (Fig. 5B).

A 5,000-fold purification and a yield of 1.7% show that UEF3 is a low abundant factor. It is important to note, however, that all activity and recovery measurements are based on the EMSA data. In fact, after passage through the mutated DNA column, we observed an increase in specific activity which indicates that interfering factors partially masked the true activity of UEF3 (Table II). In addition, after specific DNA chromatography, the purified UEF3 rapidly lost its activity. Despite these limits, the purification procedure presented here gave a purification and yield comparable to procedures employed to purify NF- κ B factors (Hansen *et al.*, 1994).

UC and LC can be partly separated by S-Sepharose chromatography, reflecting the differences in the two complexes. However, during the course of purification, we were never able to completely separate UC and LC binding activities, neither using other types of ion exchange and general affinity chromatography nor by specific DNA-affinity matrices different from the one described here. This suggests that the protein composition of UC and LC are very similar both in structure as well as DNA binding. Since the UC and LC contain the same p64 subunit, but differ in the p50 *versus* p40 subunit, we speculate that p50 and p40 may be products of very similar genes or arise from alternatively spliced mRNA from the same gene. A third possibility is that p40 is a cleaved form of p50. This degradation may, however, be physiological since we always observed an equal presence of UC and LC, both when using freshly prepared as well as older nuclear extracts (not shown).

The purification procedure gave a sufficient yield of purified UEF3 to obtain a 17-amino acid residue sequence by microsequencing of the p64 polypeptide. This sequence was obtained after tryptic digestion and therefore probably represent an internal sequence of p64. After another purification we have tried N-terminal sequencing of all of the three subunits: p40, p50, and p64. In all cases the sequencing failed, probably due to blockage of the N-terminal of these polypeptides. Searches in data bases for homologous peptide sequences revealed none, showing that p64 is a novel protein and that UEF3 as such is a novel factor.

We have mapped the binding site for UEF3 by mutation and methylation interference studies. These demonstrate that the sequence TGACAG is the core binding site for UEF3. Mutation or methylation of 5 out of 6 nucleotides in the TGACAG motif interfered with binding of UEF3. In addition, methylation interference showed that nucleotides flanking the TGACAG motif may also participate in binding. These flanking bases are, however, not conserved in the O-1 and O-17 oligonucleotides. If this weak interference is caused by nonspecific steric hindrance or by methylation of bases important for the interaction of

UEF3 with its binding site remains to be determined. Searching the transcription factor data base for known factors able to bind the UEF3 sites did not reveal any, supporting the possibility that UEF3 is a novel factor.

The UEF3 binding site, however, has a five out of six nucleotide homology to the NIP binding site (TGACAG *versus* TGGCAG). The NIP element binds a so far unknown factor that acts as an AP-1 dependent repressor of transcription in the IL-3 promoter (Mathey-Prevot *et al.*, 1990) and as an AP-1 activity modulator in the stromelysin promoter (Sirum-Connolly and Brinckerhoff, 1991). In addition, the UEF3 binding site is homologous to a negative regulatory element, the ICK-1 element, found in promoters of the human LD-78 (Nomiya *et al.*, 1993) gene and in its murine counterpart, the MIP-1A gene (Ritter *et al.*, 1995). The ICK-1 element binds four different complexes, of which one is a negative factor co-migrating with the NIP activity in EMSA (Nomiya *et al.*, 1993). The base change in the NIP element occurs at a position which does contact UEF3 (Fig. 7). However, purified UEF3 specifically binds the NIP element (Fig. 8). Our data therefore show that UEF-3 is one of the NIP-binding factors. We are now proceeding with the molecular cloning of cDNAs encoding the various subunits of UEF3, and thereby investigate the nature of this novel factor and the mechanisms involved in regulating transcriptional activity of various promoters.

Acknowledgments—We are grateful to Dario DeCesare, Claus Nerlov, and Morten Johnsen for many fruitful discussions.

REFERENCES

- Blasi, F., and Verde, P. (1990) *Semin. Cancer Biol.* **1**, 117–126
- Bauw, G., Van Damme, J., Puype, M., Vandekerckhove, J., Gesser, B., Ratz, G. P., Lauridsen, J. B., and Celis, J. F. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 7701–7706
- Cassady, A. I., Stacey, L. J., Nimmo, K. A., Murphy, K. M., Ahe, D. V. D., Pearson, D., Botteri, F. M., Nagamine, Y., and Hume, D. A. (1991) *Nucleic Acids Res.* **19**, 6839–6847
- Chodosh, L. A., Carthew, R. W., and Sharp, P. A. (1986) *Mol. Cell. Biol.* **6**, 4723–4733
- Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489
- Gottschalk, L. R., Giannola, D., and Emerson, S. G. (1993) *J. Exp. Med.* **178**, 1681–1692
- Gutman, A., and Wasylyk, B. (1990) *EMBO J.* **9**, 2241–2246
- Hansen, S. K., Baeuerle, P. A., and Blasi, F. (1994) *Mol. Cell. Biol.* **14**, 2593–2603
- Heukeshoven, J., and Dernick, R. (1988) *Electrophoresis* **9**, 28–32
- Kadonaga, J. T., and Tjian, R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5889–5893
- Larsson, L. I., Skriver, L., Nielsen, L. S., Grøndahl-Hansen, J., Kristensen, P., and Danø, K. (1984) *J. Cell Biol.* **98**, 894–903
- 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
- Matsudaira, P., and Burgess, D. R. (1978) *Anal. Biochem.* **87**, 1386–1396
- Nerlov, C., Rørth, P., Blasi, F., and Johnsen, M. (1991) *Oncogene* **6**, 1583–1592
- Nerlov, C., de Cesare, D., Pergola, F., Caracciolo, A., Blasi, F., Johnsen, M., and Verde, P. (1992) *EMBO J.* **11**, 4573–4582
- Nomiya, H., Hieshima, K., Hirokawa, K., Hattori, T., Takatsuki, K., and Miura, R. (1993) *Mol. Cell. Biol.* **13**, 2787–2801
- Ritter, L., Bryans, M., Abdo, O., Sharma, V., and Wilkie, N. M. (1995) *Mol. Cell. Biol.* **15**, 3110–3118
- Rørth, P., Nerlov, C., Blasi, F., and Johnsen, M. (1990) *Nucleic Acids Res.* **18**, 5009–5017
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Sappino, A. P., Huarte, J., Belin, D., and Vassalli, J. D. (1989) *J. Cell Biol.* **109**, 2471–2479
- Siebenlist, U., and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 122–126
- Sirum-Connolly, K., and Brinckerhoff, C. E. (1991) *Nucleic Acids Res.* **19**, 335–341
- Verde, P., Boast, S., Franze, A., Robbiati, F., and Blasi, F. (1988) *Nucleic Acids Res.* **16**, 10699–10716