Two Distinct Factors Bind to the Rabbit Uteroglobin TATA-Box Region and Are Required for Efficient Transcription

JÖRG KLUG,* STEFAN KNAPP,† IVONE CASTRO, AND MIGUEL BEATO

Philipps-Universität Marburg, Institut für Molekularbiologie und Tumorforschung, 35037 Marburg, Germany

The rabbit uteroglobin gene is expressed in a variety of epithelial cell types like the lung Clara cells and the glandular and luminal epithelial cells of the endometrium. Expression in Clara cells is on a high constitutive level, whereas expression in the rabbit endometrium is under tight hormonal control. One important element of the rabbit uteroglobin gene mediating its efficient transcription in two epithelial cell lines from human endometrium (Ishikawa) and lung (NCI-H441) is its noncanonical TATA box (TACA). Here, we show that two factors (TATA core factor [TCF] and TATA palindrome factor [TPF]) different from the TATA-box binding protein bind to the DNA major groove at two adjacent sites within the uteroglobin TATA-box region and that one of them (TCF) is specifically expressed in cell lines derived from uteroglobin-expressing tissues. The binding sites for TCF and TPF, respectively, are both required for efficient transcription in Ishikawa and NCI-H441 cells. Mutation of the TACA box, which we show is a poor TATA box in functional terms, to a canonical TATA motif does not affect TCF and TPF binding. Therefore, we suggest that the function of the unusual cytosine could be to reduce rabbit uteroglobin expression in cells lacking TCF and that the interaction of TATA-box binding protein with the weak TACA site is facilitated in TCF- and TPF-positive cells.

The expression of eucaryotic genes is mainly regulated on the level of transcription, with the formation and activation of a stable preinitiation complex being the main regulatory switch. In most cases, the formation of the preinitiation complex starts with the binding of TFIID, a multisubunit complex consisting of the TATA-box binding protein (TBP) and TBP-associated factors, to the TATA box as the most important core promoter element. After assembly of this template-committed complex as the nucleating event, TFIIA, TFIIB, RNA polymerase II/TFIIF, TFIIE, TFIIH, and TFIIJ join and form the complete preinitiation complex (10, 63). Activator proteins binding to upstream modulatory promoter elements assist in the efficient assembly of the preinitiation complex.

In the last few years, we have begun to study the mechanisms directing the regulated expression of the rabbit uteroglobin gene in various ontogenetically unrelated epithelial tissues. In the rabbit lung, the uteroglobin promoter is constitutively active at a high level (15, 28), whereas it is almost silent in the endometrium but gets dramatically activated by ovarian hormones in the preimplantation phase of pregnancy (34). In both tissues, the same defined initiation start site is used (53). DNase I footprinting (32), promoter deletion, and linker-scanning analysis (52) revealed the presence of seven functionally important elements in the promoter region up to −270 bp (Fig. 1A). Among them is at least one functional binding site for Sp1 family transcription factors (20), although the promoter sequence in general is not GC rich.

Another important element is the noncanonical TATA-box motif (TACA box [Fig. 1]). Alignment of the rabbit uteroglobin TATA-box sequence with the corresponding sequences of the human and rat genes (whose gene products were formerly called Clara cell 10-kDa secretory proteins [CC10] [60]) shows that the human and rat genes possess a canonical TATA box, whereas the rabbit promoter has a TACA box with a cytosine instead of a thymine in the third position of the consensus sequence. Though TBP/TFIID has been reported to bind to noncanonical TATA box sequences (47, 59), all available data on TACA boxes point to the fact that it is a weak recognition site for TBP (7, 24, 39). Therefore, the uteroglobin promoter is related to a subclass of non-GC-rich and TATA-less promoters which are regulated during development or differentiation and initiate transcription at only one main site (50). In order to find out why this noncanonical and weak TATA box is functionally still very important for uteroglobin transcription, we carried out a thorough investigation of this promoter element.

In previous DNase I footprinting experiments, it was shown that the rabbit uteroglobin TATA-box region (−40 to −18) is specifically protected by nuclear extract proteins from Ishikawa cells, which are derived from a human endometrial adenocarcinoma (37), but not by nuclear extract proteins from HeLa cells (32). Accordingly, the uteroglobin promoter drives efficient transcription of a reporter gene in Ishikawa cells but not in HeLa cells (32, 52).

Here, we show that two factors (TATA core factor [TCF] and TATA palindrome factor [TPF]) different from TBP are binding to the DNA major groove at two adjacent sites within the uteroglobin TATA-box region. Each of these factors is functionally important, and only one of them (TCF) is specifically expressed in cell lines derived from uteroglobin-expressing tissues. Because mutation of the noncanonical TATA box (TACA) to a canonical TATA motif does not affect TCF and TPF binding, we suggest that the TACA sequence serves to reduce rabbit uteroglobin expression in cells lacking TCF,

* Corresponding author. Mailing address: Institut für Molekularbiologie und Tumorforschung, Emil-Mannkopff-Str. 2, 35037 Marburg, Germany. Phone: (49) 6421-28-6545. Fax: (49) 6421-28-5398. Electronic mail address: Klug%VSIMT1.decnet@vax.HRZ.Uni-Marburg.DE.

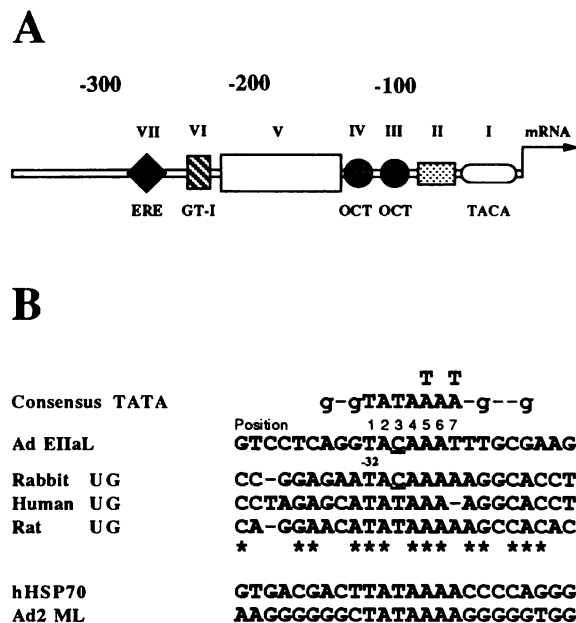† Present address: Karolinska Institute, Center for Structural Biochemistry Novum, S-14157 Huddinge, Sweden.

A

-300   -200   -100

VII VI   V   IV III II   I   mRNA

ERE GT-I   OCT OCT   TACA

B

	T T
Consensus TATA	g-gTATAAAA-g--g

Position   1 2 3 4 5 6 7

Ad EIIaL	GTCCTCAGGTA_CAAATTTGCGAAG
	-32
Rabbit UG	CC-GGAGAATA_CAAAAAGGCACCT
Human UG	CCTAGAGCATATAAA-AGGCACCT
Rat UG	CA-GGAACATATAAAAAGCCACAC
	* ** *** *** ** ***

| hHSP70 | GTGACGACTTATAAAACCCCAGGG |
| Ad2 ML | AAGGGGGGCTATAAAAGGGGGTGG |

FIG. 1. Model of the rabbit uteroglobin (UG) promoter structure. (A) Seven transcriptionally important regions (I to VII) have been mapped by DNase I footprinting (32), deletion, and linker-scanning analysis (52). Mutants affecting the noncanonical TATA element (TACA box, region I) display a very strong phenotype. Regions III and IV contain inverted repeats of an A/T-rich sequence resembling degenerated octamer (OCT) elements, which are bound by Oct1 in vitro (reference 45 and unpublished results). Region VI contains a binding site for Sp1 family transcription factors (20), and region VII contains an estrogen-responsive element (ERE [49]). (B) Comparison of the rabbit uteroglobin TATA box sequence (53) with the cognate human (60) and rat (21) sequences, the TATA box of the adenovirus EIIa late promoter (Ad EIIaL [24]), two well-characterized TATA boxes from the human HSP70 gene (hHSP70 [35]) and the adenovirus type 2 major late promoter (Ad2 ML [35]), and the consensus sequence (7). The C's converting the rabbit uteroglobin and adenovirus EIIaL TATA boxes into TACA boxes are underlined. Position 1 of the rabbit uteroglobin TATA box is located 32 bp upstream of the transcription start site. Asterisks denote conserved nucleotides within the uteroglobin sequences.

because of its apparently weak affinity for TBP. Simultaneously, expression of uteroglobin is selectively switched on in TCF-positive cells.

MATERIALS AND METHODS

Oligonucleotides and EMSAs. Oligonucleotides were synthesized on an Applied Biosystems 380A oligonucleotide synthesizer by using deoxynucleoside phoshoramidite chemistry (4). The sequences of the oligonucleotides used for electrophoretic mobility shift assays (EMSAs) are shown in Table 1. All oligonucleotides had an additional HindIII and SalI extension added to their 5' and 3' ends, respectively. As an unspecific competitor for all EMSA reactions, an oligonucleotide with a defined but random sequence was routinely used (5'-CAGCGACTAACATCGATCGC-3').

Annealing of single-stranded oligonucleotides was done by heating a mixture of complementary strands (10 nmol each) to 95°C in 100 μl of annealing buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 30 mM KCl) and cooling to room temperature over 2 h. Double-stranded DNA was labeled by filling in the ends with [α-^{32}P]dCTP and Klenow fragment. Labeled double-

TABLE 1. Sequences of oligonucleotides used for EMSAs

Oligonucleotide	Sequence[a]
ocTATA[b]	5'-TCTTGCCGGAGAATACAAAAAGGCACCTCGAGG-3'
Pal1	5'-TCTTGCCGGAGAATACAAAAAGTCACATCGAGG-3'
A/C	5'-TCTTGCCGGAGAATACAACAAGGCACCTCGAGG-3'
C/T	5'-TCTTGCCGGAGAATATAAAAAGGCACCTCGAGG-3'
Core	5'-TGGTACCAGATACAAAAAGTCTAGAGCTGCAGG-3'
Core+Pal	5'-TGGTACCAGATACAAAAAGGCACCTGCTGCAGG-3'
hHSP70	5'-TGGTACCACGACTTATAAAACCCCAGCTGCAGG-3'

[a] Underlined and boldface letters indicate uteroglobin and human HSP70 (hHSP70) promoter sequences, respectively. Double underlining indicates point mutations in ocTATA yielding the mutants listed to the left.

[b] ocTATA, O. cuniculus (rabbit) uteroglobin TATA-box region.

stranded DNA was separated from unannealed strands by gel electrophoresis on a native 6% polyacrylamide gel in Tris-borate-EDTA buffer.

For the construction of uteroglobin promoter-chloramphenicol acetyltransferase (CAT) constructs containing mutations in the TATA box region, PCR (42) was employed with the T7 primer (5'-TAATACGACTCACTATAGGG-3') as 5' primer and the lower strand of the TATA box sequence containing the desired point mutations and a suitable restriction site (EclXI) close by as 3' primers (Table 2). The fragments were amplified with the UG(−301)CAT deletion construct (32) as template. After restriction with HindIII-EclXI, the corresponding wild-type promoter fragment of UG(−301)CAT was exchanged for the mutagenized fragments. The mutations were verified by sequencing the TATA-box regions by the dideoxy chain termination method (44). The sequencing primer was located in the CAT gene, approximately 50 bp downstream of the cap site (5'-ATATCAACGGTGGTATATCCAGTG-3').

For EMSAs, 6 to 10 μg of nuclear extract was preincubated with 0.5 μg of unspecific competitor poly(dI-dC) in 10 mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid [pH 7.9])–100 mM KCl–1 mM dithiothreitol–0.5 mM MgCl$_2$–0.1 mM EDTA–8.5% glycerol for 10 min on ice. Subsequently, 0.5 ng of labeled double-stranded oligonucleotide (15,000 cpm, Cerenkov counting) was added to yield a final volume of 20 μl, and incubation was continued at room temperature for another 15 min. For competition experiments, an unlabeled oligonucleotide was added in excess as indicated before the labeled probe was added. Samples were analyzed on a 4% native polyacrylamide gel (acrylamide/bisacrylamide ratio, 40/1) in 0.05 M Tris base–0.045 M boric acid–0.5 mM EDTA. The gel was transferred to Whatman 3MM paper, dried under heat and a vacuum, and exposed to Kodak XAR5 film overnight at room temperature.

Cell culture and preparation of nuclear extracts. Ishikawa cells were grown as monolayers in minimal essential medium supplemented with 10% fetal calf serum, L-glutamine, and antibiotics. A total of 5 × 10^8 cells were collected in phosphate-buffered saline (PBS), and nuclear extracts were prepared according to the method of Shapiro et al. (46), except that the hypotonic buffer for the isolation of nuclei contained 0.5% skim milk. An aqueous stock solution containing 10% (wt/vol) skim milk (Difco) was centrifuged by sedimentation at 10,000 × g, and 1 volume of the supernatant was then added to 20 volumes of hypotonic buffer (29).

HeLa cells were grown in suspension in minimal essential medium modified for suspension culture (S-MEM; Gibco), supplemented with 5% fetal calf serum, nonessential amino acids, and antibiotics. Cells were continuously grown to a density of 3 × 10^5 to 5 × 10^5 cells per ml. A total of 0.9 × 10^9

TABLE 2. 3' Primers used for the construction of uteroglobin
promoter-CAT constructs containing mutations
in the TATA-box region

Sequence[a]	Oligonu-cleotide
3'-CTTATGTTTTTCAGTGTACT**GCCGGC**AGG-5'	Pal1
3'-CTTATATTTTTCAGTGTACT**GCCGGC**AGG-5'	C/T-Pal1
3'-CTTATGTTTTTCCTAGGACT**GCCGGC**AGG-5'	Pal2
3'-CTTATGTTGTTCCGTGGACT**GCCGGC**AGG-5'	A/C
3'-CTTATATTTTTCCGTGGACT**GCCGGC**AGG-5'	C/T
3'-CTTCCGTTTTTCCGTGGACT**GCCGGC**AGG-5'	GGCA
3'-CTTATGTT-TTCCGTGGACT**GCCGGC**AGG-5'	4A

[a] Underlined and boldface letters represent uteroglobin point mutations and
EclXI sites, respectively.

to 1.5 × 10⁹ cells were harvested in PBS during the logarithmic
phase, and nuclear extracts were prepared according to the
method of Dignam et al. (13).

CV-1 cells were grown in Dulbecco's minimal essential
medium (Gibco), supplemented with 10% fetal calf serum and
antibiotics. Nuclear extracts were prepared according to the
micromethod of Andrews and Faller (1). NCI-H441 cells were
cultivated in RPMI 1640 medium (Gibco) supplemented with
10% fetal calf serum, 0.3% glucose, and antibiotics. Nuclear
extracts were prepared according to the method of Shapiro et
al. (46) (described above). All buffers for the preparation of
nuclear extracts contained 0.5 mM phenylmethylsulfonyl fluo-
ride, 100 μM benzamidine, and 2.0 μg of aprotinin (Trasylol)
per ml.

Transfections. Ishikawa cells were transfected as described
previously (9) with the DEAE-dextran method with some
modifications. Eight micrograms of uteroglobin construct and
2 μg of pRSVL (12) were transfected, followed by a 2-min
dimethyl sulfoxide shock with 15% dimethyl sulfoxide in
medium and a chloroquine treatment for 2.5 h. The medium
was changed after 24 h, and cells were harvested 72 h after
transfection by scraping them into 300 μl of buffer A (15 mM
Tris-HCl [pH 8.0], 60 mM KCl, 15 mM NaCl, 2 mM EDTA,
0.15 mM spermine, 1 mM dithiothreitol, 0.4 mM phenylmeth-
ylsulfonyl fluoride [40]).

CV-1 cells were transfected by using the calcium phosphate
protocol (19) as described previously (30) with 4 μg of a
uteroglobin construct and 1 μg of pRSVL per 60-mm plate.
NCI-H441 cells were transfected with lipofectamine (Gibco)
according to the instructions of the supplier and with 4 μg of a
uteroglobin construct and 1 μg of pRSVL for 6 × 10⁵ cells
plated onto a 60-mm plate. DNA was mixed with 25 μl of
lipofectamine reagent, and the cells were incubated with this
mixture for 6 h before the medium was changed. CV-1 and
NCI-H441 cells were harvested 48 h after transfection by
scraping them into 150 μl of buffer A (described above). All
transfections were done at least twice.

CAT assays were performed with 100 μg of protein accord-
ing to the method of Gorman et al. (18), but with buffer A (see
above and reference 40) instead of 0.25 M Tris-HCl (pH 7.8).
CAT conversion was assayed by thin-layer chromatography,
and quantitation of [¹⁴C]chloramphenicol and its acetylated
derivatives was performed with an automated imaging scanner
(United Technologies Packard). The ratio of acetylated to
total chloramphenicol (percentage of conversion) was calcu-
lated.

Luciferase assays were done with 25 μg of protein as
described previously (3) with the alternate protocol. Light was
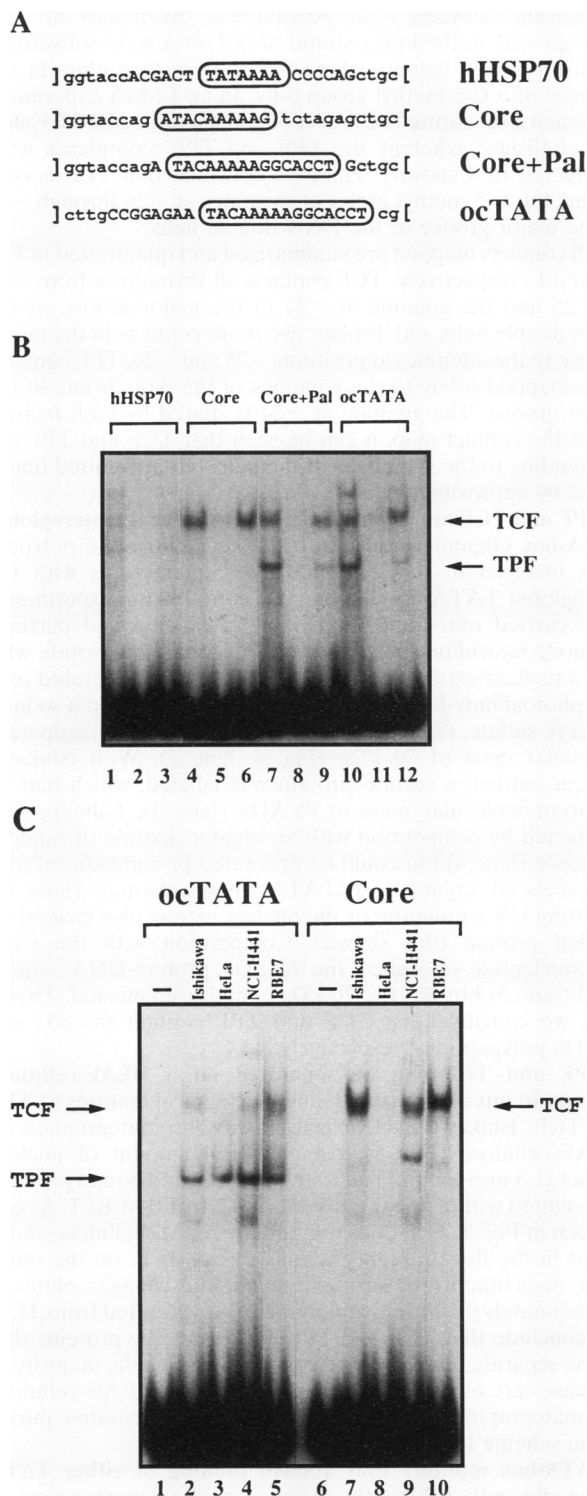measured with an AutoLumat LB 953 luminometer (Berthold)

by injecting 100 μl of a 0.4 mM luciferin solution into sample
tubes and measuring light output over 10 s. The values were
used to normalize the CAT conversion data for plate-to-plate
variations in transfection efficiency.

Probing DNA-protein interactions. The upper and lower
strands of the ocTATA oligonucleotide (described above)
were labeled by selectively filling in its 5' protruding ends (with
dTTP and [α-³²P]dCTP for the upper strand and dATP, dGTP,
and [α-³²P]dCTP for the lower strand) by standard procedures
(30). For methylation interference experiments, the Klenow
reaction mixture was phenol extracted and the oligonucleotide
was precipitated in the presence of 1 μg of glycogen and
methylated with dimethylsulfate (DMS) as described previ-
ously (31), but with an incubation time of 2.5 min at 20°C. For
KMnO₄ interference experiments, the oligonucleotide was
modified with KMnO₄ prior to labeling. One hundred nano-
grams of ocTATA oligonucleotide was mixed with 1 μg of
poly(dI-dC), modified, and reannealed as described previously
(56) but with a 10⁻⁴ M KMnO₄ solution for modification. All
labeled oligonucleotides were gel purified as described above.

Preparative mobility shift assays were set up with 1.0 × 10⁵
to 1.5 × 10⁵ cpm (Cerenkov counting) of oligonucleotide, 50 to
100 μg of nuclear extract, and 2.5 to 5.0 μg of poly(dI-dC) in
a volume of 20 μl with the Mg²⁺ concentration raised to 1 mM.
For methylation protection assays, 2 μl of 2% (vol/vol) DMS in
10 mM HEPES (pH 7.9)–100 mM KCl–1 mM MgCl₂–1 mM
dithiothreitol–0.1 mM EDTA–8.5% glycerol was added to a
completed binding reaction mixture with unmodified oligonu-
cleotide. The reactions were stopped with 2 μl of 60 mM
β-mercaptoethanol 2.5 min after incubation at 20°C and
loaded immediately onto a 4% EMSA gel (see above). After
separation of the complexes, the gel was electroblotted onto
DEAE-cellulose ion-exchange paper (DE 81; Schleicher &
Schuell) with a Trans Blot Cell (Bio-Rad). Transfer was
performed in 0.5× Tris-borate-EDTA buffer overnight in the
cold room at 18 V. The positions of free and complexed
oligonucleotide bands on the wet DE 81 paper were identified
by 1 h of autoradiography. Paper pieces containing the bands
were excised, transferred to a microcentrifuge tube, and
washed in (i) TE8 (10 mM Tris-HCl [pH 8.0], 1 mM EDTA)
and (ii) TE8–0.1 M NaCl for a few minutes each. DNA was
eluted with 300 μl of TE8–1.5 M NaCl by incubation at 65°C
for 2 h. The DE 81 paper pieces were removed by centrifuga-
tion through filter inserts for microcentrifuge tubes (Milli-
pore). After ethanol precipitation in the presence of 1 μg of
glycogen per sample, the DNA was dissolved in 90 μl of water.
Cleavage of the DNA backbone was performed by adding 10
μl of piperidine and incubating the mixture at 95°C for 30 min.
The reaction was stopped by the addition of 10 μl of 3 M NaAc
followed by ethanol precipitation. Alternatively a strong ade-
nine-weak guanine cleavage reaction was performed with the
methylation protection samples as described previously (31),
but with 100 μl of 0.1 M NaOH–1 mM EDTA for the
DNA-backbone cleavage and concentration of the DNA after
heat treatment by ethanol precipitation. Ten thousand counts
per minute of each sample was analyzed on a 15% sequencing
gel. Autoradiography was done on the frozen gel in a light-
tight box at −80°C with an intensifying screen.

RESULTS

**A ubiquitous factor and a cell-specific factor bind to par-
tially overlapping sites of the uteroglobin TATA-box region.** In
previous EMSA experiments with an oligonucleotide encom-
passing the rabbit (*Oryctolagus cuniculus*) uteroglobin TATA-
box region (−40 to −18 [ocTATA]) and nuclear extracts of the

A

]ggtaccACGACT TATAAAA CCCCAGctgc [hHSP70

]ggtaccag ATACAAAAAG tctagagctgc [Core

]ggtaccagA TACAAAAAGGCACCT Gctgc [Core+Pal

]cttgCCGGAGAA TACAAAAAGGCACCT cg [ocTATA

B

C

FIG. 2. EMSA of the rabbit uteroglobin TATA box. (A) The relevant sequences of the oligonucleotides used are shown. They all had the same length (41 nucleotides) after filling in with Klenow enzyme and the same end sequences (indicated by the rectangular brackets). Uteroglobin and heat shock promoter sequences are in uppercase. The TATA box and the uteroglobin 3' palindrome AG GcaCCT, when present, are boxed. hHSP70 contains the sequence of the DNase I footprint of TBP over the human HSP70 TATA box (35). (B) The oligonucleotides shown in panel A were end labeled and used in EMSA experiments with Ishikawa nuclear extract (lanes 1, 4, 7, and

human endometrial adenocarcinoma-derived cell line Ishikawa (37), two specific complexes were observed (32). One of the two complexes was specific for Ishikawa cells (Ishikawa-specific factor) and could not be detected with HeLa nuclear extract, whereas the second complex was ubiquitous. In order to define the DNA sequences within the ocTATA oligonucleotide required for the formation of these two complexes, we used oligonucleotides containing truncated uteroglobin promoter sequences (Fig. 2A) as probes in EMSA experiments (Fig. 2B). It turned out that the 10-bp TATA-box core region (Core) which lacks the palindrome (AGGcaCCT) immediately downstream of the TATA box and seven nucleotides upstream is sufficient for the generation of the Ishikawa-specific complex (Fig. 2B, lanes 4 to 6). For this reason and because the complex is not solely Ishikawa specific (described below), we propose to rename the Ishikawa-specific factor "TATA core factor" (TCF). Furthermore, an ocTATA variant (Core+Pal), with uteroglobin sequences upstream of -33 exchanged, yielded the same complex pattern as ocTATA itself (Fig. 2B, lanes 7 to 9). Since the core region alone produces only a TCF complex, we conclude that the downstream palindrome must be of crucial importance for the ubiquitous TATA factor, which we propose to call "TATA palindrome factor" (TPF). Although 10 bp centered around the TATA box are required and sufficient for the formation of the TCF complex, the downstream palindrome alone is important but not sufficient for the binding of TPF (not shown). Some nucleotides of the core region seem to be required in addition to the palindrome itself.

As a comparison, an oligonucleotide containing the corresponding TATA-box region of the human HSP70 gene (35) was used for EMSA experiments with Ishikawa nuclear extract (Fig. 2A). Apart from a series of weak and unspecific bands, no prominent specific complex could be detected (Fig. 2B, hHSP70, lanes 1 to 3). If a general TBP containing transcription factor-like TFIID would have been responsible for the complexes, as displayed by ocTATA, one would have expected to see the same complexes with hHSP70. This statement is supported by the finding that bacterially expressed and purified human TBP does bind to the human HSP70 TATA box, whereas only a weak complex is visible when the uteroglobin TATA box (ocTATA) is used (not shown).

Uteroglobin is not only expressed in the rabbit endometrium during the preimplantation phase (5) but is also expressed constitutively in the Clara cells of the lung in rabbits (2, 58) and other species (21, 27, 43, 48). Therefore, the human lung papillary carcinoma-derived cell line NCI-H441, which exhibits the morphological properties of Clara cells (17) and also recognizes the uteroglobin promoter (52), was analyzed for the presence of TCF. Furthermore, we tested another endometrium-derived rabbit cell line, RBE-7, which has been reported to express the UG gene (33). Both lines were found to be positive (Fig. 2C). Thus, TCF is specifically expressed in three cell lines derived from uteroglobin-expressing tissues, whereas it is not detectable in HeLa cells (Fig. 2C, lanes 3 and 8 [32]).

10). Competition experiments were performed with a 500-fold molar excess of an unspecific oligonucleotide of random sequence (lanes 3, 6, 9, and 12) and with the cognate unlabeled oligonucleotide (lanes 2, 5, 8, and 11). TCF and TPF are indicated by arrows. In panel C, EMSA experiments were performed with nuclear extracts from Ishikawa (lanes 2 and 7), HeLa (lanes 3 and 8), NCI-H441 (lanes 4 and 9), and RBE7 cells (lanes 5 and 10) and oligonucleotides ocTATA (lanes 1 to 5) and Core (lanes 6 to 10), respectively, as the DNA probe.

High-resolution mapping of protein-DNA contacts. After having determined the approximate binding sites of TCF and TPF, various interference and protection techniques were employed to more precisely define the DNA contacts of both factors (Fig. 3). For methylation interference experiments (Fig. 3A), the guanines and adenines of the ocTATA oligonucleotide were methylated at the N-7 and N-3 positions, respectively, with DMS. A preparative EMSA was performed with Ishikawa nuclear extract by using the modified probe. The DNA of the complex bands was recovered and analyzed on a 15% sequencing gel together with controls and sequencing reactions for orientation. For the TPF complex (Fig. 3A, lanes TPF Bound), strong interference is observed upon methylation of guanines at −23 and −24 in the upper strand and methylation of guanines at −19, −20, and −22 in the lower strand. All of these guanines are contained within the palindrome AGGcaCCT just downstream of the TATA box, which correlates with the results of the EMSA experiments with variant ocTATA oligonucleotides (described above). No significant interference was observed for the TCF complex.

For methylation protection experiments, completed preparative EMSA reaction mixtures were treated with 0.2% DMS. Complexes were separated in an EMSA gel, and the recovered DNA was cleaved either with piperidine for the analysis of guanines (Fig. 3B) or with sodium hydroxide after acid depurination for analysis of adenines and guanines (Fig. 3C). In the case of TPF, the same guanines with which methylation interfered with binding (Fig. 3A) were also protected against methylation (Fig. 3B, lanes TPF Bound). The same results were obtained when the DMS treatment was not performed prior to electrophoretic separation but was instead performed afterward within the EMSA gel (38). With some preparations of Ishikawa nuclear extract, another specific complex was formed (see Fig. 7) that migrates faster than TPF and which we call "TPF′". DNA recovered from the TPF′ complex was also analyzed in the methylation protection (Fig. 3B, lanes TPF′ Bound) and KMnO₄ interference experiments (described below). From the observation that TPF′ displays the same methylation protection pattern as TPF and also behaves like TPF in the KMnO₄ interference experiment, we conclude that TPF′ is related to TPF and most likely represents a proteolytic degradation product. For the TCF complex, significant protection of guanine −24 is observed, which could not be seen in the complementary interference experiment.

When a strong adenine-weak guanine reaction was performed with the recovered DNA samples, weak but significant protection of adenines at −25 and −26 in the upper strand was observed for the TCF complex (Fig. 3C). Compared with the strong G reaction, signal intensities were in general much weaker, although the amounts of input DNA were the same.

So far, only one guanine contact (at −24) and two weak contacts with adenines at −25 and −26 had been mapped for TCF. Therefore, it seemed likely that the major contacts were directed to the thymines, and a KMnO₄ interference experiment (56) was performed with the T-rich lower strand of ocTATA (Fig. 3D). As suspected, the oxidation of thymines at −25 through −29 interferes with the binding of TCF. The oxidation of a thymine with KMnO₄ starts with glycolization of the C-5-C-6 double bond followed by ring opening and further oxidation via an aldehyde intermediate leading eventually either to complete removal of the base or to the formation of 2′-deoxy-D-ribofuranosylurea (23). Therefore, one cannot distinguish if an observed interference is due to the removal of the C-5 methyl group, which is exposed in the major groove, or the missing N-3, which lies in the minor groove. In order to make

a decision between both possibilities, thymidines at −25 through −36 in the lower strand of ocTATA were substituted for uridine (see below), which differs from thymidine in the absence of a C-5 methyl group only. In an EMSA experiment with such a substituted ocTATA oligonucleotide, the TCF shift was abolished, whereas the TPF and TPF′ complexes were unaffected (not shown). Thus, we conclude that TCF is contacting the C-5 methyl group of thymines at −25 through −29 in the major groove of the DNA double helix.

All contacts mapped are summarized and quantitated in Fig. 3E and F, respectively. TCF contacts all thymidines from −29 to −25 and the guanine at −24 in the major groove of the DNA double helix and displays two more contacts in the minor groove to the adenines at positions −25 and −26. TPF contacts were mapped solely to the guanines of the palindrome in the major groove. The guanine at −24 is shared by both factors. From the contact map, it can be seen that TCF and TPF are not binding to the same face of the helix but are shifted out of phase by approximately 90°.

TPF and TCF are photoaffinity labeled with the uteroglobin TATA-box oligonucleotide. In order to identify the polypeptides involved in the TPF and TCF interactions with the uteroglobin TATA-box region, UV cross-linking experiments were carried out. Incubation of a ³²P-labeled and partially bromodeoxyuridine-substituted ocTATA oligonucleotide with HeLa nuclear extract and irradiation with UV generated only one photoaffinity-labeled protein, which migrated on a sodium dodecyl sulfate (SDS)-polyacrylamide gel with an apparent molecular mass of 70 kDa (Fig. 4, lane 2). With Ishikawa nuclear extract, a second protein was labeled, which had an apparent molecular mass of 85 kDa (lane 3). Labeling was unaffected by competition with an oligonucleotide of random sequence (lane 4) but could be prevented by competition with an excess of unlabeled ocTATA oligonucleotide (lane 6). Omitting UV irradiation or the nuclear extract also yielded no labeled protein (not shown). Competition with the Core oligonucleotide abolished the 85-kDa protein-DNA adduct band (lane 5) but left the 70-kDa complex unaffected. Therefore, we conclude that TCF and TPF contain the 85- and 70-kDa polypeptides, respectively.

TPF and TCF can be separated on a DEAE-cellulose column. In our attempts to define biochemical features of TPF and TCF, Ishikawa nuclear extract was chromatographed on DEAE-cellulose (Fig. 5). An analytical amount of nuclear extract (1.4 mg) was applied to the column, and bound proteins were eluted with a linear gradient of 0.1 to 1.0 M KCl. As can be seen in Fig. 5, TCF does not bind to DEAE-cellulose and is found in the flowthrough fractions. Intact TPF, on the other hand, does bind to the strong basic anion exchanger, elutes at approximately 0.4 M salt, and thus can be separated from TCF. We conclude that TCF and TPF are formed by proteins that can be separated under native conditions. Since the majority of nucleases are also found in the flowthrough, DEAE-cellulose chromatography is a suitable first step in a preparative purification scheme for TPF.

TATA-box mutants that abolish binding of either TATA factor efficiently inhibit the expression of a reporter gene in Ishikawa and NCI-H441 cells but not in CV-1 cells. Point mutations were introduced into the uteroglobin TATA-box region that were expected to affect the binding of one of the TATA factors, which alter the TACA sequence or represent a combination of both types. The ability of these mutants to drive the expression of a reporter gene was then assessed by transient transfection of the constructs into Ishikawa and NCI-H441 cells, in which the modulatory uteroglobin promoter elements are active (52), and into CV-1 cells. The latter
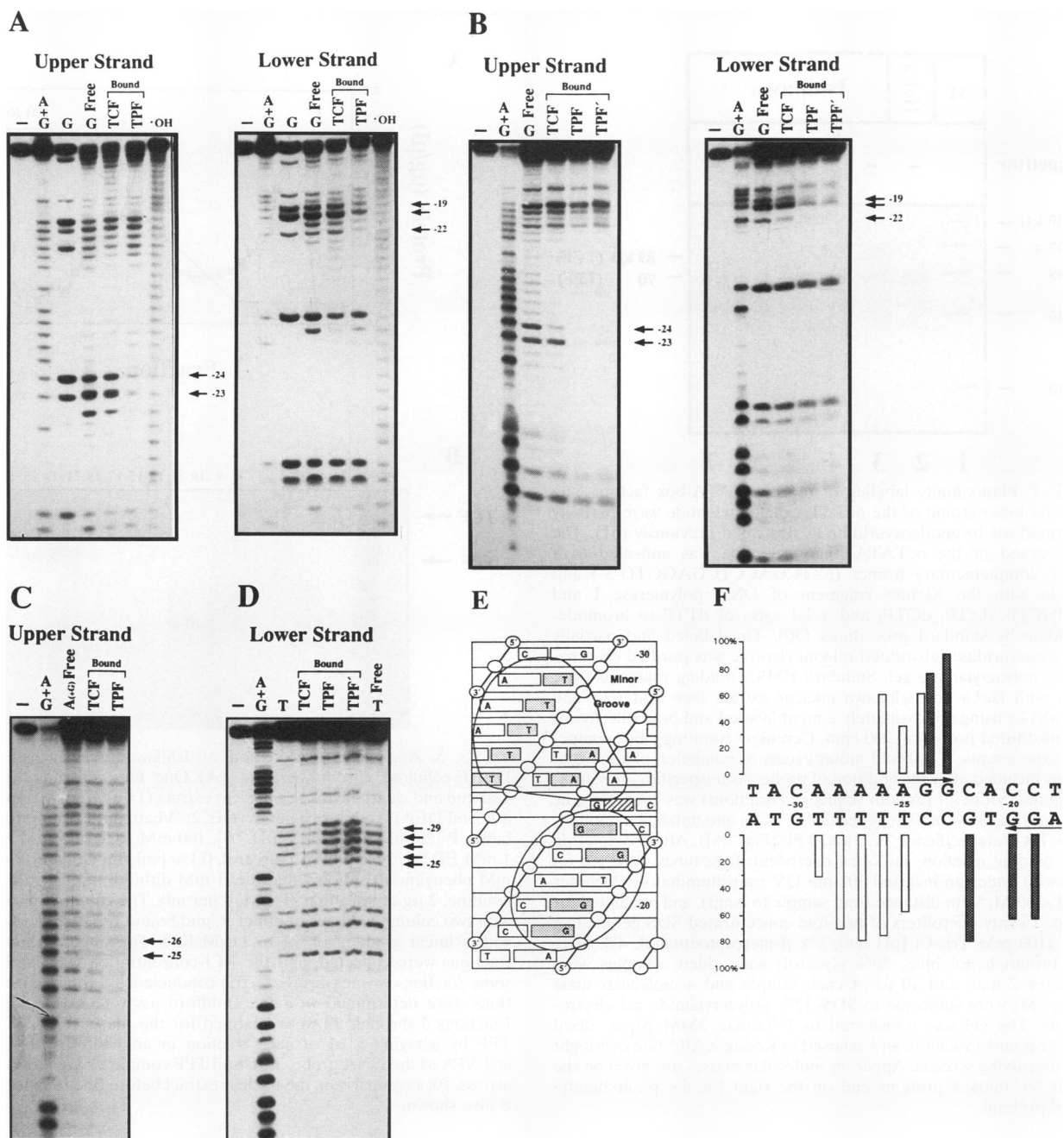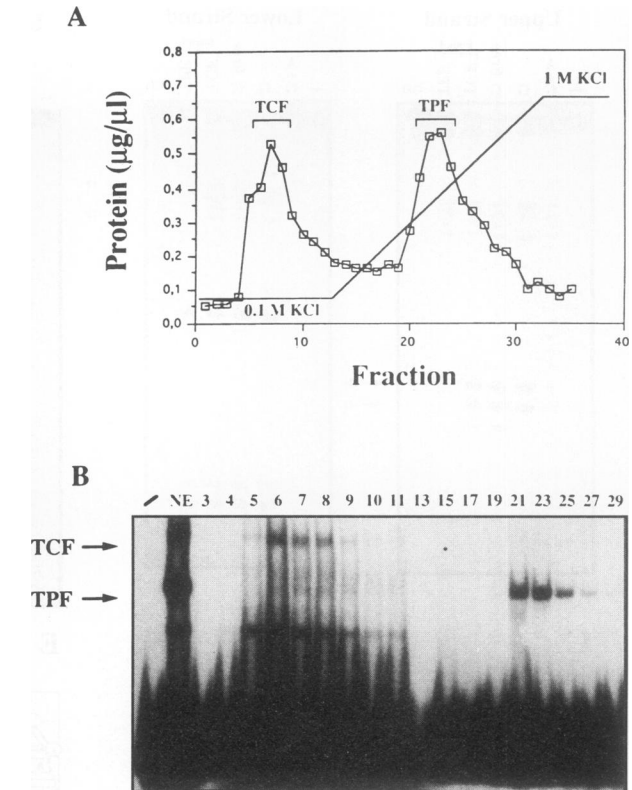
FIG. 3. Probing DNA-protein interactions in the rabbit uteroglobin TATA box region by using Ishikawa nuclear extract. (A) Methylation interference. Complexed (TCF and TPF) as well as free (G Free) oligonucleotides were analyzed on 15% sequencing gels together with appropriate controls. On all gels, the unmodified and uncleaved oligonucleotide (−) and a G+A reaction performed with naked DNA (36) were routinely included. For the methylation interference experiment, hydroxyl radical reactions (· OH [14]) were additionally performed. The guanines which do interfere with the binding of TPF are indicated with arrows. Numbers refer to the position relative to the cap site. (B) Methylation protection. Free oligonucleotide and DNA recovered from complex bands were cleaved with piperidine (strong G reaction). Protected guanines are indicated by arrows. (C) Methylation protection (strong A-weak G reaction). Free DNA [A+(G) Free] and DNA from complex bands were treated with sodium hydroxide after acid depurination. Only the upper strand is shown. The two adenines weakly protected in the TCF complex are indicated by arrows. (D) KMnO$_4$ interference. The thymines of the ocTATA oligonucleotide were modified with KMnO$_4$ prior to EMSA. The recovered oligonucleotides were treated as described for panel B. A naked DNA control was also included (T). The five thymines which interfere with the binding of TCF are indicated by arrows. (E) Sketch summarizing all contacts made by TCF and TPF. The DNA helix is depicted as a two-dimensional cylindrical projection with 10 bp per turn instead of 10.4 bp for the sake of simplicity. Phosphates are represented by open circles. The position of CG base pairs at −20 and −30 is indicated. All contacts made by TCF (shaded light grey, top) and by TPF (dark grey, bottom) are encircled. The guanine at −24, which is shared by both factors, is hatched. (F) The extent of protection of bases against modification or of interference with binding to modified bases is shown. The magnitude of the effect is represented by open (TCF) and shaded (TPF) bars. 0% indicates no protection or interference. For the guanines, where protection and interference experiments were performed, only the protection results are shown. The intensity of bands on appropriate autoradiography exposures was measured with a Desaga Quick Scan densitometer. Only effects larger than 30% were considered significant. Numbers along the DNA sequence refer to the position relative to the transcription start site. The half sites of the palindrome are indicated by arrows.

6213
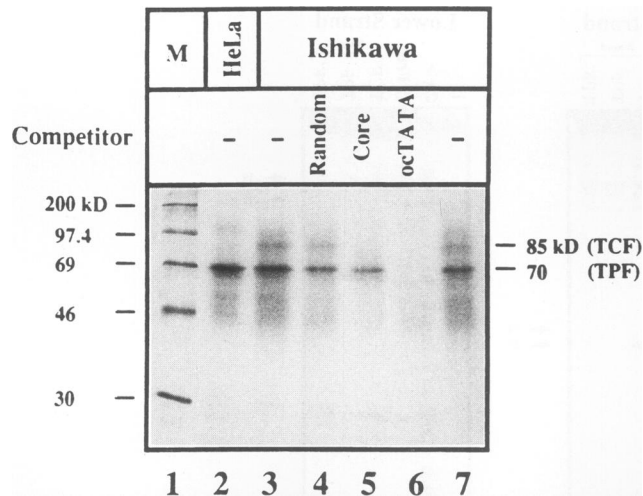
FIG. 4. Photoaffinity labeling of the two TATA-box factors. The T's in the lower strand of the ocTATA oligonucleotide were partially substituted for bromodeoxyuridine as described previously (61). The upper strand of the ocTATA oligonucleotide was annealed to a 14-base complementary primer (5'-TCGACCTCGAGGTG-3') and filled in with the Klenow fragment of DNA polymerase I and [α-32P]dCTP, dATP, dGTP, and a 1:1 ratio of dTTP to bromodeoxyuridine by standard procedures (30). The labeled and partially bromodeoxyuridine-substituted oligonucleotide was purified on a native 6% polyacrylamide gel. Standard EMSA binding reactions were set up with HeLa and Ishikawa nuclear extract (see Materials and Methods) by using approximately 2 ng of labeled and bromodeoxyuridine-substituted probe (50,000 cpm, Cerenkov counting). For competition experiments, a 100-fold molar excess of unlabeled oligonucleotide was included after the addition of probe. As unspecific competitor, an oligonucleotide of random sequence (Random) was used. Specific competitors used were Core (specific for TCF) and unlabeled, unmodified ocTATA (specific for TCF and TPF [Fig. 2A]). After completion of the binding reaction, the open microcentrifuge tubes were put on ice, placed under an inverted 302-nm UV transilluminator (Bachofer type IL-200-M; 5-cm distance from sample to lamp), and irradiated for 10 min. Twenty microliters of two-fold concentrated SDS gel-loading buffer (100 mM Tris-Cl [pH 6.8], 5% β-mercaptoethanol, 4% SDS, 0.1% bromophenol blue, 20% glycerol) was added, samples were boiled for 2 min, and 10 µl of each sample and a molecular mass marker (M) were subjected to SDS–12% polyacrylamide gel electrophoresis. The gel was transferred to Whatman 3MM paper, dried under heat and a vacuum, and exposed to Kodak XAR5 film overnight with intensifying screens. Apparent molecular masses are given on the left for the marker proteins and on the right for the photoaffinity-labeled proteins.

FIG. 5. Analytical fractionation of Ishikawa nuclear extract by DEAE-cellulose chromatography. (A) One hundred microliters of dialyzed and cleared Ishikawa nuclear extract (14 µg/µl) was applied to a 1.1-ml DEAE-cellulose column (DE52; Whatman) equilibrated with buffer P (20 mM K-HEPES [pH 7.9], 100 mM KCl, 12.5 mM MgCl₂, 1 mM EDTA, 20% [vol/vol] glycerol, 0.1% [vol/vol] Nonidet P-40, 0.5 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol, 1 mM benzamidine, 2 µg of aprotinin [Trasylol] per ml). The column was washed with two column volumes of buffer P, and bound proteins were eluted with a linear gradient of 0.1 to 1.0 M KCl. Two-hundred microliter fractions were collected, and the KCl concentration was checked in some fractions by measuring electric conductance. Protein concentrations were determined with the Bradford assay (Bio-Rad [6]). (B) Fractions 3 through 29 were analyzed for the presence of TCF and TPF by assaying 5 µl of each fraction in an EMSA with labeled ocTATA as the DNA probe. TCF and TPF complexes are indicated by arrows. As a comparison, the nuclear extract before fractionation (NE) is also shown.

do not recognize these elements and contain only TPF when assayed in an EMSA (25a). A panel of seven mutants was used (Fig. 6A). Pal1 was designed to destroy the guanine contact at −23 in the upper strand and the guanine contact at −19 in the lower strand required by TPF for binding (described above). Pal2 affects the 2-bp spacer sequence of the palindromic TPF binding site, thereby destroying the guanine contact at −22 TPF in the lower strand. The A/C mutant interrupts the oligo(dA) stretch immediately downstream of the TACA box and thus destroys one T contact required by TCF for binding (see above). The C/T mutant produces a canonical TATA box, but does not affect any nucleotide crucial for the binding of TCF or TPF. In the 4A mutant, one A of the oligo(dA) stretch is deleted, thereby reducing the number of available T contacts for TCF in the lower strand to four. The mutant C/T-Pal1 is a double mutant containing the mutations of the C/T and Pal1

constructs. In the mutant GGCA, the TATA-box homology is virtually destroyed by replacing TACA with GGCA, which leaves all mapped contacts for TCF and TPF intact. The latter mutant was designed to assess the participation of TBP in transcription from the uteroglobin promoter.

The wild-type promoter sequence in the uteroglobin promoter deletion CAT construct UG(−301)CAT (32) was exchanged for the mutant sequences, and their ability to drive the expression of the CAT reporter in Ishikawa, NCI-H441, and CV-1 cells was tested in gene transfer experiments (Fig. 6B). In CV-1 cells, none of the mutations had a deleterious effect on reporter gene activity (Fig. 6B, bottom). Only the GGCA mutant with its destroyed TATA box shows a moderate 2.5-fold reduced CAT activity. The C/T mutant on the other side, which contains a canonical TATA box, shows a sevenfold increased activity, whereas the C/T-Pal1 double mutant in-

A

WT	5´- ATACAAAAA̅G̅G̅CACCTG-3´

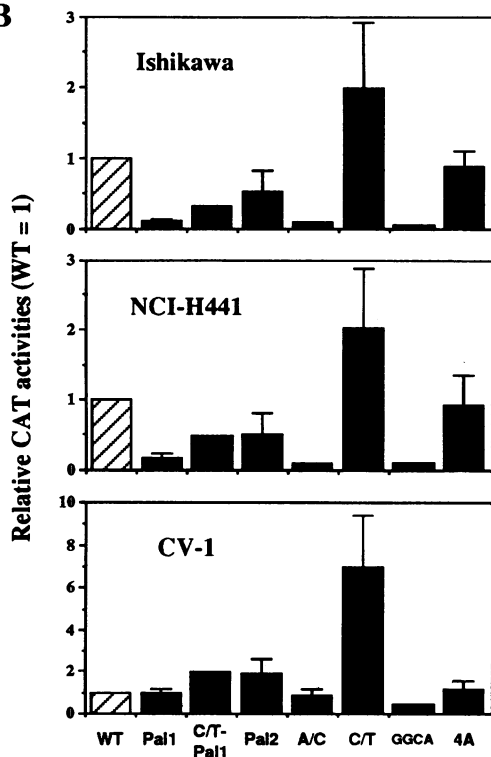| Pal 1 | 5´- ATACAAAAAGTCACATG-3´ |
C/T-Pal 1	5´- ATATAAAAAGTCACATG-3´
Pal 2	5´- ATACAAAAAGGATCCTG-3´
A/C	5´- ATACAACAAGGCACCTG-3´
C/T	5´- ATATAAAAAGGCACCTG-3´
GGCA	5´- AGGCAAAAAGGCACCTG-3´
4A	5´- ATACAA-AAGGCACCTG-3´

B

FIG. 6. Analysis of TATA-box mutants in gene transfer experiments with Ishikawa, NCI-H441, and CV-1 cells. (A) The mutations were introduced into the wild-type (WT) sequence in a construct containing 301 bp of the rabbit uteroglobin promoter followed by the CAT reporter gene and the simian virus 40 enhancer. Utilization of the correct transcription start site was shown by RNase mapping (32). The sequences of the TATA-box region are shown. Point mutations are underlined, and the half sites of the palindrome are marked with arrows in the wild-type sequence. (B) The set of mutants was transfected into Ishikawa, NCI-H441, and CV-1 cells. All transfections were normalized for plate-to-plate variations by cotransfecting and coassaying the Rous sarcoma virus-luciferase construct pRSVL (12). The ratios of the normalized CAT activities of each mutant construct to the wild-type construct (with standard deviation) are shown. The construct used is indicated below each bar of the CV-1 panel. The CAT activities (conversion) for the wild-type construct in the different cell lines were as follows: Ishikawa, 20 to 40%; NCI-H441, 10 to 30%; and CV-1, 5 to 10%. For each construct and cell line, at least two independent values were used to generate the chart. Error bars are based on at least two independent transfection series.
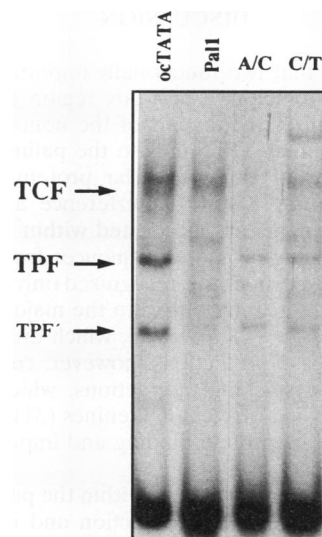

FIG. 7. EMSA of TATA-box mutants. ocTATA oligonucleotide mutants containing the point mutations of the Pal2, A/C, and C/T reporter constructs as shown in Fig. 6A were chemically synthesized and used in EMSA experiments with Ishikawa nuclear extract. TCF, TPF, and TPF′ complexes are indicated by arrows.

creases the activity only by a factor of 2. The Pal2 mutant had a moderate effect and stimulates transcription also by a factor of 2. In Ishikawa and NCI-H441 cells, the mutations in the Pal1, A/C, and GGCA constructs are strong down-mutations (approximately 10-fold). Introduction of a canonical TATA box in a mutant background with strong phenotype (C/T-Pal1) results in a twofold increase in reporter gene activity compared with the cognate single mutant Pal1. The phenotype of Pal2 is intermediate, resulting in twofold reduced activity, compared with twofold increased activity in CV-1 cells. Whereas the C/T mutant leads to a sevenfold increased activity in CV-1 cells, the effect is only a moderate twofold in Ishikawa and NCI-H441 cells. The 4A construct had no effect in all cell lines tested.

To answer the question of whether the functional phenotypes of the mutants Pal1, A/C, and C/T do correlate with in vitro binding data, the corresponding mutations were introduced into the ocTATA oligonucleotide and binding of TATA factors was examined (Fig. 7). As expected from the high-resolution mapping of protein-DNA contacts and EMSA experiments, the Pal1 mutation abolished binding of the ubiquitous factor TPF. Similarly, the A/C mutation severely inhibited the binding of TCF, whereas the C/T mutation did not lead to any major alteration, except for a stronger complex with low mobility, which is only spurious in some but not all shifts with ocTATA (Fig. 2B). In general, the complexes generated with the mutant oligonucleotides were weaker than the wild-type ocTATA oligonucleotide, which could simply be due to different specific activities of the labeled probes (ocTATA was subcloned, mutants were chemically synthesized). The previously mentioned TPF′ complex was also present in the Ishikawa extract used for this experiment. Like in the methylation protection experiment (described above), TPF′ behaves as TPF. Both complexes are undetectable with the Pal1 mutant, and both complexes are present when shifts were performed with the A/C mutant oligonucleotide, thereby strengthening the hypothesis that TPF′ is a degradation product of TPF with full DNA binding potential.

DISCUSSION

We show here that two functionally important factors bind to the rabbit uteroglobin TATA-box region at two adjacent sites. TCF recognizes the 3' part of the noncanonical TATA box (TACA), whereas TPF binds to the palindrome immediately downstream of this site. The protein-DNA contacts mapped for TCF by $KMnO_4$ interference and methylation protection experiments are all located within the Core oligonucleotide containing promoter sequences from −33 to −24, which in EMSA experiments is recognized only by TCF but not by TPF. All of these contacts map to the major groove except those of adenines at −25 and −26, which are located in the minor groove. These protections, however, could also reflect alteration in base stacking interactions, which is known to occur occasionally in a stretch of adenines (31). The alteration might be induced by protein binding and impede accessibility of the minor groove for DMS.

For TPF, only guanine contacts within the palindrome could be identified by methylation protection and interference experiments. Therefore, we conclude that TPF contacts the major groove. The palindrome was recognized as important for the binding of TPF from the EMSA experiments, although the palindromic sequence alone is not sufficient for producing a TPF complex. Thus, we currently do not know what kind of other contacts could be important for TPF binding.

The guanine contact at −24 is interesting, because it can be seen only in methylation protection but not in methylation interference experiments. This was also found, for example, for the guanine in position 4 (lower strand) of glucocorticoid-responsive elements, which methylation interferes with receptor binding (57) and which is hypermethylated by DMS after receptor binding (25). Similarly, the adenine opposite the first thymine in a glucocorticoid-responsive element is not contacted by the progesterone receptor according to missing contact probing experiments (8), although it has been reported that modification of this base interferes with the binding of glucocorticoid receptor (55).

Another interesting finding is that the same guanine (−24) is shared by TCF and TPF. A phosphate shared by RNA polymerase and λ repressor forms part of the evidence for the model that λ repressor activates transcription of the phage λcI gene by binding to the O_R2 operator site and directly contacting RNA polymerase (22). Similarly, the shared guanine suggests that TCF and TPF might interact to form a complex with DNA. A representation of the bases contacted by both factors (Fig. 3E) suggests that they could have a heterodimerization interface when bound to DNA. In fact, in some of the EMSA experiments, another specific complex with lower mobility than TCF can be observed (Fig. 2B, lane 10), which could represent a ternary complex. On the other hand, DNA binding of TCF and TPF could be independent of each other despite the shared G contact.

By linker-scanning analysis, it has been shown previously that the rabbit uteroglobin TATA box is one of the elements that must be kept intact for efficient expression of a reporter gene cloned behind the promoter (52). This is not an unexpected finding because in all linker-scanning analyses published so far, the TATA box had a strong phenotype. However, in the uteroglobin promoter, the TATA box is not canonical (TACA) and one would expect it to be a weak binding site for TBP. By comparing the activities of reporter gene constructs containing the uteroglobin promoter with a canonical TATA box (C/T mutant) and a destroyed TATA box (GGCA mutant) with that of the wild-type setting (TACA) in CV-1 cells, we show that this is indeed the case. CV-1 cells do not efficiently transcribe the uteroglobin promoter, and its activity should be determined by the affinity of TBP to the uteroglobin TATA box. As expected the canonical TATA box works best (Fig. 6B [C/T mutant]), the TACA sequence (wild-type construct) shows a 7-fold decreased activity (relative to C/T), and the GGCA mutant shows an 18-fold decreased activity. A very similar result was also shown for the TACA box of the adenovirus EIIa late promoter (Fig. 1), which binds TFIID with a 10-fold reduced activity compared with the canonical TATA mutant. Also, its ability to mediate activation by modulatory transcription factors is reduced by a factor of 10 (24).

The finding that those point mutations affecting TCF or TPF binding in vitro (A/C and Pal1) lead to 10-fold reduced activity in transient transfection experiments with Ishikawa and NCI-H441 cells correlates with the phenotype of three independent linker-scanner mutants with mutated binding sites for either TCF, TPF, or both (52). Moreover, because the C/T mutation does not affect the function of the uteroglobin promoter in Ishikawa and NCI-H441 cells and does not reduce binding of TCF, we conclude that the TACA sequence has no positive function in these cells. Rather, it is likely that the TACA sequence serves to reduce expression of the uteroglobin promoter in all cells lacking TCF because of its low affinity for TBP. In this context, it is interesting to note that only the uteroglobin promoter in rabbits contains a TACA box (Fig. 1) and that high levels of endometrial expression are only observed in the preimplantation phase of pregnant rabbits and not in other species like mice, rats, and humans (43). The finding that, in contrast to the results with Ishikawa and NCI-H441 cells, the two constructs Pal1 and A/C have no phenotype in CV-1 cells, which lack TCF, also speaks in favor of this hypothesis for TCF function. These results seem to imply that TPF is not required for transcription in these cells. However, since the activity of C/T-Pal1 is increased only twofold, we cannot exclude the possibility that TPF contributes to some extent to the activity of the uteroglobin promoter in this cell line.

The GGCA mutant has a strong phenotype in all cell lines tested. Since this mutation most likely inhibits binding of TBP, the results suggest that transcription in Ishikawa and NCI-H441 cells is also dependent on TBP binding. The main difference between these cell lines and CV-1 cells is the high level of activity of the wild-type construct with its TACA sequence, which is a weak TBP binding site. This high level of activity is drastically reduced by the Pal1 mutation, which blocks TPF binding. Since the activity of the Pal1 mutant cannot be fully restored to the wild-type level by the C/T-Pal1 double mutant, we propose that TPF, possibly in conjunction with TCF, facilitates TBP interaction with the weak TACA element on the uteroglobin promoter. Also, Swaminathan et al. (54) presented data on the use of two nonconsensus TATA boxes in the adenovirus E2 early promoter that are consistent with the idea that mammalian cells may possess one or more factors that facilitate efficient transcription of promoters with nonconsensus TATA motifs.

Since TCF does bind directly to the uteroglobin TATA-box core and TPF does bind directly to a palindrome immediately downstream of this site, the question arises of whether these factors are related to the well-known TATA-box binding protein TBP or other TATA-recognizing factors. There is general agreement that all TBP in nuclear extracts is assembled into general transcription complexes in higher eucaryotes (41). Therefore, binding activity of free TBP in an unpurified extract as measured by EMSA or DNase I footprinting experiments has never been reported. Native TBP-containing com-

plexes like SL1, TFIID, and TFIIIB, on the other hand, have low mobilities in EMSA gels because of their high molecular mass. In an EMSA experiment with a DNA probe containing a canonical TATA box, one would, therefore, expect no specific shift, as was the case with the human HSP70 TATA-box oligonucleotide. Furthermore, bacterially expressed and purified TBP did bind to this consensus TATA box in an EMSA experiment but did not bind to the rabbit uteroglobin TATA box. The mobility of the TBP-human HSP70 TATA complex was also different from those of the TCF and TPF complexes (not shown). Moreover, the apparent molecular masses of TCF and TPF, as deduced from the UV cross-linking experiment, do not fit the molecular mass of human TBP (37.5 kDa), even when taking into account the maximum molecular mass of the cross-linked DNA strand (13.5 kDa). It is also well documented that TBP binds exclusively to the minor groove of the DNA (26, 51) which is not the case for either TCF or TPF (see above).

Two other proteins that bind to a canonical TATA box have been cloned. The tissue-specific and developmentally regulated TBP-related factor TRF from *Drosophila melanogaster* (11) has a high level of homology to TBP but cannot substitute for it in a *Drosophila* in vitro transcription system containing general transcription factors. This also holds true for the human TATA-element modulatory factor TMF, which was cloned recently by recognition site screening of a HeLa cell expression cDNA library with a human immunodeficiency virus TATA-box oligonucleotide (16). In contrast to TRF, TMF is not homologous to TBP. Like in the case of TBP, the reported molecular masses for TRF (27 kDa) and TMF (123 kDa) seem to exclude any relationship to TCF or TPF.

A search of known DNA binding sequences of transcription factors for the presence of the palindrome AGGnnCCT recognized by TPF revealed a ubiquitous factor, NTF1, that interacts with the N element of an immunoglobulin heavy-chain gene promoter, which is in close proximity to an octamer element (62). Since the sequence of the N element (AGGaaC CTccccc) did contain the palindrome followed by a pyrimidine stretch and the methylation of all guanines in this sequence also interfered with binding, we tested whether NTF1 is identical to TCF or TPF. An oligonucleotide containing the N element but not the adjacent octamer motif, having the same ends and same length as ocTATA, was used in an EMSA experiment with Ishikawa nuclear extract (not shown). The observed NTF1 complex had much lower mobility than TCF, thereby making any relationship of NTF1 to TCF or TPF unlikely.

In conclusion, the TATA-box region of the rabbit uteroglobin gene exemplifies the possibility that, in addition to its usual function in binding basal transcription factors, this promoter element is recognized by specific transcription factors binding to the major groove of DNA. Thus, it will be important to elucidate the architecture of the transcription preinitiation complex of this promoter.

ACKNOWLEDGMENTS

REFERENCES

1. **Andrews, N. C., and D. V. Faller.** 1991. Rapid micropreparation technique for extraction of DNA binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res. **19:**2499.
2. **Aumüller, G., J. Seitz, W. Heyns, and C. Kirchner.** 1985. Ultrastructural localization of uteroglobin immunoreactivity in rabbit lung and endometrium and rat ventral prostate. Histochemistry **83:**413–417.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1993. Current protocols in molecular biology. Greene Publishing Associates and John Wiley & Sons, Inc., New York.
4. **Beaucage, S. L., and M. H. Caruthers.** 1981. Deoxynucleoside phoshoramidites—a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett. **22:**1859–1862.
5. **Beier, H. M.** 1968. Uteroglobin: a hormone-sensitive endometrial protein involved in blastocyst development. Biochim. Biophys. Acta **160:**289–291.
6. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
7. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50:**349–383.
8. **Cairns, C., J.-Å. Gustafsson, and J. Carlstedt-Duke.** 1991. Identification of protein contact sites within the glucocorticoid/progestin response element. Mol. Endocrinol. **5:**598–604.
9. **Cato, A. C. B., R. Miksicek, G. Schütz, J. Arnemann, and M. Beato.** 1986. The hormone regulatory element of mouse mammary tumour virus mediates progesterone induction. EMBO J. **5:**2237–2240.
10. **Conaway, R. C., and J. W. Conaway.** 1993. General initiation factors for RNA polymerase II. Annu. Rev. Biochem. **62:**161–190.
11. **Crowley, T. E., T. Hoey, J.-K. Liu, Y. N. Jan, L. Y. Jan, and R. Tjian.** 1993. A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila*. Nature (London) **361:**557–561.
12. **de Wet, J. R., K. V. Wood, M. DeLuca, D. R. Helinski, and S. Subramani.** 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. **7:**725–737.
13. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
14. **Dixon, W. J., J. J. Hayes, J. R. Levin, M. F. Weidner, B. A. Dombroski, and T. Tullius.** 1991. Hydroxyl radical footprinting. Methods Enzymol. **208:**380–413.
15. **Fernandez-Renau, D., M. Lombardero, and A. Nieto.** 1984. Glucocorticoid-dependent uteroglobin synthesis and uteroglobin mRNA levels in rabbit lung explants cultured in vitro. Eur. J. Biochem. **144:**523–527.
16. **Garcia, J. A., S. H. Ou, F. Wu, A. J. Lusis, R. S. Sparkes, and R. B. Gaynor.** 1992. Cloning and chromosomal mapping of a human immunodeficiency virus 1 "TATA" element modulatory factor. Proc. Natl. Acad. Sci. USA **89:**9372–9376.
17. **Gazdar, A. F., R. I. Linnoila, Y. Kurita, H. K. Oie, J. L. Mulshine, J. C. Clark, and J. A. Whitsett.** 1990. Peripheral airway cell differentiation in human lung cancer cell lines. Cancer Res. **50:**5481–5487.
18. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
19. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–487.
20. **Hagen, G., S. Müller, M. Beato, and G. Suske.** 1992. Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes. Nucleic Acids Res. **20:**5519–5525.
21. **Hagen, G., M. Wolf, S. L. Katyal, G. Singh, M. Beato, and G. Suske.** 1990. Tissue-specific expression, hormonal regulation and 5'-flanking gene region of the rat Clara cell 10 kDa protein: comparison to rabbit uteroglobin. Nucleic Acids Res. **18:**2939–2946.
22. **Hochschild, A., N. Irwin, and M. Ptashne.** 1983. Repressor structure and the mechanism of positive control. Cell **32:**319–325.
23. **Howgate, P., A. S. Jones, and J. R. Tittensor.** 1968. The perman-

ganate oxidation of thymidine. J. Chem. Soc. (C):275–279.

24. Huang, D.-H., M. Horikoshi, and R. G. Roeder. 1988. Activation of the adenovirus EIIa late promoter by a single-point mutation which enhances binding of transcription factor IID. J. Biol. Chem. 263:12596–12601.

25. Karin, M., A. Haslinger, A. Holtgreve, R. I. Richards, P. Krauter, H. M. Westphal, and M. Beato. 1984. Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein IIA. Nature (London) 308:513–519.

25a.Klug, J. Unpublished observations.

26. Lee, D. K., M. Horikoshi, and R. G. Roeder. 1991. Interaction of TFIID in the minor groove of the TATA element. Cell 67:1241–1250.

27. Lopez de Haro, M. S., and A. Nieto. 1983. Isolation and characterization of uteroglobin from the lung of the hare (Lepus capensis). Arch. Biochem. Biophys. 226:539–547.

28. Lopez de Haro, M. S., and A. Nieto. 1985. Glucocorticoids induce the expression of the uteroglobin gene in rabbit foetal lung explants cultured in vitro. Biochem. J. 225:255–258.

29. Maire, P., J. Wuarin, and U. Schibler. 1989. The role of cis-acting promoter elements in tissue-specific albumin gene expression. Science 244:343–346.

30. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

31. Maxam, A. M., and W. Gilbert. 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74:560–564.

32. Misseyanni, A., J. Klug, G. Suske, and M. Beato. 1991. Novel upstream elements and the TATA-box region mediate preferential transcription from the uteroglobin promoter in endometrial cells. Nucleic Acids Res. 19:2849–2859.

33. Mukherjee, A. B., L. C. Murty, and J. Y. Chou. 1993. Differentiation and uteroglobin gene expression by novel rabbit endometrial cell lines. Mol. Cell. Endocrinol. 94:R15–R22.

34. Müller, H., and M. Beato. 1980. RNA synthesis in rabbit endometrial nuclei. Hormonal regulation of transcription of the uteroglobin gene. Eur. J. Biochem. 112:235–241.

35. Nakajima, N., M. Horikoshi, and R. G. Roeder. 1988. Factors involved in specific transcription by mammalian RNA polymerase II: purification, genetic specificity, and TATA box-promoter interactions of TFIID. Mol. Cell. Biol. 8:4028–4040.

36. Negri, R., G. Costanzo, and E. D. Mauro. 1991. A rapid single-reaction method for DNA sequence determination. Anal. Biochem. 197:389–395.

37. Nishida, M., K. Kasahara, M. Kaneko, and H. Iwasaki. 1985. Establishment of a new human endometrial adenocarcinoma cell line, Ishikawa cells, containing estrogen and progesterone receptors. Acta Obst. Gynaecol. Jpn. 37:1103–1111.

38. Papavassiliou, A. G. 1993. Localisation of DNA-protein contact points by DMS resistance of complexes resolved in gel retardation assays. Nucleic Acids Res. 21:757–758.

39. Penotti, F. E. 1990. Human DNA TATA boxes and transcription initiation sites. A statistical study. J. Mol. Biol. 213:37–52.

40. Pothier, F., M. Ouellet, J.-P. Julien, and S. L. Guérin. 1992. An improved CAT assay for promoter analysis in either transgenic mice or tissue culture cells. DNA Cell Biol. 11:83–90.

41. Rigby, P. W. J. 1993. Three in one and one in three: it all depends on TBP. Cell 72:7–10.

42. Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350–1354.

43. Sandmöller, A., A. K. Voss, J. Hahn, B. Redemann-Fibi, G. Suske, and M. Beato. 1991. Cell-specific, developmentally and hormonally regulated expression of the rabbit uteroglobin transgene and the endogenous mouse uteroglobin gene in transgenic mice. Mech. Dev. 34:57–67.

44. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing

with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.

45. Sawaya, P. L., B. R. Stripp, J. A. Whitsett, and D. S. Luse. 1993. The lung-specific CC10 gene is regulated by transcription factors from the AP-1, octamer, and hepatocyte nuclear factor 3 families. Mol. Cell. Biol. 13:3860–3871.

46. Shapiro, D. J., P. A. Sharp, W. W. Wahli, and M. J. Keller. 1988. A high-efficiency HeLa cell nuclear transcription extract. DNA 7:47–55.

47. Singer, V. L., C. R. Wobbe, and K. Struhl. 1992. A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. Genes Dev. 4:636–645.

48. Singh, G., J. Singh, S. L. Katyal, W. E. Brown, J. A. Kramps, I. L. Paradis, J. H. Dauber, T. A. Macpherson, and N. Squeglia. 1988. Identification, cellular localization, isolation, and characterization of human Clara cell-specific 10kD protein. J. Histochem. Cytochem. 36:73–80.

49. Slater, E. P., G. Redeuihl, K. Theis, G. Suske, and M. Beato. 1990. The uteroglobin promoter contains a noncanonical estrogen responsive element. Mol. Endocrinol. 4:604–610.

50. Smale, S. T., and D. Baltimore. 1989. The "initiator" as a transcription control element. Cell 57:103–113.

51. Starr, D. B., and D. K. Hawley. 1991. TFIID binds in the minor groove of the TATA box. Cell 67:1231–1240.

52. Suske, G., W. Lorenz, J. Klug, A. F. Gazdar, and M. Beato. 1992. Elements of the rabbit uteroglobin promoter mediating its transcription in epithelial cells from the endometrium and lung. Gene Expr. 2:339–352.

53. Suske, G., M. Wenz, A. C. B. Cato, and M. Beato. 1983. The uteroglobin gene region: hormonal regulation, repetitive elements and complete nucleotide sequence of the gene. Nucleic Acids Res. 11:2257–2271.

54. Swaminathan, S., P. Malhotra, C. F. Manohar, R. Dhar, and B. Thimmapaya. 1993. Activation of a dual adenovirus promoter containing nonconsensus TATA motifs in Schizosaccharomyces pombe: role of TATA sequences in the efficiency of transcription. Nucleic Acids Res. 21:2737–2746.

55. Truss, M., and M. Beato. 1993. Steroid hormone receptors: interaction with DNA and transcription factors. Endocr. Rev. 14:459–479.

56. Truss, M., G. Chalepakis, and M. Beato. 1990. Contacts between steroid hormone receptors and thymines in DNA: an interference method. Proc. Natl. Acad. Sci. USA 87:7180–7184.

57. Truss, M., G. Chalepakis, E. P. Slater, S. Mader, and M. Beato. 1991. Functional interaction of hybrid response elements with wild-type and mutant steroid hormone receptors. Mol. Cell. Biol. 11:3247–3258.

58. Warembourg, M., O. Tranchant, M. Atger, and E. Milgrom. 1986. Uteroglobin messenger ribonucleic acid: localization in rabbit uterus and lung by in situ hybridization. Endocrinology 119:1632–1640.

59. Wiley, S., R. J. Kraus, and J. E. Mertz. 1993. Functional binding of the "TATA" box binding component of transcription factor TFIID to the −30 region of TATA-less promoters. Proc. Natl. Acad. Sci. USA 89:5814–5818.

60. Wolf, M., J. Klug, R. Hackenberg, M. Gessler, K. H. Grzeschik, M. Beato, and G. Suske. 1992. Human CC10, the homologue of rabbit uteroglobin: genomic cloning, chromosomal localization and expression in endometrial cells. Hum. Mol. Genet. 1:371–378.

61. Wu, C., S. Wilson, B. Walker, I. Dawid, T. Paisley, V. Zimarino, and H. Ueda. 1987. Purification and properties of Drosophila heat shock activator protein. Science 238:1247–1253.

62. Yoza, B. K., and R. G. Roeder. 1990. Identification of a novel factor that interacts with an immunoglobulin heavy-chain promoter and stimulates transcription in conjunction with the lymphoid cell-specific factor OTF2. Mol. Cell. Biol. 10:2145–2153.

63. Zawel, L., and D. Reinberg. 1993. Initiation of transcription by RNA polymerase II: a multi-step process. Prog. Nucleic Acids Res. Mol. Biol. 44:67–108.