Published online December 8, 2004

6454–6467 Nucleic Acids Research, 2004, Vol. 32, No. 21
doi:10.1093/nar/gkh981

A novel E2 box-GATA element modulates Cdc6 transcription during human cells polyploidization

Nuria Vilaboa, Rodrigo Bermejo, Pilar Martinez, Rafael Bornstein[1] and Carmela Calés*

Department of Biochemistry, Instituto de Investigaciones Biomédicas 'A. Sols', Universidad Autónoma-CSIC, Arturo Duperier, 4, 28029 Madrid, Spain and [1]Madrid Cord Blood Bank, Hospital Universitario 'Doce de Octubre', Km. 5.4, 28041 Madrid, Spain

Received August 21, 2004; Revised and Accepted November 16, 2004

ABSTRACT

Cdc6 is a key regulator of the strict alternation of S and M phases during the mitotic cell cycle. In mammalian and plant cells that physiologically become polyploid, *cdc6* is transcriptionally and post-translationally regulated. We have recently reported that Cdc6 levels are maintained in megakaryoblastic HEL cells, but severely downregulated by ectopic expression of transcriptional repressor *Drosophila melanogaster escargot*. Here, we show that *cdc6* promoter activity is upregulated during megakaryocytic differentiation of HEL endoreplicating cells, and that *Escargot* interferes with such activation. Transactivation experiments showed that a 1.7 kb region located at 2800 upstream *cdc6* transcription initiation site behaved as a potent enhancer in endoreplicating cells only. This activity was mainly dependent on a novel *cis*-regulatory element composed by an E2 box overlapping a GATA motif. Ectopic *Escargot* could bind this regulatory element *in vitro* and endogenous GATA-1 and E2A formed specific complexes in megakaryoblastic cells as well as in primary megakaryocytes. Chromatin Immunoprecipitation analysis revealed that both transcription factors were occupying the E2 box/GATA site *in vivo*. Altogether, these data suggest that cdc6 expression could be actively maintained during megakaryocytic differentiation through transcriptional mechanisms involving specific *cis*- and *trans*-regulatory elements.

INTRODUCTION

Several terminally differentiated cell lineages, from plants to mammals, repeatedly enter S phase without completing mitosis in order to become polyploid (1,2). Among them, platelet precursors, megakaryocytes, undergo endoreplication in a process traditionally known as endomitosis. Megakaryocytic polyploidization is believed to be essential for full platelet production (3). As in other endoreplicating cells, megakaryocytic endomitosis is accompanied by a specific control of G_1/S transition cell cycle machinery, which enables the establishment of such truncated cell cycles (4–7). We have also recently shown that transcriptional repressor *Snail* (*Sna*) and *Escargot* (*Esg*) interfere with the regulation of the G_1/S transition of endomitotic cycles and therefore with polyploidization of megakaryoblastic HEL cells (8). These proteins also inhibit the entrance into endocycles of embryo and larval *Drosophila* cells (9), as well as of mouse trophoblast (10). However, little is known about the putative targets of these transcriptional repressors and how they are affecting the establishment of endomitosis.

We have recently described that Cdc6 is differentially regulated during megakaryocytic polyploidization (11). Cdc6 is one of the essential factors involved in assembling and licensing origins of replication during the G_1 phase of the mitotic cell cycle [for a review see (12)]. Its expression and ability to load MCM complex onto the origin of replication are tightly regulated in the majority of somatic cells to ensure that DNA is replicated only once per cell cycle (13). Interestingly, *cdc6* expression seems to be stabilized not only in megakaryoblastic cells that undergo endoreplication (11) but also in *Arabidopsis* endoreplicating cells (14,15). As also reported in plant cells (15), what we previously found was that the maintenance of *cdc6* expression during megakaryocytic differentiation could be achieved through post-translational stabilization in cyclin E overexpressing cells, and also through transcriptional regulation of *cdc6* gene (11).

Cdc6 expression peaks during G_1 and is transcriptionally regulated in an E2F-, cell cycle-related manner, not only in mammalian (16–18) but also in *Drosophila* (19) and plant cells, as is the case of *Arabidopsis* (14,20). In mammalian cells, this control is mediated through E2F-2 and E2F-3 binding to specific sites proximal to the gene transcription initiation site (21,22). In fact, the essential regulatory region is contained within ~300 bp of the 5′ flanking region, where the E2F-binding sites are located (16–18), and no

*To whom correspondence should be addressed. Tel: +34 91 5854469; Fax: +34 91 5854401; Email: ccales@iib.uam.es
Present address:
Nuria Vilaboa, Department of Experimental Surgery, Bone Metabolism Laboratory, Hospital 'La Paz', Madrid, Spain

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

Nucleic Acids Research, Vol. 32 No. 21 © Oxford University Press 2004; all rights reserved

further proliferation-mediated transcriptional activity is obtained by using additional upstream sequences up to −9000 bp (17).

In order to investigate the transcriptional regulation of *cdc6* during megakaryoblastic cells endoreplication, we took advantage of the fact that differentiated *Esg*-expressing HEL cells lose their ability to endoreplicate and to express Cdc6 (11). We have identified potential *Esg*-binding sites (E2 boxes) in *cdc6* upstream region that could be involved in maintaining Cdc6 expression during endoreplication and inhibited by the presence of *Esg* in HEL cells. We have found in the −4537 to −2800 region upstream *cdc6* transcription initiation site three E2 boxes, one of which overlaps a GATA element. This E2 box/GATA site is the most prominent binding site, not only for *Esg* but also for endogenous proteins. GATA-1 and E2A '*in vitro*' and '*in vivo*' occupy this site in endoreplicating cells. Transactivation data indicate that this particular sequence may act as an enhancer involved in transcriptional activation of *cdc6* promoter only when megakaryocytic differentiation is stimulated. The results presented here suggest that specific transcriptional mechanisms could be involved in the maintenance of Cdc6 expression that occurs during megakaryocytic endoreplication.

MATERIALS AND METHODS

Cell culture

Cells were cultured as described previously (8,11). To induce megakaryocytic or erythrocytic differentiation, 0.15–0.20 $\times 10^6$ cells per ml were grown for 48 h in the presence or absence of 10^{-8} M *o*-tetradecanoylphorbol 13 acetate (TPA; Sigma) or 5×10^{-5} M hemin (Sigma), respectively. For RNA stabilization studies, exponentially growing and TPA-treated cells were cultured for the indicated times in the presence of actinomycin D (5 μg/ml). For proliferation curves, exponentially growing cells were subcultured into 24-well plates (0.20×10^6 cells per ml and well) with or without TPA. The number of viable cells was determined by trypan blue exclusion in a hemocytometer chamber.

Megakaryocyte culture and $CD41^+$ cells immunoselection

Cord blood (CB) mononucleated cell fraction (MNC) was obtained after separation on Ficoll-Hypaque gradients and counted. A total of 5×10^6 cells were used for nuclear extract isolation. $CD34^+$ cells were isolated from MNC (typically, 5×10^7 cells) using an immunomagnetic procedure (MiniMACS, Miltenyi Biotec, Auburn, CA). CB-purified $CD34^+$ cells (5×10^5) were cultured for 12 days in 24-well culture dishes at 50 000 cells/ml in serum-free medium (CellGro, CellGenix, Freiburg, Germany) supplemented with recombinant human thrombopoietin [(TPO); non-pegylated; 50 ng/ml; generously provided by Kirin, Gunma, Japan].

$CD41^+$ cells generated in culture were purified using colloidal magnetic dextran iron particles that were selectively bound to target cells by means of anti-CD41 bi-specific tetrameric antibody complexes (StemSep, StemCell Technologies, Vancouver, BC, Canada). These complexes recognize both dextran and CD41 surface antigen. Labeled cells were retained in a column placed in a magnetic field and after washing, $CD41^+$ cells were eluted from the column, according to the manufacturer's instructions.

Plasmids and transfection assays

pHscdc6(4537), a reporter plasmid containing a 4.54 kb fragment of the upstream region of human *cdc6* gene inserted into pGL2Basic (Promega), was kindly provided by Dr Nakamura (Tokyo Medical and Dental University, Tokyo, Japan) (17). To construct pHscdc6(−2800), a 1.2 kb KpnI/XhoI fragment and 1.64 kb BamHI/KpnI fragment isolated from pHscdc6luc were inserted in between the BamHI and XhoI sites of pBluescript $SK^{+/-}$ (Stratagene). The resulting construct was digested with SacI/XhoI to isolate a 2.9 kb fragment that was ligated in between the SstI and XhoI sites of pGL2Basic. To construct pHscdc6(−1600), 1.6 kb HindIII fragment from pHscdc6Luc was inserted into the HindIII site of pGL2Basic. To construct pHscdc6(4537/−2800), a 1.7 kb KpnI/BamHI fragment from pHscdc6(4537) was inserted in between the KpnI/BglII of pGL2Promoter (Promega). Mutagenesis of the E2 boxes sites was performed in pHscdc6(4537/−2800) using the Quick Change Method (Quick Change kit, Stratagene Corp.) by employing the following oligonucleotides:

- To prepare pHscdc6(4537/−2800) A^M: 5′-AGCTGGGACCAAGCTTGTGTACCACT-3′; and 5′-AGTGGTACACAAGCTTGGTCCCAGCT-3′.
- To prepare pHscdc6(4537/−2800) B^M: 5′-GTATGAGTTAGGAGTTGATAAG-3′; and 5′-AGGTCCTTATCAACTCCTAACT-3′.
- To prepare pHscdc6(4537/−2800) $BGATA^M$: 5′-GTATGAGTTAGCAGGTGCACAG-3′; and 5′-AGGTCCTGTGCACCTGCTAACT-3′.
- To prepare pHscdc6(4537/−2800) C^M: 5′-GAGGCTGAGGGGATCCGATCACCTGA-3′; and 5′-TCAGGTGATCGGATCCCCTCAGCGTC-3′.

Cells were transfected by electroporation at 975 μF/250 V in a 0.4 cm cuvette. A total of 5×10^6 cells were transfected with 5 μg of the various plasmids and 5 μg of renilla plasmid (Promega). After 5 h, cells were washed with phosphate-buffered saline (PBS), counted and divided into two aliquots one of which was treated with TPA. At the end of 48 h following electroporation, cells were harvested and cell extracts for assaying luciferase activity were prepared using the Dual Luciferase kit (Promega), according to the manufacturer's instructions.

Northern blot

Total RNA was extracted using TRIzol® Reagent (Life Technologies), following the procedure described by the manufacturer. All other conditions, including the source and preparation of *cdc6*-specific cDNA probe and the entire pTRI RNA 28S plasmid, which hybridizes to 28S rRNA (Ambion, Inc., Austin, TX), were as described previously (11). The autoradiographs were scanned using a densitometer (GE-800 Calibrated Densitometer; Bio-Rad) to determine the relative optical densities of *cdc6* mRNA and 28S rRNA hybridization signals. For each RNA sample, the signal for the transcripts was measured within the linear range of the densitometer, and the ratio of *cdc6* mRNA signals to 28S rRNA was calculated.

Mobility gel shift analysis

EMSA were perfomed as described previously (8) with 0.02 pmol of ^{32}P-labelled oligoprobe and 20 µg of nuclear extracts. The partially complementary oligonucleotides are as follows: 5′-GCGGCCTGACAGGTGC-3′ and 5′-CAAAG-CACCTGTCAGGC-3′ (EBE probe); 5′-GCTGGGACCA-CAGGTGTGTAC-3′ and 5′-AGTGGTACACACCTGTG-GTCC-3′ (BoxA); 5′-GTATGAGTTAGCAGGTGATAAG3′ and 5′-AGGTCCTTATCACCTGCTAACT-3′ (BoxB); 5′-GAGGCTGAGGCAGGTGGATCAC-3′ and 5′- CTCAGGT-GATCCACCTGCCTCA-3′ (BoxC); 5′-GTATGAGTTAG-CAGGTGCACAG-3′ and 5′-AGGTCCTGTGCACCTGCT-AACT-3′ (BoxB GATA^{M}); 5′-GTATGAGTTAGGAGTTGA-TAAG-3′ and 5′-AGGTCCTTATCAACTCCTAACT-3′ (BoxB Ebox^{M}); 5′-AATTCGCAGGTGACAGGTGGCAGG-TGACAGGTGT-3′ and 5′-CTAGACACCTGTCACCTGC-CACCTGTCACCTGCG-3′ (E2 box tandem).

These were used to generate double-stranded oligoprobes. Binding reactions were carried out for 20 min at room temperature in binding buffer [20% glycerol, 1 mM DTT, 20 mM HEPES (pH 7.9), 5 mM $MgCl_2$, 100 mM KCl and 0.2 mM EDTA] containing 0.02 pmol of ^{32}P-labelled oligoprobe, 20 µg of protein extracts as indicated, and 40 µg/ml poly(dI-dC). When needed, competitor oligoprobes were added in excess to the binding reaction. For supershifts experiments, 2 µl of anti-GATA-1 (C-20 Santa Cruz), anti-GATA-2 (H-116 Santa Cruz), anti-E2A (V-18 Santa Cruz) antibodies were incubated with the nuclear extracts for 2 h at 4°C before the addition of the radiolabelled probe. DNA–protein complexes were separated from unbound labelled oligoprobes on 6% non-denaturing polyacrylamide gels in Tris-Borate/EDTA buffer. After drying, the shifted complexes were visualized using autoradiography.

Western blot

Nuclear extracts (30 µg) were subjected to SDS–PAGE and proteins transferred onto BioTrace PVDF membranes (Pall Corporation, Ann Arbor, MI) for 1 h at 2 mA/cm^2 on a semidry transfer apparatus (Amersham). Ponceau staining was routinely performed on membranes to check sample loading control. After blocking in PBS containing 0.1% Tween-20 (T-PBS) and 5% skimmed dry milk, filters were incubated overnight at 4°C with anti-human Cdc6 mouse mAb (Ab3; Oncogene, Darmstadt, Germany) at a 1:500 dilution in T-PBS. After washing and incubation with goat anti-mouse IgG conjugated to horseradish peroxidase (Dako, Glostrup, Denmark), signals were detected using the enhanced chemiluminescence system (Pierce, Rockford, IL).

Chromatin immunoprecipitation assays

The procedure was performed as described previously (23), with the following modifications. Nuclear extracts were sonicated with three cycles of 15 s each at 50% maximum power and 8 mA amplitude with a sonicator (Soniprep 150, MSE), and Protein-G Sepharose (Oncogene Research Products) was used instead of Protein-A Sepharose. Incubation with antibodies (anti-GATA-1, rat monoclonal N6 or goat polyclonal C-20, and anti-E2A, rabbit polyclonal V-18, all from Santa Cruz Biotechnology) or the corresponding non immune isotype immunoglobulins (Sigma) was carried out overnight at 4°C. The extracted DNA was resuspended in 30 µl water/10^7 cells. An aliquot of 5 µl and the equivalent to 1/1000 of the input fraction was used for each PCR (40 cycles of 30 s at 94°C; 30 s at 56°C; 30 s at 72°C and one last cycle of 10 min at 72°C), containing 25 pmol of each upper and lower primers. The following primers were used:

- For cdc6 E2 box/GATA site: 5′-ATTCTGTCTTCTTGT-GATGGTATTGC-3′; and 5′-GGTCGTTTAGTTCATGGT-TAGATTGG-3′.
- For cdc6 (−2000): 5′-TTGTATTTTAGTAGAGACGGG-GTTTC-3′; and 5′-GATGGTATTTCTGATGTCTTAG-TAGG-3′.
- For γ-globin site: 5′-AGAGAAAAACTGGAATGACT-GAATCG-3′; and 5′-AAGCGAGTGTGTGGAACTGCT-GAAGG-3′.

PCR products were separated in a 1.5% TBE-agarose gel and DNA was visualized by Ethidium Bromide staining.

Flow cytometry determinations

Cell surface immunofluorescence staining was performed by incubating cells at room temperature with the appropriate dilution in phosphate-buffered saline (PBS)/0.1% NaN_3/ 10% fetal calf serum (FCS) of anti-human GPIIIa (CD61) fluorescein isothiocyanate (FITC)-conjugated antibody (Becton Dickinson) for 20 min. To determine DNA content, cells were harvested by centrifugation at 500 *g* for 5 min, washed in ice-cold PBS and then resuspended in PBS containing 50 µg/ml propidium iodide (PI) and 0.1% Nonidet P-40. Determination of CD61 expression and cell cycle analysis were performed using a FACScan analyser using CELLQUEST software (both from Becton Dickinson).

RESULTS

Constitutive expression of Escargot determines downregulation of Cdc6 expression at RNA level

We have recently shown that cdc6 expression is stabilized during megakaryocytic polyploidization but downregulated in megakaryoblastic cells unable to establish endoreplication cycles, and among them, HA1 cell line, derived from HEL cells by constitutive expression of *D.melanogaster Esg* (11). We have also shown that ectopic expression of transcriptional repressor *Esg* in HEL cells inhibits megakaryocytic endomitosis (8). In this paper, HA1 cells have been designated as HEL(*Esg*+) in order to stress that it is the ectopic *Esg* expression that differentiates parental and derived cells. Interestingly, Cdc6 protein levels were profoundly diminished in HEL(*Esg*+) cells, compared to those in parental HEL cells, suggesting that *Esg* could be interfering with *cdc6* gene expression [(11) and Figure 1D]. In an independent search to find putative *Esg* targets that could be affecting endoreplication, we compared cDNAs of endoreplicating HEL and of HEL(*Esg*+) cells by DNA microarray hybridization. The analysis revealed that *cdc6* was one of the genes significantly downregulated in differentiated HEL(*Esg*+) (N. Vilaboa and C. Calés, unpublished data). To confirm these data, a northern-blot analysis was carried out to compare the levels of *cdc6* RNA in HEL and HEL(*Esg*+) after TPA treatment,

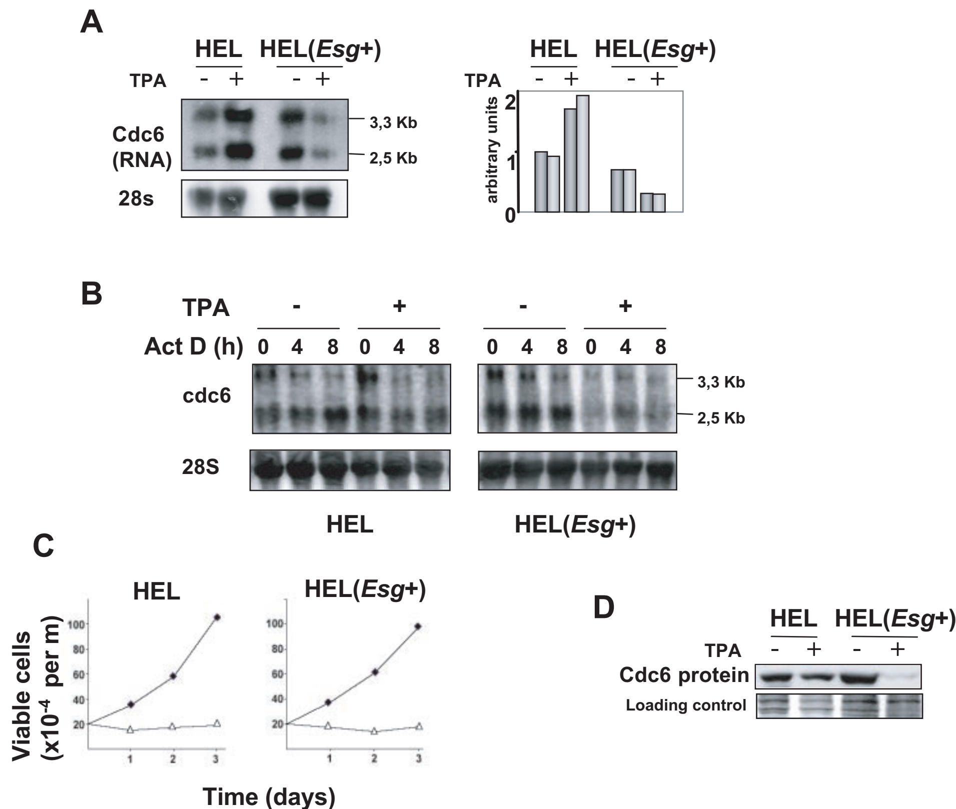

Figure 1. *cdc6* mRNA levels are downregulated in differentiated HEL(*Esg*+) cells. (**A**) HEL and HEL(*Esg*+) were cultured for 48 h in the absence (−) or in the presence (+) of 10^{-8} M TPA for 48 h. Total RNA was extracted and analyzed by northern blotting with full-length *cdc6* and 28S rRNA-specific cDNA probes. Quantification of *cdc6* mRNA signals normalized by 28S RNA is also shown (dark bars, 3.3 kb specie and light bars, 2.5 kb specie). (**B**) Cells were cultured as in (A), and then treated for the indicated times with 5 µg/ml Act D. RNA extraction and analysis was performed as in (A). (**C**) HEL and HEL(*Esg*+) cells were seeded in triplicate wells and cultured in the absence (closed circles) or in the presence (open triangles) of TPA. At the indicated times, cells were harvested and viable cells were counted as presented in Materials and Methods. (**D**) Cells were cultured as in (A) and nuclear extracts were isolated, subjected to SDS–PAGE and analyzed by western blotting with anti-Cdc6 antibody. A portion of total protein staining of transferred gel is shown as a loading control.

which triggers megakaryocytic differentiation in both cell lines. We found that *cdc6* RNA levels were diminished after culturing HEL(*Esg*+) in the presence of TPA (∼50%), while a ∼2-fold induction was reached in differentiated endoreplicating HEL cells (Figure 1A). In order to assess whether *cdc6* RNA levels increase in TPA-treated HEL cells was due to gene transcription maintained or to RNA stabilization, a similar experiment was performed in the presence of transcription inhibitor actinomycin D (ActD). As can be seen in Figure 1B, as early as 4 h after ActD treatment, *cdc6* RNA levels decreased in untreated HEL and HEL(Esg+) cells. Interestingly, in differentiated HEL cells, the increase obtained after TPA treatment was not observed for a long period in the presence of ActD, thus indicating that *cdc6* transcription is indeed active during HEL megakaryocytic differentiation. Other possibilities that could account for the different *cdc6* RNA levels, i.e. a different proliferation state of parental HEL and Esg-expressing cells, was discarded, as both cell lines stopped proliferating after TPA treatment [(8) and Figure 1C], but only parental HEL cells maintain *cdc6* expression at RNA (as shown above) and protein [(11) and Figure 1D) level.

These results also indicate that ectopic expression of *Esg* is able to interfere with the mechanisms that determine the maintenance of *cdc6* expression during megakaryocytic polyploidization.

A 1.7 kb *cdc6* 5′-upstream region containing putative Esg-binding E2 box motifs is responsible for the reduced *cdc6* promoter activation in HEL(*Esg*+) cells

Esg belongs to the *snail* family of zinc finger transcriptional repressors, which function appears to involve binding to E boxes recognized by a variety of basic helix–loop–helix (bHLH) transcriptional activators (24). We thus searched a 4537 bp regulatory *cdc6* region for putative sites that could be recognized by ectopic *Esg*. As shown in Figure 2A, we found three E2 boxes with perfect match to the described Esg consensus-binding site A/GCAGGTG, located at −4501 to −4495 (ACAGGTG), −3535 to −3529 (GCAGGTG) and −2995 to −2989 (GCAGGTG) (referred to as Boxes A, B and C, as indicated). Two other E2 boxes corresponding to much less specific Esg binding sites (C/TCAGGTG or G/ACAGNTG) were detected, at −1567 to −1561 (CCAGGTG) and at −2675 to −2679 (ACAGATG).

Transactivation assays were performed by transfecting HEL and HEL(*Esg*+) cells with a construct containing firefly

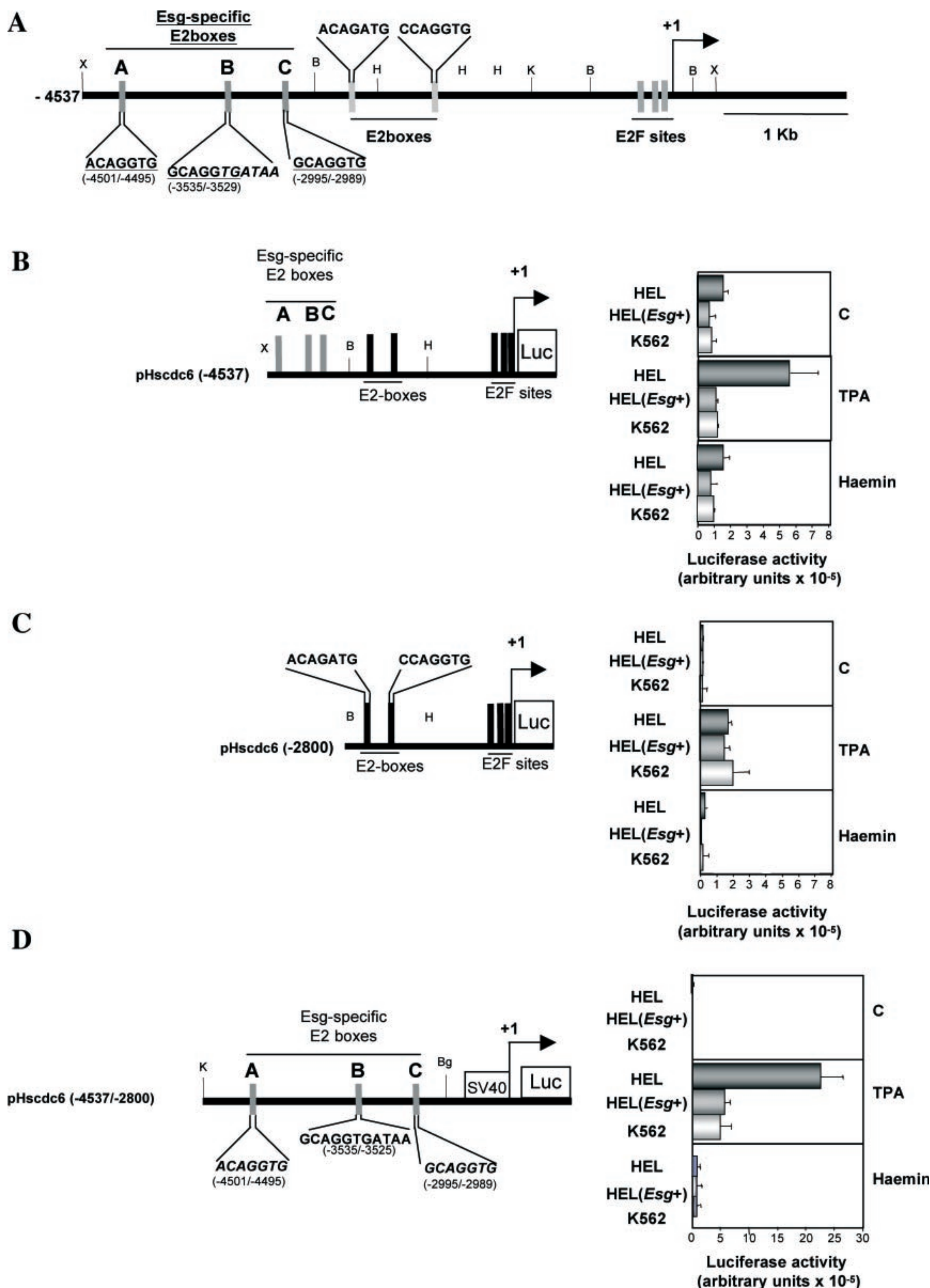

Figure 2. *cdc6* upstream region containing Esg-binding E2 boxes is responsible for reduced *cdc6* expression in non-endoreplicating cells. (**A**) Scaled diagram of the human *cdc6* gene showing the E2 boxes with perfect match to *Esg* consensus binding sites (referred to as *Esg*-specific E2 Boxes A, B and C) and E2 boxes corresponding to much less specific *Esg* binding sites and E2F-binding elements. The arrow indicates the transcription initiation site. Relevant restriction sites are also shown: X, XhoI; B, BamHI; H, HinDIII; K, KpnI; and B, BglII. (**B** and **C**) HEL, HEL(*Esg*+) and K562cells were transfected with pHscdc6(−4537) (B) or pHscdc6(−2800) (C), cultured in the absence or in the presence of TPA or hemin, as indicated, collected and assayed for luciferase activity. (**D**) HEL, HEL(*Esg*+) and K562 and cells were transfected with pHscdc6 (−4537/−2800), cultured in the absence or in the presence of TPA or hemin, as indicated, collected and assayed for luciferase activity. The data represent the mean ± SD of at least three independent experiments.

luciferase gene under the control of −4537 to +98 region of *cdc6* upstream region [pHsCdc6–luc(−4537), as in (17)]. As expected, the −4537 to +98 region was able to drive luciferase expression in exponentially growing cells. However, after TPA treatment, luciferase activity was significantly increased in parental HEL but not in differentiated HEL(*Esg*+) cells. Such activation was only seen after megakaryocytic differentiation, as the addition of hemin resulted in erythroid differentiation of both HEL and HEL(*Esg*+). This did not result in the induction of luciferase activity (Figure 2B). These results suggest that *cdc6* transcriptional activity is induced in endoreplicating HEL cells and that *Esg* may be modulating *cdc6* expression by repressing its promoter activity.

We have previously found that *cdc6* RNA levels are also downregulated in megakaryoblastic K562 cells. These cells respond to TPA, as HEL cells do, in terms of megakaryocytic differentiation but lack the ability to establish endoreplication cycles (11). In order to investigate whether increase in *cdc6* promoter activity was restricted to endomitotic differentiation, pHsCdc6–luc(−4537) construct was transfected in K562 cells and luciferase activity was determined before and after TPA or hemin treatment. Interestingly, no significant induction of the reporter activity could be detected after erythroid or megakaryocyte differentiation of K562 cells, thus suggesting a relationship between the ability to undergo endoreplication and TPA-driven stimulation of *cdc6* promoter activity.

To assess the involvement of *Esg*-specific E2 boxes on *cdc6* transcriptional regulation, a 1.7 kb fragment containing all three sites was deleted and the resulting pHsCdc6–luc(−2800) construct was transfected in HEL, HEL(*Esg*+) and K562 cells. Luciferase activity was then measured in undifferentiated, TPA- and hemin- treated cell extracts (Figure 2C). While hemin treatment did not result in any induction of luciferase activity, the transcriptional activity of this promoter region was increased at a same extent in all three lines after TPA treatment although no relationship with the ability of the cell to undergo endoreplication could be drawn. Similar results were obtained with pHsCdc6–luc(−1600), a deletion construct containing the E2F-binding elements responsible for cell-cycle-dependent *cdc6* expression regulation and lacking the E2 boxes corresponding to much less specific Esg binding sites at −1567 to −1561 and at −2675 to −2679 (data not shown).

These results suggest that the −4537 to −2800 region containing the *Esg*-specific E2 boxes could be responsible for endoreplication-driven stimulation of *cdc6* promoter activity and that *Esg* could be responsible for its inhibition when overexpressed in HEL cells.

In order to explore this hypothesis, a 1.7 kb fragment containing −4537 to −2800 of *cdc6* 5′ upstream region (only containing *Esg*-specific E2 boxes A, B and C), was inserted into a luciferase reporter vector under the control of the SV40 promoter [pHscdc6 (−4537/−2800)]. This fragment was then transfected into HEL, HEL(*Esg*+) and K562 cells. As seen in Figure 2D, this upstream region of *cdc6* gene lacked any detectable transcriptional activity in exponentially growing cells. However, after TPA treatment, this fragment was able to drive luciferase expression in all three lines. Interestingly, the degree of transcriptional activation was much higher in endoreplicating HEL than in non-endoreplicating HEL(*Esg*+) and K562 cells. Additionally, transcriptional activity of this 1.7 kb fragment was associated with megakaryocytic differentiation only, since no relevant luciferase activity could be detected in HEL, HEL(*Esg*+) or K562 cells that underwent erythrocytic differentiation in the presence of hemin.

Taken together, these results suggest that *cdc6* promoter activity is induced in endoreplicating HEL cells and that *Esg*-specific E2 boxes located at −4537 to −2800 may be critical for such a transcriptional regulation.

Escargot preferentially binds one of three E2 boxes located in the 5′ upstream region of *cdc6* gene

We hypothesized whether *Esg* would bind E2 boxes A, B and C present in −4537 to −2800 enhancer region '*in vitro*'. HEL and HEL(*Esg*+) cells were cultured in the presence or the absence of TPA and collected 48 h later to isolate nuclear extracts. At this time of differentiation, HEL cells were actively undergoing endoreplication cycles, which takes place up to 4 days (Figure 3A) and even up to 1 week (data not shown). Nuclear extracts were then analyzed by electrophoretic mobility shift assays (EMSA) using labeled double-stranded oligonucleotides containing E2 boxes A, B or C, and 9–11 surrounding nucleotides (Box A, Box B and Box C, as shown in Figure 3B). An oligonucleotide containing the consensus ACAGGTG binding site for *Esg* (hereafter referred to as the Esg Binding Element, EBE) was also included as a positive control. As already described (8), *Esg* present in HEL(Esg^+) nuclear extracts results in a specific pattern of two main, specific shifted bands when complexed with EBE (Figure 3B, lanes 1 and 2). When HEL(*Esg*+) nuclear extracts were incubated with Box A, Box B and Box C (Figure 3B, lanes 3–8), *Esg* complexes were prominently observed with Box B but barely detected with Boxes A and C. The pattern of shifted proteins was very similar in control and TPA differentiated HEL(*Esg*+). Thus, it appears that *Esg* is able to bind, at least '*in vitro*', to one out of three putative binding sites in *cdc6* 5′ upstream region. In addition, two main extra bands (complexes I and II) were detected to bind Box B, but not EBE, suggesting the existence of endogenous nuclear factors that would also bind Box B. To test this, EMSA was then performed with nuclear extracts from parental TPA-untreated and -treated HEL cells. As observed in the *Esg*-expressing cells, binding to Box B was most prominent than to Boxes A and C, and indeed EBE. More interestingly, complexes I and II were also detected to bind Box B while as expected, no *Esg* complexes were present (Figure 3D). Thus, it appears that HEL cells contain endogenous factors that also recognize this site.

Endogenous nuclear factors specifically bind the E2 box and an overlapping GATA site at −3535/−3529 of *cdc6* promoter

As Box B seemed to contain the most preferred DNA binding site for ectopic *Esg*, as well as for the putative endogenous nuclear factors, we wanted to confirm their binding specificity through competition experiments with unlabeled Box B, using extracts from HEL(*Esg*+) cells. Examination of Box B sequence revealed the existence of a GATA motif overlapping the described E2 box, which prompted us to perform binding assays on Box B in the presence of unlabeled $BEbox^M$ or $BGATA^M$ mutated oligonucleotides, as well as with an unrelated oligo containing four E2 boxes in tandem (E2 box

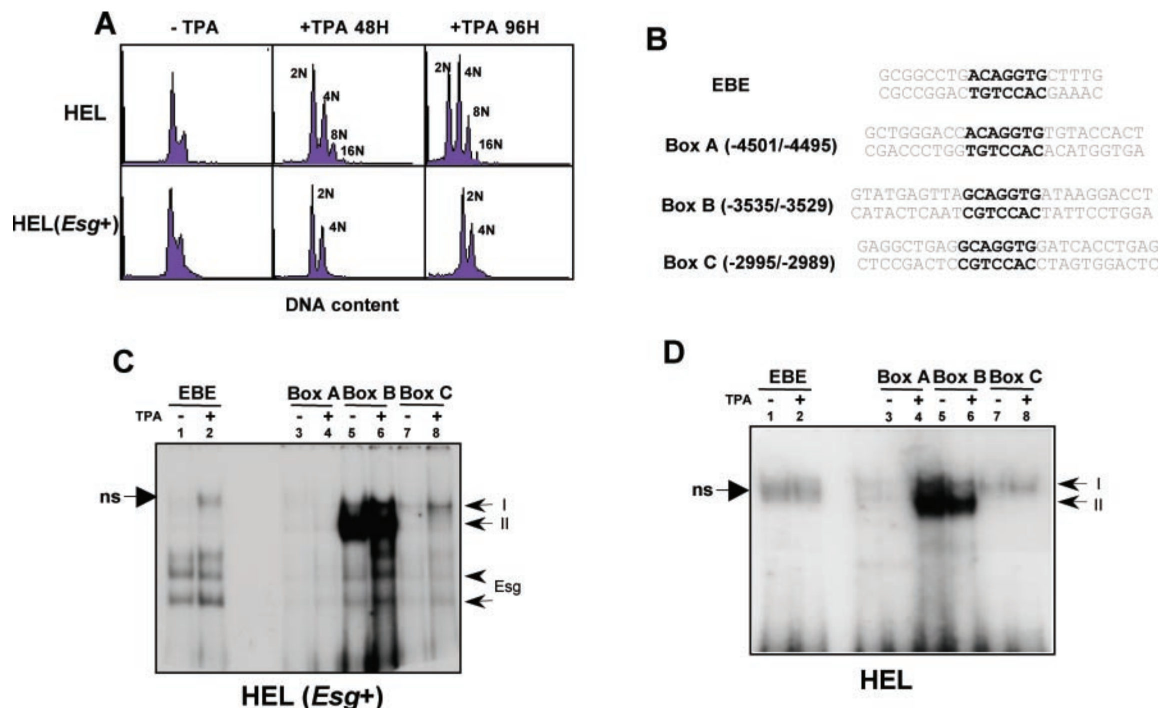

Figure 3. *Esg* and endogenous factors preferentially bind the E2 box/GATA site located at −3535/−3525 upstream *cdc6* gene. (**A**) HEL and HEL(*Esg*+) were cultured in the absence (−) or in the presence (+) of 10^{-8} M TPA for 48 or 96 h as indicated. DNA content pattern of these cells is shown as analyzed by flow cytometry. Vertical axis, relative number of cells; horizontal axis, relative red fluorescence (FL2) in a logarithmic scale, indicating DNA content per cell. (**B**) Synthetic oligonucleotides corresponding to the EBE, Box A, Box B and Box C used for EMSAs. (**C**) Nuclear extracts obtained from HEL(*Esg*+) cells cultured in the absence (−) or in the presence (+) of TPA were incubated with labeled EBE, Box A, Box B or Box C and analyzed using EMSA. (**D**) Nuclear extracts obtained from HEL cells cultured in the absence (−) or in the presence (+) of TPA were incubated with labeled EBE, Box A, Box B or Box C and analyzed using EMSA. Arrows indicate I, II and Esg complexes, as well a non-specific complex (ns).

tandem) (Figure 4A). First, competition with an excess of unlabeled Box B indicated that not only *Esg* but also the endogenous complexes I and II were specific (Figure 4B). Second, competition with an excess of unlabeled BEbox^M, BGATA^M or E2 box tandem revealed that: (i) *Esg* binds through the E2 box, since it was only competed with the oligos containing a wild-type E2 box (BGATA^M and E2 box tandem), but not with the BEbox^M; (ii) complex II binds through the GATA motif and does not require the E2 box, as it only competes with the oligo that contains wild-type GATA binding site (BEbox^M); (iii) complex I binds through the E2 box, but also requires the surrounding intact GATA motif, since it competes with oligos containing intact E2 boxes (E2 box tandem), but not with the oligos containing either a mutated E2 box (BEbox^M) or the mutated GATA motif (BGATA^M).

In order to further characterize the complexes binding requirements, in particular that of complex I, we performed EMSA analysis to determine the binding pattern to mutated oligos BEbox^M and BGATA^M (Figure 4C and D, respectively). As can be observed, only complex II could bind the oligo containing the mutated E2 box (Figure 4C). This is in agreement with complex II requiring only GATA motif and complex I requiring an intact E2 box, as it was also observed with *Escargot*. Inversely, only *Esg* complexes could be detected when labeled oligo BGATA^M was used (Figure 4D). Thus, this further strengthened the observation that complex I requires both intact E2 box and GATA motifs to bind to Box B.

To further confirm the specificity of endogenous complexes, competition experiments were carried out with nuclear extracts from parental HEL cells, untreated or TPA differentiated. As shown in Figure 4E, the specificity of complexes I and II in both undifferentiated and differentiated cells was identical to that detected in HEL(*Esg*+) cells.

These results indicate that at least two endogenous complexes are present in HEL cells before and after TPA treatment: complex I which binds the E2 box, but requires intact adjacent GATA sequence and complex II which specifically binds to and only requires GATA motif.

GATA-1 and E2A proteins bind to the E2 box/GATA site '*in vitro*' and '*in vivo*'

In an attempt to identify endogenous factors contained in complexes I and II, we investigated whether individual proteins with specific ability to bind the E2 box/GATA motif were involved in the binding to Box B site. One obvious candidate was GATA-1, since this protein is expressed in megakaryoblastic cells and affects megakaryocytic development. On the other hand, an independent one-hybrid experiment aimed to look for putative endogenous factors that would bind the EBE element allowed us to identify E2A (both E12 and E47) and E2-2 bHLH proteins as endogenous factors recognizing the G/ACAGGTG E2 box (N. V. and C. C., unpublished data). Thus, EMSA analysis was performed using HEL nuclear extracts in the presence or absence

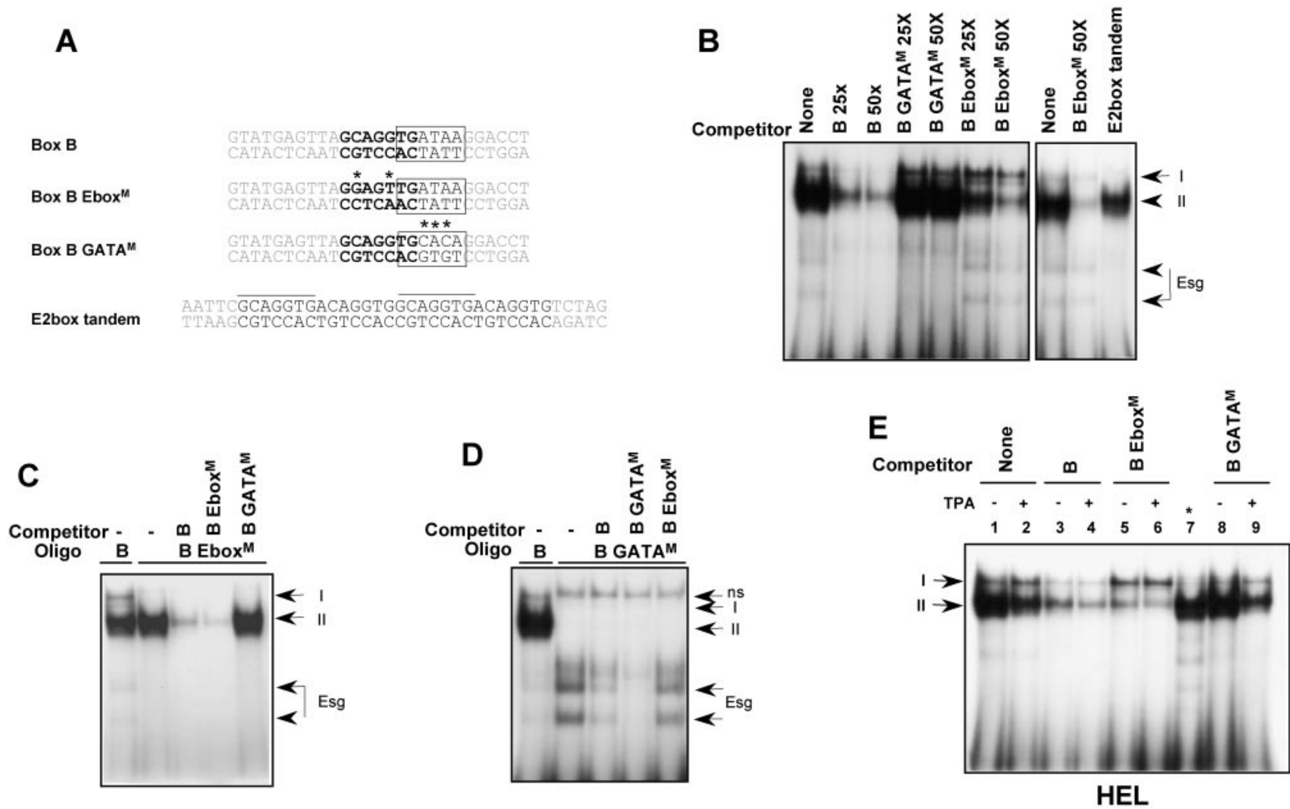

Figure 4. Endogenous nuclear factors require both E2 box and GATA motifs to bind the −3535/−3525 E2 box/GATA site. (**A**) Synthetic oligonucleotides corresponding to Box B, Box B Ebox^M, Box B GATA^M and the E2 box tandem used for EMSAs. E2 box appears in boldface and GATA site is boxed. Asterisks show nucleotide modified to obtain mutant E2 box and GATA elements. Esg-specific E2 boxes are marked with a line on the E2 box tandem oligonucleotide. (**B**) Nuclear extracts from HEL(*Esg*+) cells were incubated with labeled Box B in the absence (None) or the presence of an excess of unlabelled Box B (B), Box B Ebox^M (B Ebox^M), Box B GATA^M (B GATA^M) or E2 box tandem (25 and 50 times, as indicated) and analyzed using EMSA. (**C** and **D**) Binding reactions were carried out with the following labeled oligonucleotides: Box B Ebox^M (C) or Box B GATA^M (D) in the absence (−) or the presence of 50-fold excess of unlabeled Box B, Box B Ebox^M or Box B GATA^M. A reaction containing labeled Box B is shown in the first lane of both (C) and (D) as a control. Arrows indicate I, II and Esg complexes, as well a non-specific complex (ns). (**E**) Nuclear extracts from HEL cells cultured in the absence (−) or in the presence (+) of TPA were incubated with labeled Box B in the absence (None, lanes 1 and 2) or the presence of 50-fold excess of unlabeled Box B (lanes 3–4), Box B Ebox^M (lanes 5 and 6) or Box B GATA^M (lanes 8 and 9) and analyzed using EMSA. For comparison purposes, nuclear extracts from HEL(*Esg*+) cells were incubated with labeled Box B (*, lane 7).

of specific antibodies against these proteins. As shown in Figure 5A, complex II faced severe competition in the presence of anti GATA-1 antibody, whereas no effect was detected with an equivalent amount of non-immune serum, or a GATA-2 antibody. Also, pre-incubation with antiE2A antibody, but not with non-immune serum, resulted in downregulation of complex I (Figure 5B).

These results suggest that while Esg binds the E2 box motif, E2A bHLH factors could be part of the E2 box binding complex I, dependent on the intact E2 box/GATA sequence of Box B site. On the other hand, GATA-1 containing complex II would bind the GATA motif, as illustrated in Figure 5C.

Owing to the '*in vitro*' nature of EMSA experiments, we hypothesized whether the endogenous proteins were bound to the E2 box/GATA site '*in vivo*'. Chromatin Immunoprecipitation (ChIP) experiments were then performed by isolating cross-linked chromatin from HEL nuclear extracts. Both anti-GATA-1 and anti-E2A antibodies, as well as the corresponding non-immune Ig controls, were used to immunoprecipitate the fragmented chromatin and after reversal of the cross-link, PCR was performed to detect the presence of the E2 box/GATA site of *cdc6* upstream region. As can be seen in Figure 5D, in both anti-GATA-1 and anti-E2A immunoprecipitated fractions, a 249 bp band corresponding to the expected amplimer for this site could be detected, whereas no amplification was obtained in the non-immune Ig-precipitated fractions. A similar specific presence of site B could be detected with two independent anti-GATA-1 antibodies. Anti-GATA-1, unlike anti-E2A, antibodies were also able to immunoprecipitate a fragment corresponding to a region of gamma-globin promoter, already shown to bind GATA-1 (25). No amplified DNA was obtained when PCR was performed using primers located approximately −2000 upstream *cdc6* promoter. Both GATA-1 and E2A binding to B site could also be detected when ChIP was performed in HEL cells treated with TPA (Figure 5E), indicating that both proteins are bound to the endogenous regulatory site in differentiated cells. Similar results were obtained with TPA-treated and -untreated HEL(*Esg*+) cells (Figure 5E).

These data indicate that GATA-1 and E2A proteins are able to bind the E2 box/GATA site '*in vivo*'.

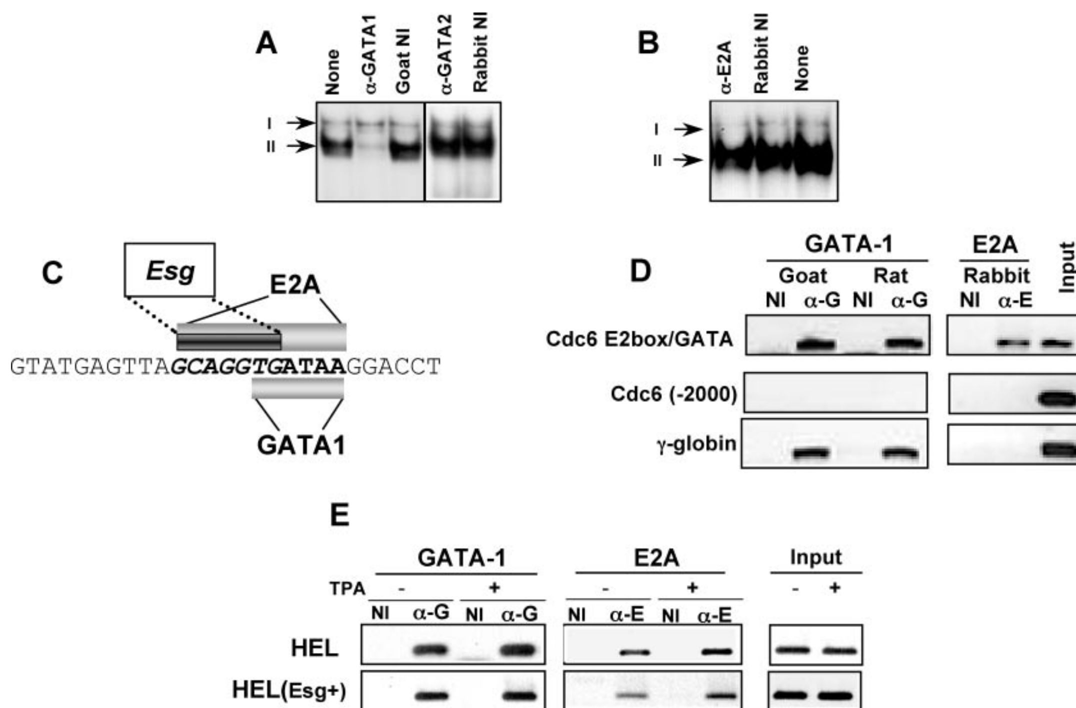

Figure 5. Endogenous GATA-1 and E2A bind the −3535/−3525 E2 box/GATA site both *in vitro* and *in vivo*. (**A**) Nuclear extracts from HEL(*Esg*+) cells were incubated with labeled Box B in the absence (None) or in the presence of goat anti-GATA-1 antibody (α-GATA1), non-immune goat serum (Goat NI), rabbit anti-GATA-2 antibody (α-GATA2) or non-immune rabbit serum (Rabbit NI), and analyzed using EMSA. (**B**) A similar experiment was performed in the presence of rabbit anti-E2A antibody (α-E2A) or non-immune rabbit serum (Rabbit NI). Arrows show complexes I and II. (**C**) Schematic representation of the binding requirements of GATA1, E2 box and *Escargot* complexes to E2 box/GATA motif. (**D**) ChIP assay was performed using fragmented cross-linked chromatin from HEL cells and anti-GATA-1 antibodies, αG (Goat C20, or Rat N6), anti-E2A antibody, αE (RabbitV18) or the corresponding non-immune immunoglobulins, NI. One-sixth of total purified DNA was amplified by PCR with primers specific for the fragment including the E2 box-GATA site (Cdc6 E2 box/GATA) as well as an unrelated fragment (Cdc6-2000) of *cdc6* upstream region. A fragment in the A-γ-globin locus (γ-globin), is included as a positive control. PCRs were also performed using the fragmented chromatin purified prior to antibody incubation (Input). (**E**) ChIP assays were performed using fragmented cross-linked chromatin from HEL and HEL(*Esg*+) cells cultured in the absence (−) or in the presence (+) of TPA and non-immune goat or rabbit Ig (NI), goat anti-GATA-1 (α-G) or rabbit anti-E2A antibodies (α-E).

The enhancer activity of −4537 to −2800 *cdc6* 5′ upstream region in megakaryocytic endoreplicating cells is mainly mediated though the E2 box/GATA-binding motif

We next hypothesized whether the E2 box overlapping a GATA motif was implicated in HEL endoreplication-driven enhancer activity displayed by the entire 1.7 kb fragment. Mutations at the E2 box as well as at GATA-binding element in site B were then introduced in pHscdc6 (−4537/−2800). The resulting plasmids (BEbox^M and BGATA^M) as well as the wild-type construct (Figure 6A) were transfected in HEL, HEL(*Esg*+) and K562 cells and luciferase activation was determined in TPA-treated and -untreated cells. As already observed, TPA treatment of cells transfected with the wild-type construct resulted in upregulation of its reporter activity in all three lines and the degree of transcriptional activation was highest in the endoreplicating HEL (Figure 6B). However, the introduction of mutations in the E2 box or the GATA motifs provoked significant changes in the TPA-promoted reporter activity. Thus, the construct containing the E2 box mutation led to a 50% decrease in TPA-promoted inducibility of the 1.7 kb fragment in HEL cells. Interestingly, when BGATA^M construct was transfected, a strong loss of activity was obtained (>80%). It seems therefore that an intact E2 box-GATA motif is required for optimal enhancer activity in endoreplicating HEL cells.

As expected from our above findings suggesting that *Esg* might be interfering with the transcriptional activation through the E2 box, transfection of BEbox^M construct in HEL(*Esg*+) cells resulted in higher TPA induction of reported activity than the observed with the wild-type version. In fact, the transcriptional activation was almost identical to that obtained in HEL cells transfected with the same BEbox^M construct. Similar to that observed in HEL cells, transfection of BGATA^M construct in HEL(*Esg*+) cells caused a severe decrease in TPA promoted inducibility when compared to the wild-type fragment. Interestingly, comparison of BGATA^M reporter activity in HEL and HEL(*Esg*+) cells showed the inhibitory effect of *Esg*, in accordance with the fact that *Esg* does not bind through the GATA binding site. This further strengthens the idea that the E2 box-GATA motif is mainly responsible for the enhancer activity displayed by pHscdc6 (−4537/−2800) in endoreplicating cells.

Additionally, we observed that TPA induced luciferase activity did not suffer any significant modification when the BEbox^M and the BGATA^M constructs were transfected in non-endoreplicating K562, when compared with the wild-type fragment. This suggests that this site is only functional

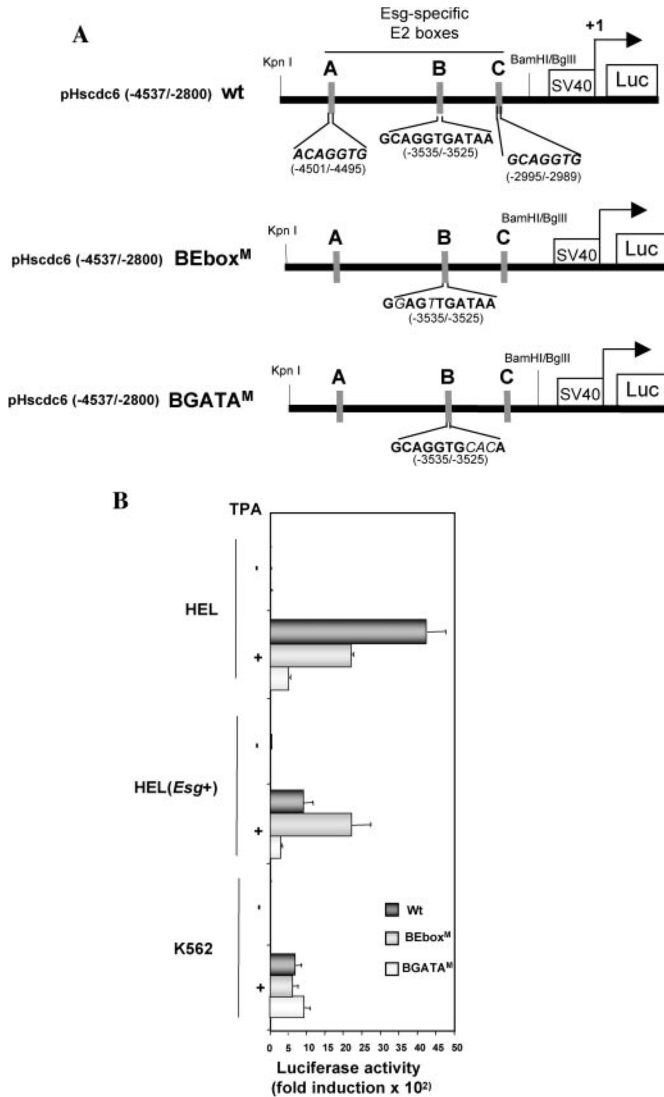

Figure 6. Enhancer activity of −4537 to −2800 *cdc6* 5′ upstream region in megakaryocytic endoreplicating cells is mainly mediated through the E2 box/GATA motif. (**A**) Diagrams of pHscdc6(−4537/−2800) (wt) and Ebox (BEbox^M) and GATA (BGATA^M) mutant constructs. (**B**) HEL, HEL(*Esg*+) and K562 cells were transfected with wild-type or mutant constructs. Transfected cells were cultured in the absence (−) or in the presence (+) of TPA, collected and assayed for luciferase activity. The data represent the mean ± SD of at least three independent experiments.

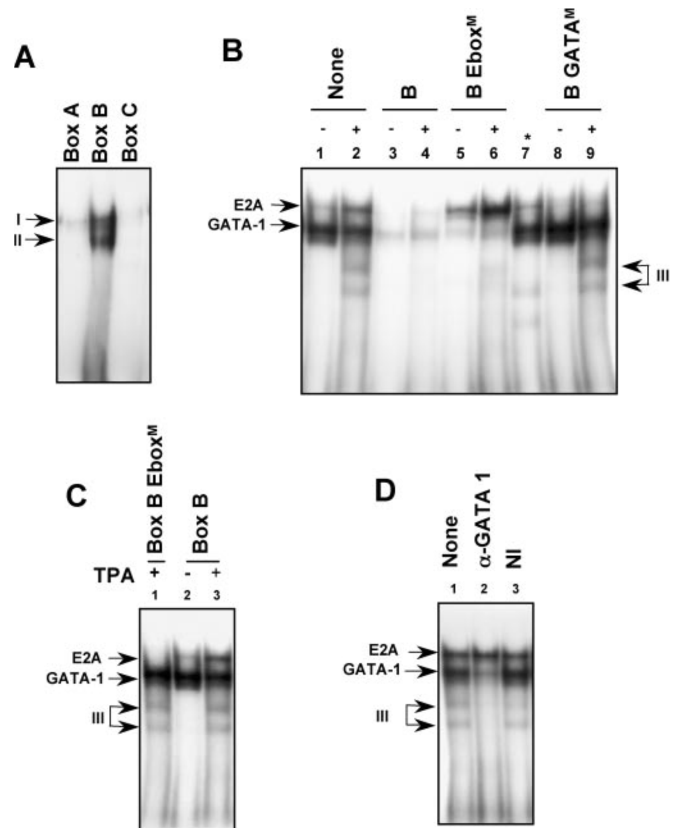

Figure 7. Non-endoreplicating K562 cells show additional GATA-1-containing complexes after megakaryocytic differentiation. (**A**) Nuclear extracts obtained from exponentially growing K562 cells were incubated with labeled Box A, Box B or Box C oligonucleotides and analyzed by EMSA. (**B**) Nuclear extracts from K562 cells cultured in the absence (−) or in the presence (+) of TPA were incubated with labeled Box B in the absence (None, lanes 1 and 2) or in the presence of 50-fold excess of unlabeled Box B (lanes 3 and 4), Box B Ebox^M (lanes 5 and 6) or Box B GATA^M (lanes 8 and 9) and analyzed using EMSA. For comparison purposes, nuclear extracts from HEL(*Esg*+) cells were incubated with labeled Box B (*, lane 7). (**C**) Nuclear extracts from K562 cells cultured in the absence (−) or in the presence (+) of TPA were incubated with labeled Box B Ebox^M or Box B and analyzed using EMSA. (**D**) Nuclear extracts from TPA-treated K562 cells were incubated with labeled Box B in the absence (None) or the presence of goat anti-GATA-1 antibody (α-GATA1) or non-immune goat serum (NI) and analyzed using EMSA. Arrows indicate the already described E2A and GATA-1 containing complexes, as well as GATA-1-containing complex III.

in HEL cells and that it is related to their ability to undergo endoreplication.

Altogether, these results show that the −4537 to −2800 region of *cdc6* upstream region can be considered as an enhancer that activates *cdc6* expression in an endoreplication-related manner during megakaryocytic differentiation and that this is mainly mediated by the E2 box/GATA motif.

GATA-1 displays a distinct binding pattern to the −3535/−3525 E2 box/GATA site, depending on the cells ability to undergo endoreplication

As mentioned earlier, both HEL and K562 cells respond to TPA in terms of megakaryocytic differentiation, although only HEL cells become polyploid (7). Interestingly, TPA treatment results in downregulation of *cdc6* gene expression, at both RNA and protein level, in K562 cells, whereas its expression is maintained in HEL cells (11). Furthermore, the TPA-promoted enhancer activity of the 1.7 kb fragment containing −4537 to −2800 of *cdc6* 5′ upstream region containing the E2 box-GATA site was diminished when compared to HEL cells (see Figure 2D) and was not affected when E2 box^M and GATA^M constructs were transfected in K562 cells. We first investigated whether extracts from these cells would bind this motif, as observed for HEL and if so, whether they would present differences in the pattern of proteins able to bind the E2 box/GATA site of *cdc6* upstream region when differentiated with TPA treatment. EMSA was then performed on nuclear extracts from K562 cells and, as observed in HEL cells, binding to Box B was most prominent than to Boxes A and C (Figure 7A). As shown in Figure 7B, nuclear extracts

from K562 cells treated with TPA for 48 h, showed differences when compared with TPA-treated HEL cells (see Figure 4E, lane 2). Thus, although the pattern of GATA-1 and E2A-containing complexes was present in both undifferentiated and TPA-treated K562 cells, two extra bands (complex III) could only be detected in differentiated K562 cells. As seen in Figure 7B, competition experiments revealed that the specificity of GATA-1 and E2A-containing complexes was identical to that detected in HEL cells. However, BEbox^M oligonucleotide was able to compete in the binding of complex III to BEbox. Indeed, complex III remained detectable when an excess of unlabeled BGATA^M oligonucleotide was used as competitor, indicating that these faster migrating bands contained a protein whose binding to Box B site occurred through the GATA-binding element. Accordingly, GATA-1 and complex III, but not E2A complex, could be detected when EMSA was performed with TPA-treated K562 nuclear extracts and labeled BEbox^M oligonucleotide (Figure 7C). In an attempt to determine whether GATA-1 protein was part of complex III, supershift experiments were performed with TPA-treated K562 nuclear extracts. As seen in Figure 7D, the presence of anti-GATA-1 antibody abolished GATA-1-containing complex, as already described, as well as complex III, suggesting that complex III contained GATA-1 protein.

Altogether, these results indicate that GATA-1 and E2A specifically bind the E2 box/GATA motif present in *cdc6* 5′ upstream region in K562 cells, as it occurs in HEL cells. They also indicate that additional GATA-1 complexes bind to this motif in non-endoreplicating K562 cells after TPA-promoted differentiation, when compared to differentiated, endoreplicating HEL cells.

The pattern of GATA-1- and E2A-containing complexes that bind the E2 box/GATA site is identical in megakaryoblastic cells and primary megakaryocytes

In order to determine whether the pattern of binding of endogenous nuclear factors to the E2 box/GATA motif is associated with megakaryocytic lineage, EMSA experiments were performed with labeled site B oligonucleotide and nuclear extracts from different human cell lines. It is worth noting that only megakaryoblastic cells showed the same characteristic pattern of shifted proteins. HeLa (epithelial origin) or U937 (promonocytic) did not contain any detectable complex, whereas osteosarcoma-derived U2OS, and embryonic kidney 293 contained endogenous proteins able to bind this oligonucleotide, but the pattern of bound proteins was distinct from that detected in megakaryoblastic HEL, HEL(*Esg*+) K562 cells (Figure 8A).

We finally investigated whether primary megakaryocytes contained endogenous nuclear factors were able to bind '*in vitro*' the E2 box/GATA site. Umbilical cord blood CD34+ cells were cultured in the presence of thrombopoietin (Tpo) and the content of megakaryocytes was assessed by flow cytometry analysis of lineage-specific GpIIIa glycoprotein (CD61) expression. Typically, >70% of cultured cells were found to be CD61 positive (Figure 8C, Tpo culture). As can be seen in Figure 8B, only two retarded, specific bands could be detected when nuclear extracts from megakaryocytic-enriched cultured cells were incubated with labeled oligonucleotide B. Megakaryocytes were further purified by immunoselection with anti-GpIIb (CD41) antibodies. As shown in Figure 8C, separation resulted in >90% of CD61 positive, of which a significant 30% corresponded to polyploid cells (DNA content >8N). EMSA was performed with nuclear extracts isolated from CD41+ cells (MK), as well as from the MNCs from which they are derived, and that contain blood cells mainly (>90% lymphocytes and monocytes). The pattern of complexes bound to oligonucleotide B in MK extracts (Figure 8D, lanes 1 and 2) was identical to that found in the corresponding Tpo culture (Figure 8B), and also to that detected in megakayoblastic HEL cell line (lanes 7 and 8). No retarded bands were detected in non-megakaryocytic MNC cells (lane 3). Addition of anti-GATA-1 (lane 5) and anti-E2A (lane 10) antibodies to nuclear extracts from megakaryocytic-enriched cultured cells was able to abolish the formation of complexes II and I, respectively, indicating that these complexes contained E2A and GATA-1. Interestingly, Cdc6 was present in the terminally differentiated primary megakaryocytes (Figure 8E).

Taken together, these results indicate that the binding of E2A and GATA-1 containing complexes to the E2 box/GATA motif is a characteristic of the megakaryocytic lineage.

DISCUSSION

The data presented in this work indicate that a 1.7 kb fragment located at −4537 to −2800 bp upstream *cdc6* gene is a transcription enhancer active in megakaryocytic but not in erythrocytic differentiated cells nor in exponentially growing cells. We have found that the stimulation of *cdc6* promoter activity during megakaryocytic endoreplication is dependent on the integrity of a novel *cis*-regulatory element defined by an E2 box overlapping a GATA motif (GCAGGTGATAA) located at −3535/−3525. Our data also show evidence that this motif is recognized '*in vitro*' and '*in vivo*' by GATA-1 and E2A containing complexes that bind the E2 box/GATA site in a specific manner. Finally, our data indicate that in human primary megakaryocytes, GATA-1 and E2A transcription factors are able to bind this motif, at least '*in vitro*', with identical pattern to that detected in megakaryoblastic cell lines.

The E2 box/GATA motif in *cdc6* upstream region appears to be critical for the endoreplication-related stimulation of the promoter activity. First, when either the E2 box or the GATA elements are mutated, the transcriptional activity of the entire 1.7 kb region is severely impaired. Second, ectopic expression of transcriptional repressor *Esg*, which specifically binds the E2 box element of the E2 box/GATA motif, leads to a significant inhibition of the enhancer activity in HEL cells. This parallels the loss of Cdc6 expression (11) and of endoreplication ability (8). Interestingly, ectopically expressing mouse Snail HEL cells, which also have lost the ability to undergo endomitosis, show identical pattern of binding of nuclear complexes to the E2 box/GATA motif than HEL(*Esg*+) cells (N. Vilaboa, R. Bermejo and C. Calés, unpublished data). To our view, it is highly relevant for *Esg* to interfere with the transcriptional activity of *cdc6* gene in differentiated HEL cells, especially as such inhibitory effect seems to affect the endogenous gene transcription as well as an exogenous, reporter gene. It has recently been reported that mouse Sna can attenuate the cell cycle in epithelial MDCK cells, as well as in

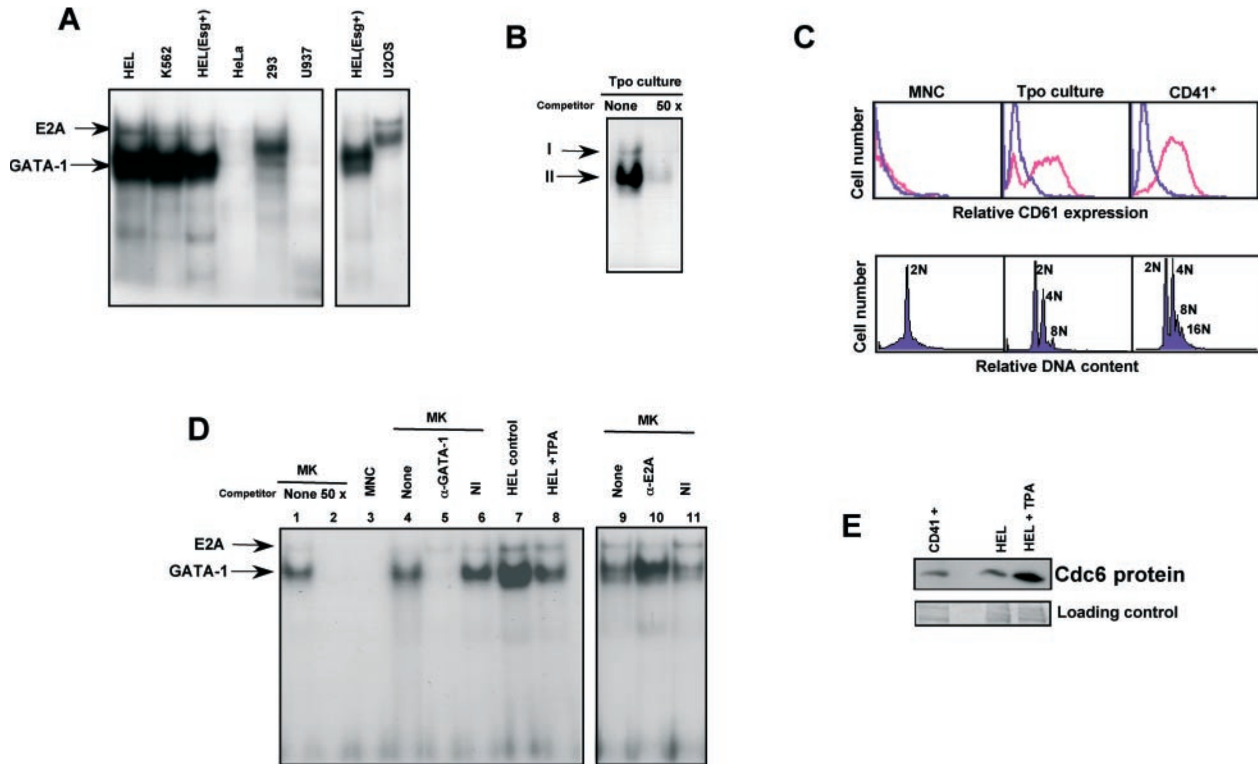

Figure 8. GATA-1 and E2A containing complexes are characteristic of megakaryocytes. (**A**) Nuclear extracts from exponentially growing HEL, K562, HEL(*Esg*+), HeLa, 293 U937 and U2OS cells were incubated with labeled Box B oligonucleotide and analyzed using EMSA. Arrows indicate E2A and GATA-1 containing complexes. (**B**) Nuclear extracts from megakaryocytic-enriched primary human cells (Tpo culture) were incubated with labeled Box B in the absence (None) or the presence (50×) of a 50-times excess of unlabeled Box B and analyzed using EMSA. (**C**) MNC isolated from human Umbilical Cord Blood, megakaryocytic-enriched cells obtained after culture of MNC-derived CD34+ cells in the presence of Tpo (Tpo-culture) and immuno-selected megakaryocytes (CD41+) were labeled with FITC-conjugated anti-GpIIIa (CD61) antibodies (upper panel) or Propidium Iodide (lower panel) and analyzed using flow cytometry. (**D**) Nuclear Extracts from CD41+ cells (MK), MNC and HEL cells untreated (control) or TPA-treated (+TPA), were incubated with labeled Box B and analyzed using EMSA. Lane 2: labeling reaction was performed with MK extracts in the presence of 50 times excess unlabeled Box B. Lanes 4–6 and 9–11: labelling reaction was performed with MK extracts in the absence (None) or the presence of non-inmune serum (NI) or the indicated antibodies (α-GATA-1 or α-E2A). (**E**) Nuclear extracts were obtained from CD41+ and untreated or TPA-treated HEL cells. Proteins were subjected to SDS–PAGE and analyzed by western blotting with anti-cdc6 antibody. A portion of total protein staining of transferred gel is shown as a loading control.

early embryo development, through partial repression of cyclin D2 promoter activity (26). *Esg* does not affect cell cycle progression in proliferating megakaryoblastic HEL cells (8). In contrast, *Esg* inhibition of megakaryocytic endoreplication seems to be related to its ability to interfere with transcriptional activity of *cdc6* promoter in differentiated HEL (*Esg*+) cells.

It has already been described that a CAGGTG E2 box and a GATA site define a composite, bipartite regulatory element involved in the regulation of a number of erythrocytic and megakaryocytic specific promoters (27,28) including GATA-1 gene itself (29). The formation of a multimeric complex formed by GATA-1, Tal1, E47, Lmo2 and Ldb1A seems to rely on a specific spacing of ∼8–10 bp (∼1 helix turn) between the E2 box and the GATA motifs (28). The *cis*-acting element we have found to direct *cdc6* transcription during megakaryocytic endoreplication appears to be somehow distinct from that found in promoters of lineage specific genes. Notably, there is no spacing between the E2 box and GATA motifs, and in fact they overlap each other. To our knowledge, this is the first time that this type of E2 box/GATA site is reported and that a differentiation-related role in transcriptional control of a licensing factor involved in polyploidization

Ebox-GATA element **CANNTGATAR**

H.Sapiens **(-3556) *G*CAGGTGATAA(-3546)**

R.Norvegicus **(-5942)*G*CACTTGATAG(-5932)**

M.Musculus **(+22974)*G*CAGGTGATAG(+22984)**

Figure 9. A Ebox-GATA element is present in cdc6 locus of mouse and rat. *In silico* search for the presence of an E2box-GATA motif through ∼50 000 bp of rat chromosome10 (87 675 000–87 725 000 bp) and mouse chromosome 1 (98 530 000–98 610 000 bp) which spans *cdc6* gene and surrounding upstream and downstream 15 000 bp. Both human and rat elements are located upstream first exon, whereas mouse element lies in the 3′-untranslated region. Numbers refer to the positions respective to reported first exon.

regulation assigned to such a *cis*-regulatory element. It is also highly relevant that an identical E2 box/GATA motif was also found in the mouse *cdc6* gene locus, although in the 3′ untranslated end of the gene instead of the upstream region, and a similar E box/GATA motif in a similar position of the 5′ end of the rat locus (Figure 9). This suggests that similar mechanisms

may be acting in rodent megakaryocytic polyploidization. The E2box/GATA element is also present in the human GATA-1 locus at chromosome X (N. V., R. B. and C. C., unpublished data). This last observation might be indicative of an additional lineage-specific regulation of GATA-1 expression during megakaryocytic maturation.

This paper provides the first evidence that E2A and GATA-1 containing complexes, individually or in coordination, could be regulating *cdc6* expression. While both GATA-1 and E2A have been proposed to affect cell cycle control (30–32), to date only E2A has been described to directly affect the expression of cdk inhibitors, such as p21, p15 and p16 (33,34). That *cdc6* expression might be specifically controlled during megakaryocytic endoreplication by a lineage-specific transcriptional regulator, such as GATA-1 is reinforced by the earlier proposed implication of this transcription factor in determining proper megakaryocytic polyploidization (35). For instance, GATA-1 deficient megakaryocytes do not undergo endomitosis and show evidence of retarded nuclear maturation (35,36). Moreover, X-linked macrothrombocytopenia is due to various mutations resulting in GATA-1 lack-of-function and an increased number of dysplastic micromegakaryocytes (37,38). Micromegakaryocytes are hypoploid megakaryocytes (39–41), most probably unable to undergo endoreplication cycles. Disruption of the GATA binding motif strongly diminished the *cdc6* promoter activity in endoreplicating cells while mutation of the E2 box decreased it to a lesser extent. This is in accordance to the reported predominant importance of the GATA site compared to the E2 box present in the composite E-box–10 bp GATA element 3.7 kb upstream GATA-1 promoter (29). Also, in a non-megakaryocytic cell type, a transcriptional complex between GATA-4 and a bHLH factor (dHAND) requires the GATA site to be intact, whereas the E-box motif is dispensable (42). Finally, it is also worth noting that the '*in vitro*' binding of E2A-containing complex to the E2 box/GATA motif requires an intact GATA site, especially when it would be expected that the binding of a bHLH protein took place exclusively through the E2 box element.

One interesting result of this paper is the fact that GATA-1 complexes bound to the E2 box/GATA-1 motif '*in vitro*' show significant differences between endoreplicating and non-endoreplicating differentiated megakaryocytic cells. In particular, a faster migrating GATA complex is detected in non-endoreplicating differentiated K562 cells. The enhancer activity of the E2 box/GATA motif-containing region is highly reduced in these cells. Certainly, it will be of great interest to investigate whether GATA-1 forms multimeric complexes with other proteins during megakaryocytic differentiation and whether the composition of the complexes and/or post-translational modifications of GATA-1 are different depending on the cell ability to endoreplicate. GATA-1 has been shown to form multimeric complexes with a variety of other transcription factors (i.e. FOG, EKLF, LMO2, SP1, CBP and Fli-1) and the resulting complexes can even display stimulatory or repressive transcriptional activity (43).

We have found that human cell lines of different lineages contain endogenous factors able to bind this motif, at least '*in vitro*', and that the resulting complex patterns are distinct of those found in megakaryoblastic cells. Noteworthy, our '*in vitro*' experiments indicate that GATA-1 and E2A-containing complexes are able to bind to the E2 box/GATA motif present in cultured human primary megakaryocytes, reinforcing the hypothesis that *cdc6* promoter activity could be regulated by differentiation-related transcription factors such as E2A and GATA-1.

ACKNOWLEDGEMENTS

We acknowledge Dr Nakamura for the generous gift of pHscdc6 construct. We also thank Dr Aurora Sánchez Pacheco for help with ChIP experiments setting and Dr Miguel Vidal for helpful discussions. This work was supported by grants PM98-0046 from the Spanish Ministry of Science and Technology, CAM 08.3/0037/2001 from the 'Comunidad Autónoma de Madrid' and FIS 01/0090-02 from 'Fondo de Investigaciones Sanitarias' to C.C. N.V. was partly supported by a Postdoctoral Fellowship from 'Comunidad Autónoma de Madrid' (Spain).

REFERENCES

1. Edgar,B.A. and Orr-Weaver,T.L. (2001) Endoreplication cell cycles: more for less, *Cell*, **105**, 297–306.
2. Ravid,K., Lu,J., Zimmet,J.M. and Jones,M.R. (2002) Roads to polyploidy: the megakaryocyte example. *J. Cell Physiol*., **190**, 7–20.
3. Zimmet,J. and Ravid,K. (2000) Polyploidy: occurrence in nature, mechanisms, and significance for the megakaryocyte-platelet system. *Exp. Hematol*., **1**, 3–16.
4. Garcia,P. and Cales,C. (1996) Endoreplication in megakaryoblastic cell lines is accompanied by sustained expression of G_1/S cyclins and downregulation of cdc25C. *Oncogene*, **13**, 695–703.
5. Zimmet,J.M., Ladd,D., Jackson,C.W., Stenberg,P.E. and Ravid,K. (1997) A role for cyclin D3 in the endomitotic cell cycle. *Mol. Cell. Biol*., **17**, 7248–7259.
6. Datta,N.S., Williams,J.L. and Long,M.W. (1998) Differential modulation of G_1-S-phase cyclin-dependent kinase 2/cyclin complexes occurs during the acquisition of a polyploid DNA content. *Cell Growth Differ*., **9**, 639–650.
7. Garcia,P., Frampton,J., Ballester,A. and Cales,C. (2000) Ectopic expression of cyclin E allows non-endomitotic megakaryoblastic K562 cells to establish re-replication cycles. *Oncogene*, **19**, 1820–1833.
8. Ballester,A., Frampton,J., Vilaboa,N. and Cales,C. (2001) Heterologous expression of the transcriptional regulator escargot inhibits megakaryocytic endomitosis. *J. Biol. Chem*., **276**, 43413–43418.
9. Fuse,N., Hirose,S. and Hayashi,S. (1994) Diploidy of *Drosophila* imaginal cells is maintained by a transcriptional repressor encoded by escargot. *Genes Dev*., **8**, 2270–2281.
10. Nakayama,H., Scott,I.C. and Cross,J.C. (1998) The transition to endoreduplication in trophoblast giant cells is regulated by the mSNA zinc finger transcription factor. *Dev. Biol*., **199**, 150–163.
11. Bermejo,R., Vilaboa,N. and Cales,C. (2002) Regulation of CDC6, geminin, and CDT1 in human cells that undergo polyploidization. *Mol. Biol. Cell*, **11**, 3989–4000.
12. Diffley,J.F. and Labib,K. (2002) The chromosome replication cycle. *J. Cell. Sci*., **115**, 869–872.
13. Fujita,M., Yamada,C., Goto,H., Yokoyama,N., Kuzushima,K., Inagaki,M. and Tsurumi,T. (1999) Cell cycle regulation of human CDC6 protein: intracellular localization, interaction with the human mcm complex, and CDC2 kinase-mediated hyperphosphorylation. *J. Biol. Chem*., **274**, 25927–25932.
14. Castellano,M.M., del Pozo,J.C., Ramirez-Parra,E., Brown,S. and Gutierrez,C. (2001) Expression and stability of *Arabidopsis* CDC6 are associated with endoreplication. *Plant Cell*, **13**, 2671–2686.
15. Ramos,G.B., de Almeida Engler,J., Ferreira,P.C. and Hemerly,A.S. (2001) DNA replication in plants: characterization of a cdc6 homologue from *Arabidopsis thaliana*. *J. Exp. Botany*, **52**, 2239–2240.
16. Hateboer,G., Wobst,A., Petersen,B.O., Le Cam,L., Vigo,E., Sardet,C. and Helin,K. (1998) Cell cycle-regulated expression of mammalian CDC6 is dependent on E2F. *Mol. Cell. Biol*., **18**, 6679–6697.

17. Ohtani,K., Tsujimoto,A., Ikeda,M. and Nakamura,M. (1998) Regulation of cell growth-dependent expression of mammalian CDC6 gene by the cell cycle transcription factor E2F. *Oncogene*, **17**, 1777–1785.
18. Yan,Z., DeGregori,J., Shohet,R., Leone,G., Stillman,B., Nevins,J.R. and Williams,R.S. (1998) Cdc6 is regulated by E2F and is essential for DNA replication in mammalian cells. *Proc. Natl Acad. Sci. USA*, **95**, 3603–3608.
19. Asano,M. and Wharton,R.P. (1999) E2F mediates developmental and cell cycle regulation of ORC1 in *Drosophila*. *EMBO J*., **18**, 2435–2448.
20. de Jager,S.M., Menges,M., Bauer,U.M. and Murra,J.A. (2001) *Arabidopsis* E2F1 binds a sequence present in the promoter of S-phase- regulated gene AtCDC6 and is a member of a multigene family with differential activities. *Plant Mol. Biol*., **47**, 555–568.
21. Humbert,P.O., Verona,R., Trimarchi,J.M., Rogers,C., Dandapani,S. and Lees,J.A. (2000) E2f3 is critical for normal cellular proliferation. *Genes Dev*., **14**, 690–703.
22. Schlisio,S., Halperin,T., Vidal,M. and Nevins,J.R. (2002) Interaction of YY1 with E2Fs, mediated by RYBP, provides a mechanism for specificity of E2F function. *EMBO J*., **21**, 5775–5786.
23. Forsberg,C.E., Downs,K.M., Christensen,H.M., Im,H., Nuzzi,P.A. and Bresnick,E.H. (2000) Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc. Natl Acad. Sci. USA*, **97**, 14494–14499.
24. Nieto,M.A. (2002) The snail superfamily of zinc-finger transcription factors. *Nature Rev. Mol. Cell. Biol*., **3**, 155–166.
25. Duan,Z., Stamatoyannopoulos,G. and Li,Q. (2001) Role of NF-Y in *in vivo* regulation of the gamma-globin gene. *Mol. Cell. Biol*., **21**, 3083–3095.
26. Vega,S., Morales,A.V., Ocana,O.H., Valdes,F., Fabregat,I. and Nieto,M.A. (2004) Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev*., **18**, 1131–1143.
27. Cohen-Kaminsky,S., Maouche-Chrétien,L., Vitelli,L., Vinit,M.-A., Blanchard,I., Yamamoto,M., Peschle,C. and Romeo,P.-H. (1998) Chromatin imunoselection defines a TAL-1 target gene. *EMBO J*., **17**, 5151–5160.
28. Wadman,I.A., Osada,H., Grutz,G.G., Agulnick,A.D., Westphal,H., Forster,A. and Rabbitts,T.H. (1997) The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J*., **16**, 3145–3157.
29. Vyas,P., McDevitt,M.A., Cantor,A.B., Katz,S.G., Fujiwara,Y. and Orkin,S.H. (1999) Different sequence requirements for expression in erythroid and megakaryocytic cells within a regulatory element upstream of the GATA-1 gene. *Development*, **126**, 2799–2811.
30. Peverali,F.A., Ramqvist,T., Saffrich,R., Pepperkok,R., Barone,M.V. and Philipson,L. (1994) Regulation of G_1 progression by E2A and Id helix–loop–helix proteins. *EMBO J*., **13**, 4291–4301.
31. Dubart,A., Romeo,P.H., Vainchenker,W. and Dumenil,D. (1996) Constitutive expression of GATA-1 interferes with the cell-cycle regulation. *Blood*, **87**, 3711–3721.
32. Whyatt,D.J., Karis,A., Harkes,I.C., Verkerk,A., Gillemans,N., Elefanty,A.G., Vairo,G., Ploemacher,R., Grosveld,F. and Philipsen,S. (1997) The level of the tissue-specific factor GATA-1 affects the cell-cycle machinery. *Genes Funct*., **1**, 11–24.
33. Prabhu,S., Ignatova,A., Park,S.T. and Sun,X.-H. (1997) Regulation of the expression of cyclin-dependent kinase inhibitor p21 by E2A and Id proteins. *Mol. Cell. Biol*., **17**, 5888–5896.
34. Pagliuca,A., Gallo,P., De Luca,P. and Lania,L. (2000) Class A helix–loop–helix proteins are positive regulators of several cyclin-dependent kinase inhibitors' promoter activity and negatively affect cell growth. *Cancer Res*., **60**, 1376–1382.
35. Vyas,P., Ault,K., Jackson,C.W., Orkin,S.H. and Shivdasani,R.A. (1999) Consequences of GATA-1 deficiency in megakaryocytes and platelets. *Blood*, **93**, 2867–2875.
36. Shivdasani,R.A., Fujiwara,Y., McDevitt,M.A. and Orkin,S.H. (1997) A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J*., **16**, 3965–3973.
37. Nichols,K.E., Crispino,J.D., Poncz,M., White,J.G., Orkin,S.H., Maris,J.M. and Weiss,M.J. (2000) Familial dyserythropoietic anaemia and thrombocytopenia due to inherited mutation in GATA1. *Nature Genet*., **24**, 266–270.
38. Freson,K., Matthijs,G., Thys,C., Mariën,P., Hoylaerts,M.F., Vermylen,J. and Van Geet,C. (2002) Different substitutions at residue 218 of the X-linked transcription factor GATA-1 lead to altered clinical severity of macrothrombocytopaenia and anemia and are associated with variable skewed X inactivation. *Hum. Mol. Genet*., **11**, 147–152.
39. Hegyi,E., Nakazawa,M., Debili,N., Navarro,S., Katz,A., Breton-Gorius,J. and Vainchenker,W. (1991) Developmental changes in human megakaryocyte ploidy. *Exp. Hematol*., **19**, 87–94.
40. Olson,T.A., Levine,R.F. and Kelleher,J. (1992) Familial thrombocytopenia with micromegakaryocytes. *Am. J. Pediatr. Hematol. Oncol*., **14**, 248–254.
41. Levine,R.F., Olson,T.A., Shoff,P.K., Miller,M.K. and Weisman,L.E. (1996) Mature micromegakaryocytes: an unusual developmental pattern in term infants. *Br. J. Haematol*., **94**, 391–399.
42. Dai,Y.-S., Cserjesi,P., Markham,B.E. and Molkentin,J.D. (2002) The transcription factors GATA4 and dHAND physically interact to synergistically activate cardiac gene expression through a p300-dependent mechanism. *J. Biol. Chem*., **277**, 24390–24398.
43. Cantor,A.B. and Orkin,S.H. (2002) Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene*, **21**, 3368–3376.