3822–3830 Nucleic Acids Research, 1995, Vol. 23, No. 19

Cell cycle regulation of *cdc25C* transcription is mediated by the periodic repression of the glutamine-rich activators NF-Y and Sp1

Jörk Zwicker, Claudia Gross, Frances C. Lucibello, Mathias Truss, Frauke Ehlert, Kurt Engeland and Rolf Müller*

Institut für Molekularbiologie und Tumorforschung (IMT), Philipps-Universität Marburg, Emil-Mannkopff-Strasse 2, D-35033 Marburg, Germany

Received August 8, 1995; Accepted September 7, 1995

EMBL accession no. Z29077

ABSTRACT

The late S/G_2-specific transcription of the human *cdc25C* gene is dependent on an initiator-proximal repressor element (CDE) and an upstream activating sequence (UAS) of undefined nature. We now show that these upstream sequences harbour multiple *in vivo* protein binding sites that interact with transcriptional activators and form separable, context-independent functional modules. Major components of the UAS are a *bona fide* Sp1 site and three direct sequence repeats (Y_C-boxes). The Y_C-boxes interact with the CCAAT-box binding protein NF-Y and are critically dependent on synergistic interactions for efficient transcription activation. The NF-Y complexes, as well as Sp1, are constitutive activators, whose activation function is periodically repressed through the CDE. These observations indicate that the cell cycle regulation of *cdc25C* transcription is mainly due to the CDE-mediated repression of glutamine-rich activators.

INTRODUCTION

The mechanisms of cell cycle-dependent transcription are still poorly understood, although in some instances clear pictures are beginning to emerge. In mammalian cells, this applies particularly to factors of the E2F/DP family, which play important roles in the cell cycle-regulated activation of a set of genes in late G_1 and S, as well as the induction of DNA replication [for reviews see (1,2)]. The E2F/DP dimers become active in late G_1 through the cell cycle-regulated phosphorylation of pRB and its relatives, and their ensuing dissociation from the E2F/DP complex. Several genes induced in S-phase, including B-*myb*, *E2F-1* and *DHFR* have been shown to be under the control of E2F. Other cell cycle-controlled transcription factors in mammalian seem to exist, but their putative association with the cell cycle is poorly defined, as in the case of Sp1, c-Jun or Ets family members (for a review see ref. 2). Even less is known about the mechanisms involved in late S/G_2-specific transcription in mammalian cells.

We have chosen the human *cdc25C* gene as a model system to investigate the regulation of late S/G_2-specific transcription in mammalian cells, because of its clear cell cycle-regulated expression with peak levels in G_2 (3,4). Cdc25 was originally discovered in *S.pombe* as a cell cycle gene with an essential function in G2→M progression (for a review see ref. 5). Higher eukaryotes contain at least 3 genes with a significant degree of similarity to cdc25, the *cdc25C* gene being the direct homologue of the yeast gene. CDC25C is a protein tyrosine/threonine phosphatase which directly activates the cdc2/cyclin B complex prior to the entry into mitosis (reviewed in ref. 5), and therefore plays a key role in the control of cell cycle progression.

To elucidate the mechanisms involved in cell cycle-regulated transcription of the cdc25 gene, the human promoter was cloned and a comprehensive structure–function analysis was performed (4). Transient expression studies and *in vivo* footprinting studies led to the identification of a novel regulatory element, termed cell cycle dependent element (CDE), which is located directly adjacent to one of the transcription initiation sites and plays a key role in the periodic transcription of the *cdc25C* gene. The CDE is bound by a protein (complex) in G_0/G_1 that acts as a transcriptional repressor and is released in S/G_2. The CDE apparently does not interfere with basal transcription from the core promoter. Its function is dependent on a stretch of upstream sequences that is needed for transcriptional activation and can be replaced with the SV40 early enhancer region. This led to the hypothesis that the CDE may function by regulating the activity of an upstream activating sequence (UAS) in a cell cycle-dependent fashion. We now show the region upstream of the CDE indeed acts as an UAS, and identify multiple constitutive *in vivo* protein binding sites for the glutamine-rich activators CBF/NF-Y and Sp1 representing the targets for CDE-mediated repression.

MATERIALS AND METHODS

Cell culture, DNA transfection and luciferase assays

WI-38 cells (6) were obtained from the ATCC, and NIH3T3 cells were kindly provided by R. Treisman (ICRF, London). All cell lines were cultured as described (4). Synchronisation of NIH3T3

* To whom correspondence should be addressed

cells in G_0 transfection by the DEAE dextran technique were carried out as previously published (7,4). Determination of luciferase activity and standardisation of results were performed as described (8,4). Final results were expressed as RLUs/2 × 10^5 transfected recipient cells.

Sequence analysis and luciferase constructs

DNA sequencing was performed by the dideoxynucleotide (9) using Sequenase (USB). Uncertain sequences and GC-rich stretches were verified by 'cycle sequencing' using *Tth* polymerase (Pharmacia). All deletions and mutations were generated by site directed mutagenesis (10) as described (4), and cloned into the promoterless luciferase vector pXP2 (11). The SV40 basal promoter fragment and *cdc25C* UAS fragments were generated by PCR with compatible ends for cloning into pXP2. The SV40 enhancer-*cdc25C* basal promoter construct SV-C20 has previously been described (4). The other SV40 enhancer-*cdc25C* constructs and CMV enhancer-*cdc25C* basal promoter construct were generated in an analogous was using PCR-amplified fragments. The CMV fragment spanned the region from –600 to –41 (12). All cloned PCR-amplified fragments wereverified by DNA sequencing.

Genomic footprinting

Genomic footprinting (13) of WI-38 cells was performed as described (4). The following primers were used:
Experiment in Figure 1A:
1st primer, T_m = 63.6°C, 5′-d(CTGCGTCAGCCAATCTCC)-3′;
2nd primer, T_m = 78.2°C, 5′-d(TGGCCTATCGTTGGGCTCGCAG)-3′;
3rd primer, T_m = 87.1°C, 5′-d(GGGCTCGCAGATCACCTGGGGGCG)-3′.
Experiment in Figure 1B:
1st primer, T_m =56.0°C, 5′-d(AGGGGAAAGGAGGTAGTT)-3′;
2nd primer, T_m = 74.0°C, 5′-d(TAGATTG CAGCTCTGCCTTCCGAC)-3′;
3rd primer, T_m = 83.0°C, 5′-d(CCTTCCGACTGGGTAGGCCAACGTCG)-3′;
Experiment in Figure 1C:
1st primer, T_m = 57.6°C, 5′-d(CACTAGTAAGGCGCGGT)-3′;
2nd primer, T_m = 76.5°C, 5′-d(GTTTAAATCTCCCGGGGTTCGTGG)-3′;
3rd primer, T_m = 86.2°C, 5′-d(GGGGTTCGTGGGGCTGAGGGAACGAG)-3′.

Electrophorectic mobility shift assays

Preparation of nuclear extracts and electrophorectic mobility shift assays (EMSAs) were performed as described (14,15), using 0.7 μg/μl poly(dI:dC) as the non-specific competitor. Oligonucleotide probes were labelled by filling-in of 5′ overhangs using the Klenow fragment of *E.coli* DNA polymerase I in the presence of α-^{32}P deoxynucleoside triphosphates. Protein–DNA complexes were separated on native 5% polyacrylamide gels. Oligonucleotides had the following sequences:
MHC class II promoter NF-Y site: 5′-(GATTTTTCTGATTGGTTAAAAGT) and 5′-(GACTTTTAACCAATCAGAAAAAT); *Y_c-box 1 (CBS1/2):* 5′-(GAGGCGAGCGGGGATAGGTTACTG) and 5′-(AGCCCAGTAACCTATCCCCGCTCG); *Y_c-box 2 (CBS3/4):* 5′-(GGCGCGCGCGGAGATTGGCTGACG) and 5′-(GCTGCGTCAGCCAATCTCCGCGCG); *Y_c-box 3 (CBS5/6):* 5′-(TCTGCGAGCCCAACGATAGGCCAT) and 5′-(GGTCATGGCCTATCGTTGGGCTCG); *mutated Y_c-box 3 (CBSm5/6):* 5′-(TCTGCGAGCCCAACAGTACTCCAT) and 5′-(GGTCATGGAGTACTGTTGGGCTCG); *mutated Y_c-box 3 (CBS5/m6):* 5′-(TCTGCAAACCCAACGATAGGCCAT) and 5′-(GGTCATGGCCTATCGTTGGGTTTG); *CBS5:* 5′-(GCCCAACGATAGGCCATG) and 5′-(GGTCATGGCCTATCGTTG); *CBS6:* 5′-(GGATCTGCGAGCCCAACC) and 5′-(GGGGTTGGGCTCGCAGAT).

The rd oligonucletide corresponds to an unrelated 21 bp sequence of the MMTV-LTR (–1017 to –997) (16).

RESULTS

Identification of *in vivo* protein binding sites in the *cdc25C* upstream sequence

As shown in a previous study (4), a stretch of ~50 nucleotides (nt) upstream of the CDE is essential for cell cycle regulation of the *cdc25C* promoter, and a further ~100 nt are needed for maximal activity and optimal regulation. We therefore analysed the region between position –300 and the CDE by genomic DMS footprinting of both strands for *in vivo* protein binding sites as a basis for further functional and biochemical analyses, using methylated DNA from normally growing, G_0/G_1 (quiescent or sorted) and G_2 (stimulated or sorted) cells. This analysis revealed the presence of nine constitutive protein binding sites (CBS1 to 9) in the coding strand (Figs 1A, 1B and 2A) located around positions –27 to –28 (CBS1), –39 to –44 (CBS2), –58/–59 (CBS3), –70 (CBS4), –91 to –96 (CBS5), –103/–105 (CBS6), –112 to –128 (CBS7), –154 to –161 (CBS8) and –218 to –225 (CBS9). Cell cycle-independent protection of CBS7, 8 and 9 was also clearly seen in the non-coding strand (Figs 1C and 2A) around positions –117 to –125 (CBS7), –152 to –157 (CBS8) and –217 (CBS9), while no protection was seen in the regions of CBS1 to 4. A few partially protected G-residues were detected between CBS8 and CBS9 (Figs 1C and 2A), but are probably of minor importance because of their location within a functionally dispensable region of the sequence (see below). In addition, as expected, the CDE (–12 to –16 in the coding strand, and position –14 in the non-coding strand) was protected in both strands in a cell cycle-regulated manner, i.e., occupied in G_0/G_1 and not protected in G_2.

A comparison of the CBS elements with known transcription factor binding sites showed a perfect match of CBS8 with a classical binding site for Sp1 (17). In agreement with this observation, CBS8 bound Sp1 with high affinity in EMSA experiments using nuclear extract from HeLa or NIH3T3 cells (data not shown). On the other hand, CBS3 was found to resemble a reverse CCAAT-box (GATTG) which is the core motif of a binding site for the transcription factor CBF/NF-Y (18–21). No other similarities with known binding sites were obvious. A closer inspection of the *cdc25C* promoter sequence, however, revealed the presence of three nearly identical sequence repeats (termed Y_c-boxes; see Figs 1B, 2A and B) composed of the protected sites CBS1 and 2 (Y_c1), CBS3 and 4 (Y_c2), and CBS5 and 6 (Y_c3). Strikingly, the protected G-residues are located at very similar positions in the three Y_c-boxes, especially in the CCAAT-box-like element, raising the possibility that similar proteins or protein complexes are bound to these sites. The CCAAT-like motif will subsequently be referred to as the Y_c-box

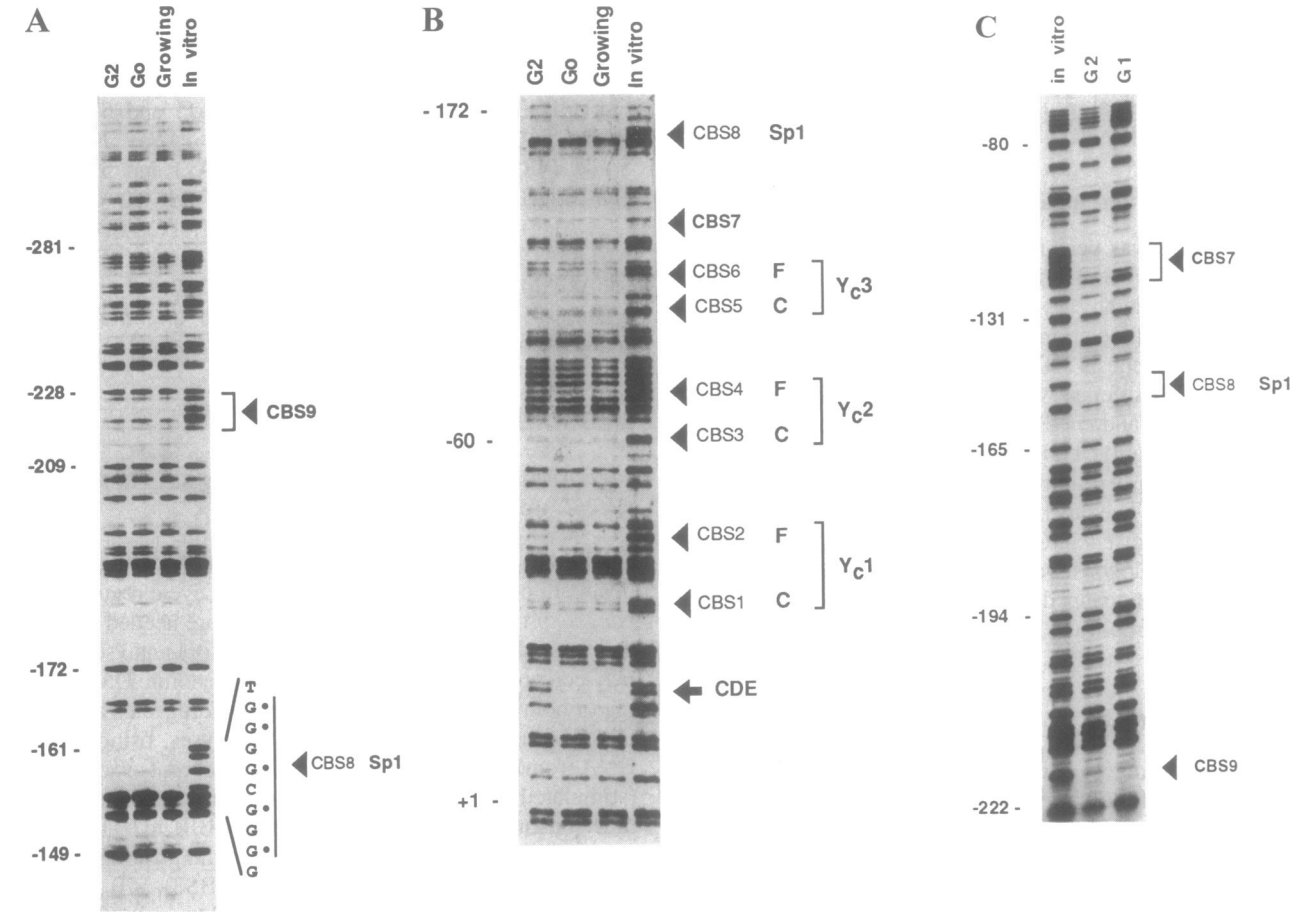

Figure 1. Identification of protein binding sites in WI-38 cells by *in vivo* DMS footprinting of a region spanning nucleotides –300 and +2 in the coding strand (**A** and **B**) and nucleotides –222 and –60 in the non-coding strand (**C**). One G_2-specific binding site (CDE) and nine constitutive binding sites (CBS 1 to 9), which are occupied in both G_0/G_1 and G_2 cells, can be identified. All protected G residues are marked in the sequence shown in Figure 2A. *In vitro*, naked DNA methylated *in vitro*; G_0, quiescent cells obtained by serum deprivation for 3 days; G_2, cells synchronised in G_0 and stimulated with FCS to re-enter the cell cycle (A and B) or obtained by sorting of a normally cycling population.

core (Y_c-C), and the adjacent protected region as the Y_c-box flanking sequence (Y_c-F).

Correlation of *in vivo* protein binding and function

We next asked the question which of the constitutive *in vivo* protein binding sites identified above are important with respect to transcriptional activation and cell cycle regulation. Towards this end, we generated a series of luciferase reporter constructs with 5′-terminal truncations or point mutations in various CBS elements, and tested these in a transient expression assay in both G_0 and growing cells.

CBS9, Sp1 site and CBS7. Almost identical results were obtained with both C290 and C172, with respect to both expression levels and cell cycle regulation (~14-fold), indicating that the region between positions –290 and –172 including CBS9 is of minor importance. In contrast, a further truncation down to –125 removing the Sp1 site reduced the promoter activity in both G_0 and growing cells by ~30%. In agreement with the truncation analysis, mutation of the Sp1 site in C290 also led to a reduction of ~35% in activity, but not in cell cycle regulation (C290mC8 in Fig. 4A). The fact that C290mC8 (which retains an intact CBS9 site; Fig. 3) shows a very similar activity as C125 (which lacks CBS9; Fig. 4A) also supports the conclusion that CBS9 is of minor importance, if any, at least with regard to the functions analysed here. This element was therefore not dealt with any further in the present study. In contrast, truncation of a further 10 nt (C115) deleting most of CBS7 led to a reduction in activity in both G_0 and growing cells by ~37% (Fig. 3), indicating that this element is relevant with respect to both basal activity and cell cycle regulation.

Y_c-box 3. Truncation of Y_c3-F in construct C98 resulted in a drop of ~30% with respect to both basal activity and regulation (Fig. 3), while the subsequent deletion of Y_c3-C in C75 had no noticeable effect (Fig. 3). On the other hand, mutation of Y_c3-C in construct C290 led to an ~2-fold reduction in both activity in growing cells and regulation (Fig. 4A). To address this apparent paradox we generated further C5 mutants in (Y_c3-C) in the background of promoter constructs differing in their 5′ termini, and analysed these in growing cells (Fig. 4C). In these experiments, the C5 mutation led to the expected ~50% reduction in the C290, C125 and C115 backgrounds, but not in C98. This suggests that the function of Y_c3-C is dependent on the presence of Y_c3-F, i.e., an intact Y_c-box 3, and thus offers an explanation for the seemingly paradoxical result obtained in the terminal deletion analysis.

Y_c-boxes 1 and 2. Further truncations led to a progressive decline in both activity and cell cycle regulation. Truncation of Y_c1-F or

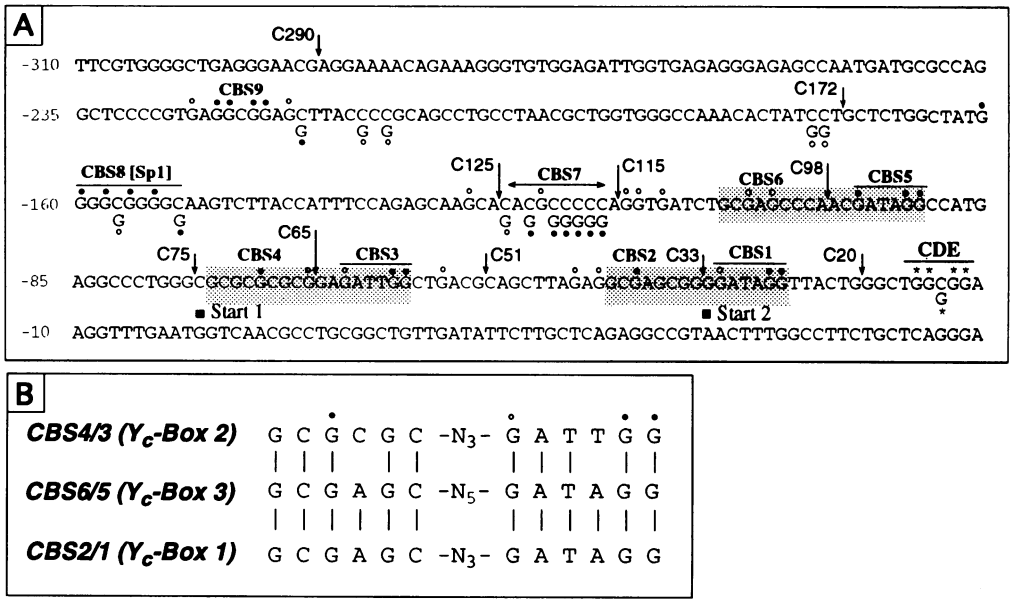

Figure 2. (**A**) Nucleotide sequence of the *cdc25C* upstream region. Protected G residues detected by *in vivo* footprinting (see Fig. 1) are marked by filled (●) and open (○) circles to denote strong and partial constitutive protection, respectively. Cell cycle-regulated protein binding to the CDE is indicated by asterisks. Y_c-boxes 1, 2 and 3 are shaded, Y_c-box 1 being the most downstream one. The CDE-proximal site of transcription initiation is marked by a solid square. Arrows show the 5′ end points of the deletion constructs used in subsequent figures. (**B**) Alignment of Y_c-boxes 1, 2 and 3. Filled and open circles indicate G residues that show strong (●) or partial (○) protection in all three sequences.

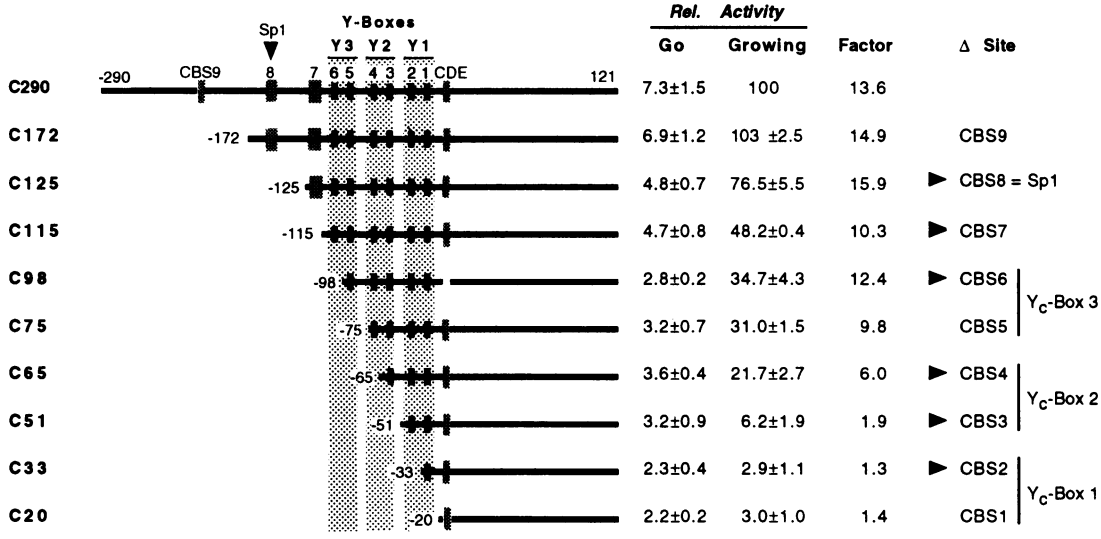

Figure 3. Transient expression analysis of terminally truncated *cdc25C* promoter–luciferase constructs in quiescent (G_0) versus growing NIH3T3 cells. Plasmids were named to indicate the 5′-truncation (see also Fig. 2A). All plasmids harbour a 121 bp region downstream of the first initiation site. Mean normalised values (C290 in growing cells = 100%) of three independent experiments and standard deviations are given. Factor is the ratio of the values in growing and G_0 cells. Δ Site indicates which site was deleted from a given construct with respect to the preceding one (one line above). Arrow heads point to those sites whose deletion led to a significant drop in activity (≥30%). Black boxes, core (C) and flanking (F) elements in Y_c-boxes; grey boxes, other elements.

Y_c2-F led to a reduction in activity of 53% and 30%, respectively, and a dramatic effect was seen upon removal of Y_c2-C, which resulted in a reduction of 71% (Fig. 3). Similarly, cell cycle regulation was severely affected and basically impaired after deletion of Y_c2-C (Fig. 3). Very similar observations were made upon mutation of Y_c1-C or Y_c2-C in construct C75 (Fig. 4B), which resulted in an ~75% decreased activity in growing cells and a severe loss of regulation (~79%). In addition, a clear, albeit less severe effect was seen upon mutating Y_c2-F (35% reduction in overall activity). This data point to Y_c-boxes 1 and 2 as functionally crucial sites in the *cdc25C* promoter that synergize in transcriptional activation. Mutation of Y_c1-C or Y_c2-C in construct C290 also led to a significant reduction in both activity in growing cells and regulation, but the effect was less severe than

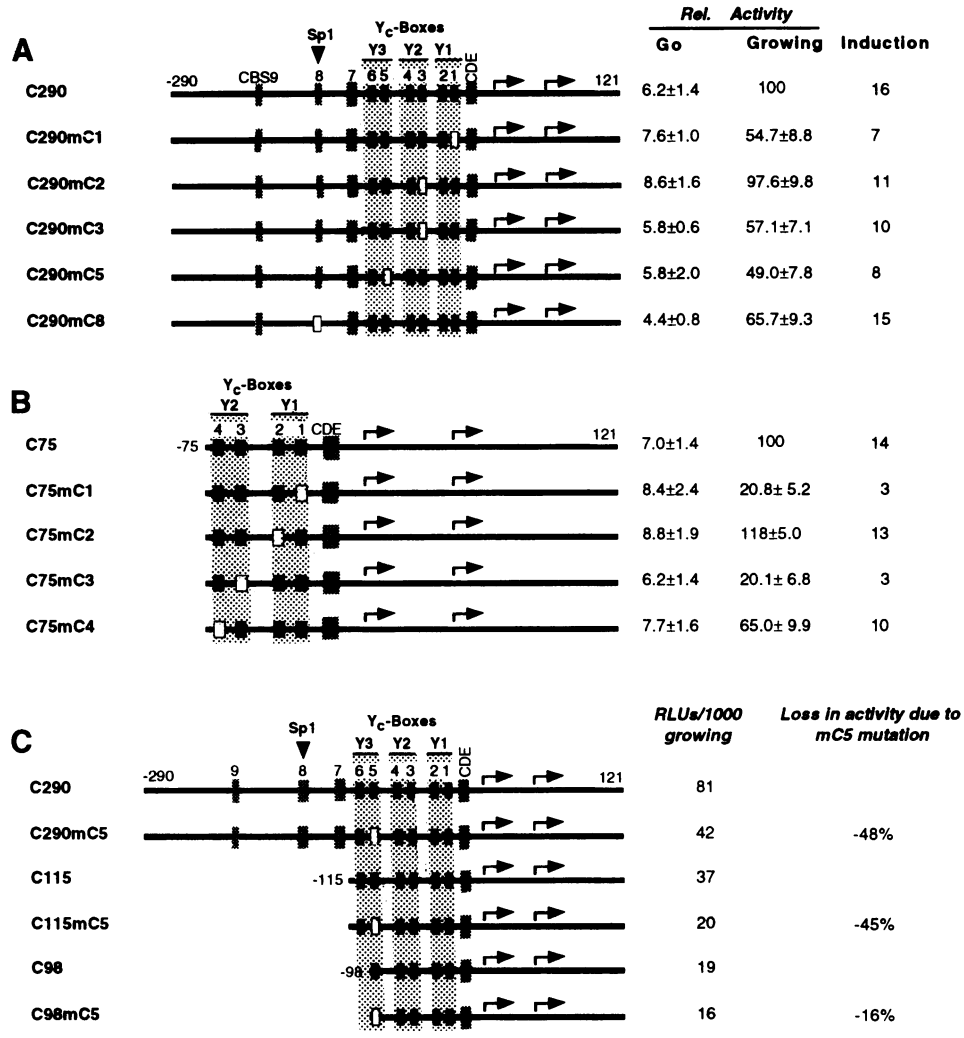

Figure 4. Transient expression analysis in quiescent (G_0) and growing NIH3T3 cells of *cdc25C* promoter–luciferase constructs harbouring specific mutations in CBS elements. Black boxes, core (C) and flanking (F) elements in Y_c-boxes; grey boxes, other elements; open boxes, mutated elements. The analysis and evaluation was performed as in Figure 3, except that in panel C absolute RLUs rather than normalised values are shown.

in the C75 background (2- versus 4-fold), which is presumably due to the presence of additional activator elements, including another Y_c-box (Y_c3) in C290.

In summary, the genomic footprinting and functional studies have identified Y_c-boxes 1, 2 and 3, CBS7 and the Sp1 site as activator binding sites involved in transcriptional activation. Within the Y_c-boxes, both core motifs and the 5′ flanking regions are functionally relevant, with a relatively relaxed sequence requirement for the latter. The truncation and mutation analyses also point to a correlation between activity in growing cells, i.e., strength of the UAS and the magnitude of cell cycle regulation, suggesting a functional connection between activation and CDE-mediated repression. Our results also indicate that Y_c-boxes 1 and 2 are both necessary and sufficient for high levels of expression and efficient cell cycle regulation.

Delineation of separable, context-independent activator modules

We next asked the question which regions of the *cdc25C* upstream sequence might act as separable activator modules, which elements might be involved in functional cooperations (apart from the synergism of Y_c-boxes described above), and whether any of the activator elements might be cell cycle-regulated itself. Towards this end, various fragments of the *cdc25C* sequence were linked to a basal SV40 early promoter-luciferase reporter and tested in a transient transfection assay (Fig. 5). As expected, the parental SV-TATA construct showed very little activity in both G_0 and growing cells. In contrast, fusions with *cdc25C* fragments harbouring either Y_c-boxes 2 and 1 (Y_c2/Y_c1-SV), CBS7 and Y_c-box 3 (CBS7/Y_c3-SV) or CBS9 and the Sp1 site (CBS9/Sp1-SV) showed a clearly elevated activity (8- to 9-fold relative to SV-TATA). In contrast, when tested individually Y_c1 and 3, as well as CBS7, showed only poor transactivation. This confirms our findings of the deletion and mutation analyses (Figs 3 and 4) which indicated a cooperation among Y_c1 and 2. The SV40 hybrid promoter analysis extends these observations by demonstrating that the cooperation of Y_c-boxes is not promoter-specific. In addition, this analysis shows that there is another kind of cooperation of the Y_c-box binding factor(s), as suggested by the synergism of Y_c3 with CBS7. Finally, the data in Figure 5

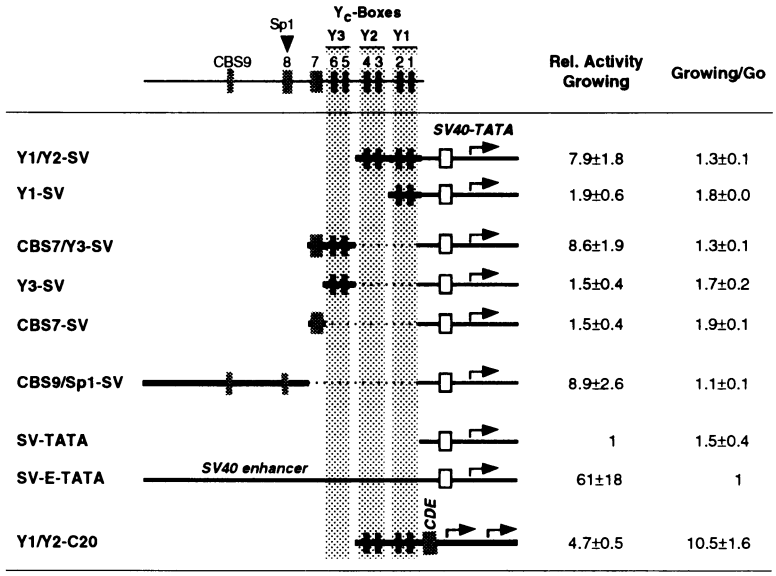

Figure 5. Transient expression analysis in quiescent (G_0) and growing NIH3T3 cells of *cdc25C*-SV40 basal promoter–luciferase constructs. *cdc25C* sequences are shown as bold lines, SV40 sequences as thin lines. Dotted lines indicate deleted sequences. Black boxes, core (C) and flanking (F) elements in Y_c-boxes; grey boxes, other elements; open boxes, SV40 TATA-box. The analysis and evaluation was performed as in Figure 3, except that values are based on four independent experiments. Activities are given relative to SV-TATA (normalised to 1); ratios of activities in growing/G_0 cells were calculated relative to SV-E-TATA (normalised to 1).

clearly indicate that none of the activator modules was able to confer a significant cell cycle regulation (≤2-fold) on the SV40 basal promoter. These results suggest that the *cdc25C* UAS is not a direct target of cell cycle regulation, thus pointing to a central role of CDE-mediated repression.

A central role for CDE-mediated repression in cell cycle-regulation of the *cdc25C* UAS

To test in a direct way the hypothesis that the CDE is the principal regulatory element in the *cdc25C* gene with respect to cell cycle regulation, we linked the activator modules identified above to the *cdc25C* basal promoter in the presence of a wild-type or a point-mutated CDE (G→T transversion at position –13; referred to as mutant RT7). CDE-dependent cell cycle regulation had previously been shown only for a short UAS fragment harbouring Y_c1 and 2 (4). The data in Figure 6 confirm the results of the chimaeric SV40-promoter study in Figure 5, in that the three tested regions (Y_c2/Y_c1, CBS7/Y_c3 and CBS9/Sp1) showed clear activation in the context of the *cdc25C* basal promoter (10- to 40-fold enhancement in growing cells relative to C20). The CBS7/Y_c3 domain, however, showed a clearly decreased activity in the C20 context as compared to the two other UAS regions tested (Y_c2/Y_c1 and CBS9/Sp1; Fig. 6), which is different to the corresponding chimaeric SV40 promoter context (Fig. 5). It is likely that this result points to differences in the mechanism by which the *cdc25C* and SV40 basal promoters are activated.

The data in Figure 6 also show that the Y_c2/Y_c1 module was strongly cell cycle regulated (16.1-fold), and that this regulation was largely dependent on the presence of an intact CDE. Likewise, both the CBS7/Y_c3 domain and the CBS9/Sp1 site-containing fragment were cell cycle-regulated (~6-fold) and this regulation was clearly CDE-dependent. Most significantly, the entire UAS in C290, including all the protein binding sites identified in the present study (CBS1 to 9), was found to be very efficiently repressed in G_0 by the CDE, as indicated by the 10.4-fold increased expression of construct C290RT7 relative to C290.

Identification of NF-Y as a Y_c-box binding protein

The work described above led to the identification of the Y_c-boxes as major elements involved in the periodic, CDE-controlled transcription of the *cdc25C* gene. In order to elucidate the mechanism of CDE-mediated repression it is of paramount importance to identify the proteins that interact with these elements. As mentioned above, the core element within Y_c-box 2 (Y_c2-C) closely resembles an inverse CCAAT-box, and shows a high degree of similarity with the core motifs in Y_c-boxes 1 and 3 (Y_c1-C and Y_c3-C), respectively (see Fig. 2B). A number of different CAAT-box binding factors, including NF-Y (also termed CBF or CP-1) (18–21), various C/EBP isoforms (22,23) and NF-I (also known as CTF) (24,25) have been described, but Y_c2-C shows the greatest similarity with the consensus NF-Y binding site (CCAAT) (18,20). In order to test the hypothesis that NF-Y interacts with the Y_c-boxes within the *cdc25C* UAS we performed electrophoretic mobility shift assays (EMSAs) using synthetic oligonucleotides encompassing Y_c1, Y_c2 or Y_c3, or a *bona fide* NF-Y site from the MHC class II promoter (26), with nuclear extracts from normally cycling HeLa cells.

As shown in Figure 7A, a complex of similar mobility was detected with all four oligonucleotide probes. Complex formation was greatly reduced by self-competition, but not affected by a competitor of random sequence (data not shown). Protein binding to the Y_c-box oligonucleotides was also efficiently competed by the Eα-Y site, and efficient cross-competition was seen when either Y_c1 or Y_c3 were used as the probe and Y_c2 as the competitor (Fig. 7A). In contrast, no interaction was seen in the same assay with *bona fide* binding sites for the CCAAT-box binding factors NF-I, and only weak binding was seen with

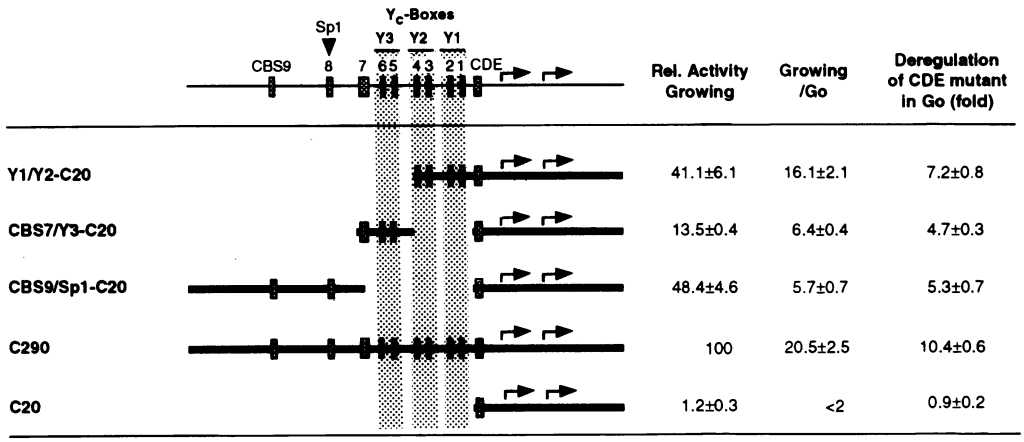

Figure 6. Transient expression analysis in quiescent (G_0) and growing NIH3T3 cells of various *cdc25C* UAS fragments linked to the C20 basal promoter construct with either a wild-type or a point-mutated (RT7 constructs; G→T transversion at –13) CDE. Black boxes, core (C) and flanking (F) elements in Y_c-boxes; grey boxes, other elements. The analysis and evaluation was performed as in Figure 5. The right-most column indicated the extent of deregulation seen in the RT7 constructs in G_0 cells (ratio of RT7/wt activities).

C/EBP binding site (data not shown). This is exactly what would be expected for a NF-Y site, since NF-Y also binds with reduced affinity to C/EBP sites (27). These results suggest that the same proteins interact with all three Y_c-boxes and the Eα-Y NF-Y site. Only in the case of Y_c2 two additional, minor complexes of higher mobility were detected, and these were competed less efficiently by the Eα-Y site (Fig. 7A and data not shown). The other minor bands seen in Figure 7 represent unspecific complexes (data not shown).

To analyse the contribution of the core element and the flanking region within the Y_c-boxes to the binding of NF-Y we used probes with mutations or deletions in the respective motifs within Y_c3. The data shown in Figure 7B clearly indicate that the core motif is absolutely essential for the interaction with NF-Y, since both its mutation (Y_c3-mC) and its deletion (Y_c3–ΔC) abolished binding. The flanking region behaved differently, in that its truncation (Y_c3-ΔF) led to a loss of binding, but not its mutation (Y_c3-nFl). This observation indicates that the presence of the core element is not sufficient for binding of NF-Y, which requires an additional upstream sequence of relaxed specificity. Very similar observations were made with the two other Y_c-boxes (data not shown). The results of these binding studies are in good agreement with the functional data (Figs 3 and 4) which also showed that not only the core motif but also the 5′ flanking sequence within the Y_c-boxes was important, but that the requirement for a specific sequence of the flanking region seemed to be rather relaxed.

In order to obtain further evidence in support of the hypothesis that all Y_c-boxes bind the same protein complex as the *bona fide* NF-Y site we performed supershift assays using a polyclonal antibody specific for the B-subunit of NF-Y (kindly provided by D. Mathis, Strasbourg) (26). The results with each of the four oligonucleotide probes were very similar: complete shift with the polyclonal antibody and no change with non-immune IgG (Fig. 8). Furthermore, a monoclonal antibody (MAb) against the A-subunit of NF-Y (also provided by D. Mathis) also supershifted the complexes observed with Y_c-boxes 1, 2 and 3 in the same way as the complex formed with the *bona fide* MHC II NF-Y site (data not shown). Taken together, the data obtained by competition, mutation/deletion and antibody supershift analyses clearly suggest that the protein complex interacting with Y_c-boxes 1, 2 and 3 is NF-Y.

DISCUSSION

We have previously shown that the cell cycle-regulated transcription of the human *cdc25C* gene is dependent on a repressor element (CDE) located in the region of the basal promoter (4). In addition, it has been shown that a region of ~60 bp upstream of the CDE was indispensable for both efficient transcription and cell cycle regulation, and this was further improved by the presence of another ~100 bp of upstream sequence. The region 5′ to the CDE sequences could be replaced by the heterologous SV40 early enhancer region, suggesting that the *cdc25C* sequences 5′ to the CDE behave as an UAS whose function in transcriptional activation is blocked by a CDE-binding protein or protein complex. We now provide strong evidence that this is indeed the case. The *cdc25C* sequences upstream of the CDE are shown to harbour multiple *in vivo* protein binding sites that interact with constitutive transcriptional activators, which in turn are repressed through a CDE-directed mechanism.

Three Y_c-boxes and a Sp1 site are the principal elements in the *cdc25C* UAS

The genomic footprinting analysis performed in the present study led to the identification of nine constitutive protein binding sites, CBS1 to CBS9, within a region of ~200 bp upstream of the CDE (Figs 1 and 2). The site farthest upstream, CBS9, did not appear to be functionally important in the assays used, and was therefore not analysed further. CBS8 represents a classical Sp1 site(17), and was found to be functionally relevant, since its truncation or mutation led to an ~30% drop in activity (Figs 3 and 4A). The CBS7 site, consisting mainly of an oligomeric (dC:dG) stretch, was found to be of similar importance as the Sp1 site (Fig. 3). This site gave a particularly strong and extensive *in vivo* footprint (in the non-coding strand), but the nature of the protein(s) binding to this site in nuclear extracts or *in vivo* is unknown. The three

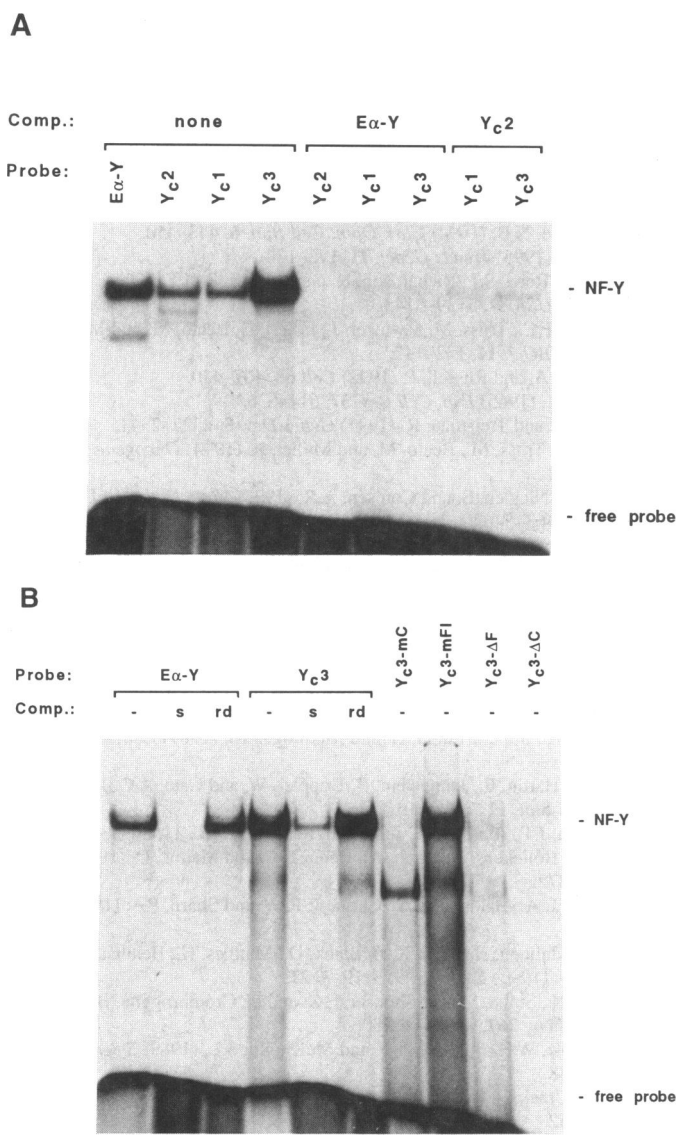

Figure 7. Electrophoretic mobility shift analysis (EMSA) of protein complexes interacting with Y_c-boxes 1, 2 and 3 or a *bona fide* NY-1 site (Eα-Y) in the absence or presence of competitors (~100-fold excess over probe). Rd, random sequence. (**A**) Competition experiments. (**B**) EMSA of protein complexes interacting with probes containing the complete Y_c-box 3 in either the wild-type configuration or with mutations in either the core motif (probe Y_c3-mC) or flanking region (probe Y_c3-mFl), or with probes with deletions of either the core motif (probe Y_c3-ΔC) or flanking region (probe Y_c3-ΔF). s, self-competition; rd, competition with a random oligonucleotide (unrelated MMTV sequence).

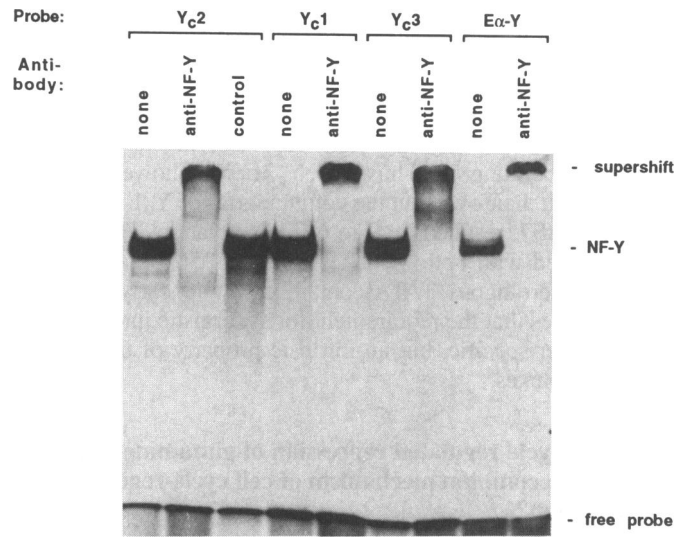

Figure 8. Antibody supershift analysis of protein complexes interacting with Y_c-boxes 1, 2 and 3 or a *bona fide* NY-1 site (Eα-Y) using a polyclonal antibody raised against the B subunit of NF-Y (IgG fraction); control, irrelevant polyclonal antibody (IgG).

Y_c-boxes found in the *cdc25C* UAS are bipartite and are each composed of a reverse CCAAT-box-like element, the core (Y_c-C), and a 5′ flanking G/C-rich sequence (Y_c-F). The Y_c-boxes were found to be of particular interest for the following reasons: (i) they are functionally important (Figs 3 and 4), and in the cell cycle-regulated C75 construct Y_c1 and Y_c2 represent the only *in vivo* protein binding sites and activator elements upstream of the CDE (Fig. 4B); (ii) the Y_c-boxes synergise in transcriptional activation, either with each other or with CBS7 (Figs 4B and 5); (iii) they represent binding sites for the CCAAT-box binding factor NF-Y [Figs 7 and 8; (18,20), also referred to as CBF (21)], but these binding sites are unusual in that they deviate from the proposed consensus sequence (18,20).

The Y_c-boxes are unusual NF-Y binding sites

The functionally most crucial sequence within each of the Y_c-boxes is the core element, since mutations within this region impair its function. In contrast, the 5′ flanking regions showed a very relaxed requirement for a specific sequence with respect to both function and binding of NF-Y. Thus, the truncation of the 5′ flanking region led to a significant drop in activity, while its mutation did not. These observations suggest that NF-Y binding to Y_c-boxes not only requires the presence of the core motif, but also a stretch of DNA upstream of the reverse CCAAT-box. The lack of a stringent sequence requirement within this region indicates that it may not be involved in specific protein–DNA interactions but may rather serve to stabilise the binding to the CCAAT-like box. It is possible that this requirement for the presence of additional sequences is due to the deviation of the Y_c-boxes from the consensus NF-Y site in both the core region (for Y_c1 and Y_c3; see Fig. 2B) and the flanking nucleotides, which leads to a significant loss in binding affinity (18).As a consequence of this decreased affinity, the interaction between NF-Y and the Y_c-boxes may be dependent on extended contacts. In fact, binding to a consensus NF-Y site has been observed with only nine nucleotides upstream from the reverse CCAAT-box (18), and is thus not dependent on further upstream sequences as in the case of the Y_c-boxes.

A situation that is very similar to the observations made in the present study has actually been described in a recent study of the minute virus of mice P4 promoter (28). In this promoter, a NF-Y site with a CCAAC instead of the consensus CCAAT core has been identified. Binding of NF-Y to this element is, as in the case of the Y_c-boxes, critically dependent on the presence of a 5′ located stretch of DNA. In addition the yeast homologue of NF-Y, HAP2/3/4/5, has also been reported to interact with a reverse CCAAT-box and a GCTC-box located upstream (29). Very

recently, the C-subunit of NF-Y/CBP has been cloned (30) so that recombinant holo-NF-Y is now available and can be used to address some of the remaining questions.

The Y_c-boxes are also unusual in view of the fact that they do not function efficiently as single copies or in the absence of another cooperating site. Thus, Y_c1, Y_c2 or Y_c3 showed only little activation on their own, but the combinations of Y_c1 and Y_c2, or Y_c3 and CBS7, resulted in 4- to 6-fold activation. This synergy was observed with both the TATA-less *cdc25C* basal promoter and the heterologous TATA-containing SV40 early promoter. This indicates that the requirement for synergistic interactions is not promoter-specific, but an intrinsic property of the Y_c-box-NF-Y complexes.

Is the cell cycle-regulated repression of glutamine-rich activators a common mechanism of cell cycle-regulated transcription?

The proteins interacting with the *cdc25C* UAS not only bind constitutively (Fig. 1), but also activate transcription in a way that is not significantly regulated by the cell cycle (Fig. 5). The NF-Y complexes interacting with the *cdc25C* UAS, as well as Sp1, rather represent targets for the cell cycle-regulated repression by the CDE. Interestingly, the major activation domains in both Sp1 (31) and NF-Y (32,33) are glutamine-rich, and both factors are therefore likely to contact a similar set of basal transcription factors, TAFs or other components of the preinitiation complex (34). It will certainly be important to consider the question as to whether this repression mechanism is restricted to certain classes of activation domains, and thus activator-basal complex contacts. In support of such a hypothesis is the observation that the heterologous SV40 enhancer is less efficiently repressed than the *cdc25C* UAS, and that the CMV enhancer is hardly affected by the CDE (our unpublished observations). On the other hand, it is unlikely that transcription factors with glutamine-rich activation domains are the only targets of CDE-mediated repression, since different classes of activators may contact overlapping sets of basal factors and TAFs (34).

A final point relates to our observation that the promoters of various other cell cycle genes, such as the cdc2 (35) and cyclin A genes (36), also contain CDE-like elements in close proximity to their transcription initiation sites, and recently performed *in vivo* footprinting experiments have indeed indicated cell cycle-regulated protein binding to these elements (37). Interestingly, these genes also harbour multiple Sp1 sites and reverse CCAAT-boxes upstream of the CDE-like elements (35,36), pointing to a similar mechanism of CDE-regulated activation. In addition, cell cycle genes repressed by a distinct, but related mechanism through the transcription factor E2F also show a conspicuous preference for Sp1 sites and CCAAT-boxes 5′ to the E2F sites, as for example in the *E2F-1* gene (38,39). It is therefore likely that the molecular basis for E2F- and CDE-mediated negative regulation is very similar and that the repression of glutamine-rich activators like Sp1 and the CCAAT-box binding factor NF-Y is a common mechanism of cell cycle-regulated transcription.

ACKNOWLEDGEMENTS

We are grateful to R. Treisman for NIH3T3 cells, to D. Mathis for NF-Y antibodies, to Dr M. Krause for synthesis of oligonucleotides and to C. Schalk for FACS analyses. This work was supported in part by the Stiftung P. E. Kempkes and the DFG (Mu601/7-1). C.G. and F.E. were recipients of fellowships from the Graduiertenkolleg 'Zell- und Tumorbiologie'. J.Z. was supported by a fellowship from the Boehringer Ingelheim Fonds.

REFERENCES

1 La Thangue, N.B. (1994) *Curr. Opin. Cell Biol.* **6**, 443–450.
2 Müller, R. (1995) *Trends Genet.* **11**, 173–178.
3 Sadhu, K., Reed, S.I., Richardson, H. and Russell, P. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5139–5143.
4 Lucibello, F.C., Truss, M., Zwicker, J., Ehlert, F., Beato, M. and Müller, R. (1995) *EMBO J.* **14**, 132–142.
5 Millar, J.B.A. and Russell, P. (1992) *Cell* **68**, 407–410.
6 Hayflick, L. (1965) *Exp. Cell Res.* **37**, 614–636.
7 Pollock, R. and Treisman, R. (1991) *Genes Dev.* **5**, 2327–2341.
8 Herber, B., Truss, M., Beato, M. and Müller, R. (1994) *Oncogene* **9**, 1295–1304.
9 Sanger, F., Nicklen, S. and Coulsen, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
10 Good, L. and Nazar, R. (1992) *Nucleic Acids Res.* **20**, 4934.
11 Nordeen, S.K. (1988) *BioTechniques* **6**, 454–457.
12 Boshart, M., Weber, F., Jahn, G., Dorsch-Häsler, K., Fleckenstein, B. and Schaffner, W. (1985) *Cell* **41**, 521–530.
13 Pfeifer, G.P., Steigerwald, S., Mueller, P.R. and Riggs, A.D. (1989) *Science* **246**, 810–813.
14 Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
15 Barberis, A., Superti-Furga, G. and Busslinger, M. (1987) *Cell* **50**, 347–359.
16 Mink, S., Härtig, E., Jennewein, P., Doppler, W. and Cato, A.C.B. (1992) *Mol. Cell. Biol.* **12**, 4906–4918.
17 Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) *TIBS* **11**, 20–23.
18 Dorn, A., Bollekens, J., Staub, A., Benoist, C. and Mathis, D. (1987) *Cell* **50**, 863–872.
19 Chodosh, L.A., Baldwin, A.S., Carthew, R.W. and Sharp, P.A. (1988) *Cell* **53**, 11–24.
20 van Hujisduijnen, H.R., Li, X.Y., Black, D., Matthes, H., Benoist, C. and Mathis, D. (1990) *EMBO J.* **9**, 3119–3127.
21 Maity, S.N., Sinha, S., Ruteshouser, E.C. and de Crombrugghe, B. (1992) *J. Biol. Chem.* **267**, 16574–16580.
22 Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1989) *Science* **243**, 1681–1688.
23 Cao, Z., Umek, R.M. and McKnight, S.L. (1991) *Genes Dev.* **5**, 1538–1552.
24 Santoro, C., Mermod, N., Andrews, P.C. and Tjian, R. (1988) *Nature* **334**, 218–224.
25 Meisterernst, M., Gander, I., Rogge, L. and Winnacker, E.L. (1988) *Nucleic Acids Res.* **16**, 4419–4435.
26 Mantovani, R., Pessara, U., Tronche, F., Li, X.Y., Knapp, A.M., Pasquali, J.L., Benoist, C. and Mathis, D. (1992) *EMBO J.* **11**, 3315–22.
27 Wuarin, J., Mueller, C. and Schibler, U. (1990) *J. Mol. Biol.* **214**, 865–874.
28 Gu, Z., Plaza, S., Perros, M., Cziepluch, C., Rommelaere, J. and Cornelis, J.J. (1995) *J. Virol.* **69**, 239–246.
29 Guarente, L. (1992) *Transcriptional Regulation*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 1007–1036.
30 Sinha, S., Maity, S.N., Lu, J. and de Crombrugghe, B. (1995) *J. Biol. Chem.* **92**, 1624–1628.
31 Courey, A.J. and Tjian, R. (1988) *Cell* **55**, 887–898.
32 Li, X.-Y., van Hujisduijnen, H.R., Mantovani, R., Benoist, C. and Mathis, D. (1992) *J. Biol. Chem.* **267**, 8984–8990.
33 Coustry, F., Maity, S.N. and de Crombrugghe, B. (1995) *J. Biol. Chem.* **270**, 468–475.
34 Goodrich, J.A. and Tjian, R. (1994) *Curr. Opin. Cell Biol.* **6**, 403–409.
35 Dalton, S. (1992) *EMBO J.* **11**, 1797–1804.
36 Henglein, B., Chenivesse, X., Wang, J., Eick, D. and Brechot, C. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5490–5494.
37 Zwicker, J., Lucibello, F.C., Wolfralm, L.A., Gross, C., Truss, M., Engeland, K. and Müller, R. (1995) *EMBO J.* **14**, 4514–4522.
38 Johnson, D.G., Ohtani, K. and Nevins, J.R. (1994) *Genes Dev.* **8**, 1514–1525.
39 Hsiao, K.-M., McMahon, S.L. and Farnham, P.J. (1994) *Genes. Dev.* **8**, 1526–1537.