CAD, a c-Myc Target Gene, is Not Deregulated in Burkitt's Lymphoma Cell Lines

Susanna M. Mac and Peggy J. Farnham*

Department of Oncology, McArdle Laboratory for Cancer Research, University of Wisconsin–Madison, Madison, Wisconsin

Although the Myc family of transcription factors is upregulated in many human tumors, it is unclear which genes are targets for the deregulated Myc. Previous studies suggest that hamster and rat carbamoyl phosphate synthase, aspartate transcarbamylase, dihydroorotase *Cad* genes are regulated by c-Myc. In fact, of all putative target genes thought to be activated by c-Myc, only the *Cad* gene showed loss of growth regulation in rat cells nullizygous for c-Myc. However, it was unknown whether upregulation of CAD, which performs the first three rate-limiting steps of pyrimidine biosynthesis, contributes to c-Myc's role in human neoplasia. To explore this possibility, we cloned the human *cad* promoter. We found that c-Myc could bind to an E box in the human *cad* promoter in gel shift assays and that growth regulated transcription from the human *cad* promoter was dependent on this c-Myc binding site. However, the increased amount of c-Myc found in Burkitt's lymphoma cell lines did not lead to increased *cad* mRNA levels. Thus, we suggest that although c-Myc is clearly important for the normal transcriptional control of the *cad* promoter, it is unlikely that increased levels of CAD are important mediators of c-Myc–induced neoplasia. Therefore, an understanding of the mechanism by which overexpressed c-Myc contributes to the development of Burkitt's lymphoma requires the identification of additional c-Myc target genes. *Mol. Carcinog. 27:84–96, 2000.*

Key words: neoplasia; transcriptional activation; E box

INTRODUCTION

The proto-oncogene c-*myc* encodes an essential nuclear phosphoprotein [1]. Experimental overexpression of c-Myc can transform primary rat embryo fibroblasts in cooperation with Ras and predispose transgenic mice to various cancers [2]. In humans, the c-*myc* gene is amplified in cancers of the lung, breast, and colon [3]. In Burkitt's lymphoma, the c-*myc* gene is translocated to an immunoglobulin locus, resulting in higher levels of the c-Myc protein and/or mutated c-Myc proteins [4]. It has been postulated that increased c-Myc levels may contribute to neoplasia by maintaining the cycling of cells that would normally exist in a differentiated state, allowing more time for additional mutations to occur [4].

The c-Myc protein is a transcriptional regulator that requires dimerization with a protein partner, max [5]. c-Myc's carboxyl terminus contains a basic helix-loop-helix leucine zipper (bHLHZip) domain that mediates DNA binding and heterodimerization, while the amino terminus contains a transactivation domain and mediates interactions with proteins such as p107 and TATA binding protein [6]. c-Myc/max dimers bind to, and increase transcriptional activity of, promoters containing E boxes (consensus sequence = CACGTG). These heterodimers must compete with max homodimers for binding to E boxes, as well as with Max paired with the Mad family of proteins (Mad, Mxi1, Mad3, Mad4) [7–10]. In contrast to c-Myc/max heterodimers, the Max complexes lacking c-Myc repress transcription [7–10]. Additionally, other ubiquitous homo- and heterodimeric bHLHZip proteins, such as USF, can bind E boxes [11]. These proteins activate some, but not all, of the E box containing promoters that are activated by c-Myc/Max [12,13]. Finally, evidence exists that c-Myc can repress transcription in an E box–independent manner through protein-protein interactions at elements such as initiators [14,15].

Since deregulation of c-Myc activity can cause neoplastic transformation, recent efforts have focused on identifying potential target genes that encode proteins which may contribute to a neoplastic phenotype. Although E boxes that contribute to transcriptional activity have been identified in a variety of promoters, a biological connection between c-Myc and many of the potential target genes is unclear. Recent studies using c-Myc–null cell lines suggest that absence of

*Correspondence to: 1400 University Ave., McArdle Laboratory, University of Wisconsin Medical School, Madison, WI 53706. E-mail: farnham@oncology.wisc.edu

Received 25 June 1999; Revised 10 September 1999; Accepted 17 September 1999

Abbreviations: bHLHZip, basic helix-loop-helix leucine zipper; CAD, carbamoyl phosphate synthetase, aspartate transcarbamylase, dihydroorotase; SSC, sodium chloride sodium citrate; SDS, sodium dodecyl sulfate; PCR, polymerse chain reaction; TBE, Tris-borate-EDTA.

c-Myc does not affect the regulation of most of the putative c-Myc target genes [16]. These observations may suggest that c-Myc's role in mediating neoplastic transformation may not be as a transcriptional activator. This hypothesis is supported by the findings that a c-Myc protein that has lost the ability to repress transcription, but still retains the ability to activate a synthetic promoter, has also lost the ability to transform cells [15]. However, the activation of one gene which was previously postulated to be regulated by c-Myc was affected in the c-Myc-null cells. The growth–regulated increase in mRNA for carbamoyl phosphate synthetase, aspartate transcarbamylase, dihydroorotase (CAD), a trifunctional protein that is rate-limiting for pyrimidine biosynthesis, was abolished in the c-Myc–null cells [16]. Thus, c-Myc does appear to be a critical regulator of the rat *cad* gene. The data from the rat cells are supported by our previous work that showed that a c-Myc binding site is required for G_1/S-phase induction of transcription from the hamster *cad* promoter [17]. Also, a dominant negative c-Myc, that heterodimerizes with Max but doesn't bind DNA can abolish the growth regulation of the hamster *cad* promoter, supporting the proposal that c-Myc/Max heterodimers are responsible for the G_1/S phase–specific increase in the transcriptional activity of the *cad* gene [17].

Based on the studies of the rodent *cad* promoters, it was possible that deregulation of human CAD may contribute to the ability of c-Myc to cause cancer. However, thus far, only rodent *cad* promoters have been characterized. To further investigate a possible role for CAD in human cancers, we cloned the human *cad* promoter and found it to be highly homologous to the hamster sequences. Results of protein binding experiments and transcriptional assays supported the hypothesis that the human *cad* gene is a c-Myc target gene. In contrast, our results also indicated that increased levels of CAD may not contribute to c-Myc–mediated neoplasia in Burkitt's lymphomas.

MATERIALS AND METHODS

Southern Blot Analysis

Cosmid DNA (gift of George Stark) was digested to completion by BglII, EcoRI, or HindIII and electrophoresed on an agarose gel. DNA fragments from this gel were transferred to a Zeta-probe filter in 0.4 M NaOH buffer. After transfer, the filter was rinsed for 10 min in 2× sodium chloride sodium citrate (SSC) at room temperature and then incubated at 42°C in a prehybridization solution (50% formamide, 0.12 M Na_2HPO_4, 0.25 M NaCl, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA) for 15–30 min at 42°C before the addition of a radioactive probe containing hamster *cad* promoter sequences (spanning from nt −333 to +222). Incubation was continued for 4 h to overnight at 42°C on a mechanical rocker. The filter was then rinsed for 10 min each at room temperature in 2 × SSC/0.1% SDS, 0.5 × SSC/0.1% SDS, 0.1 × SSC/0.1% SDS, and 0.1 × SSC/0.1% SDS at 65°C, before exposure to film.

Plasmids

The cosmid cHU CAD69, containing 31 kb of human *cad* genomic DNA, was previously isolated by screening a human leukocyte DNA library with the hamster *cad* cDNA, pUC CAD 142. hCAD2.0, a derivative of cHU CAD69, contains a 2.0-kb BglII fragment that corresponds to the human *cad* promoter. Human *cad* promoter constructs hCAD+25, hCAD+85, and hCAD+115, containing 0, 1, and 2 E boxes, respectively, were created by using hCAD2.0 as template for polymerase chain reaction (PCR) with primers specific for human *cad*, as underlined in Figure 1B, plus flanking PspAI and XhoI sites (primer pairs A and B, A and C, and A and D, respectively). The PCR products were digested by PspAI and XhoI, separated from primers and template by electrophoresis on a 1% Seaplaque agarose gel, and ligated into PspAI- and XhoI-cut PGL2Basic vector (Promega). Promoter constructs hCAD1E and hCAD2E, which contain the 5′ or 3′ E box plus flanking sequences cloned after the minimal promoter, were made by inserting oligonucleotides with XhoI and HindIII ends (5′ E box–5′ CTCGAG CGTTAG CCACGT GGACCG ACTAAG CTT 3′, 3′ E box–5′ CTCGAG CCGTCC TCACGT GGTTCC AGTAAG CTT 3′) into similar sites in the *CAD* minimal construct, hCAD+25, at the +25 position. To create a plasmid to be used in RNase protection assays, hCAD+115 was digested with PspAI and XhoI. The resulting fragment was then isolated by using a 1% Seaplaque agarose gel and ligated into PspAI- and SalI-cut PBSM13 vector, creating PBSCAD115. This plasmid was subsequently digested by EcoRI prior to transcription by T3 RNAP (RNA polymerase). Another plasmid containing internal human *cad* sequences was constructed in several steps. First, primers were chosen from sequence of the 3′ end of the human *cad* gene (as selected from the sequence in GenBank M38561). These primers were then used with reverse transcriptase–PCR, using HeLa RNA as a template. After the production of double-stranded DNA, the resulting fragment was digested with BamHI and then cloned into the BamHI site in PBSM13's multiple cloning site. This plasmid, PBSCAD3', was subsequently digested by EcoRI prior to transcription by T3 RNAP, resulting in a 170-base probe and a 110-base protected area. The protected sequences correspond to bases 6512–6622, which is in the ATCase-encoding portion of the human *cad* gene of the sequence in GenBank accession number D78586.

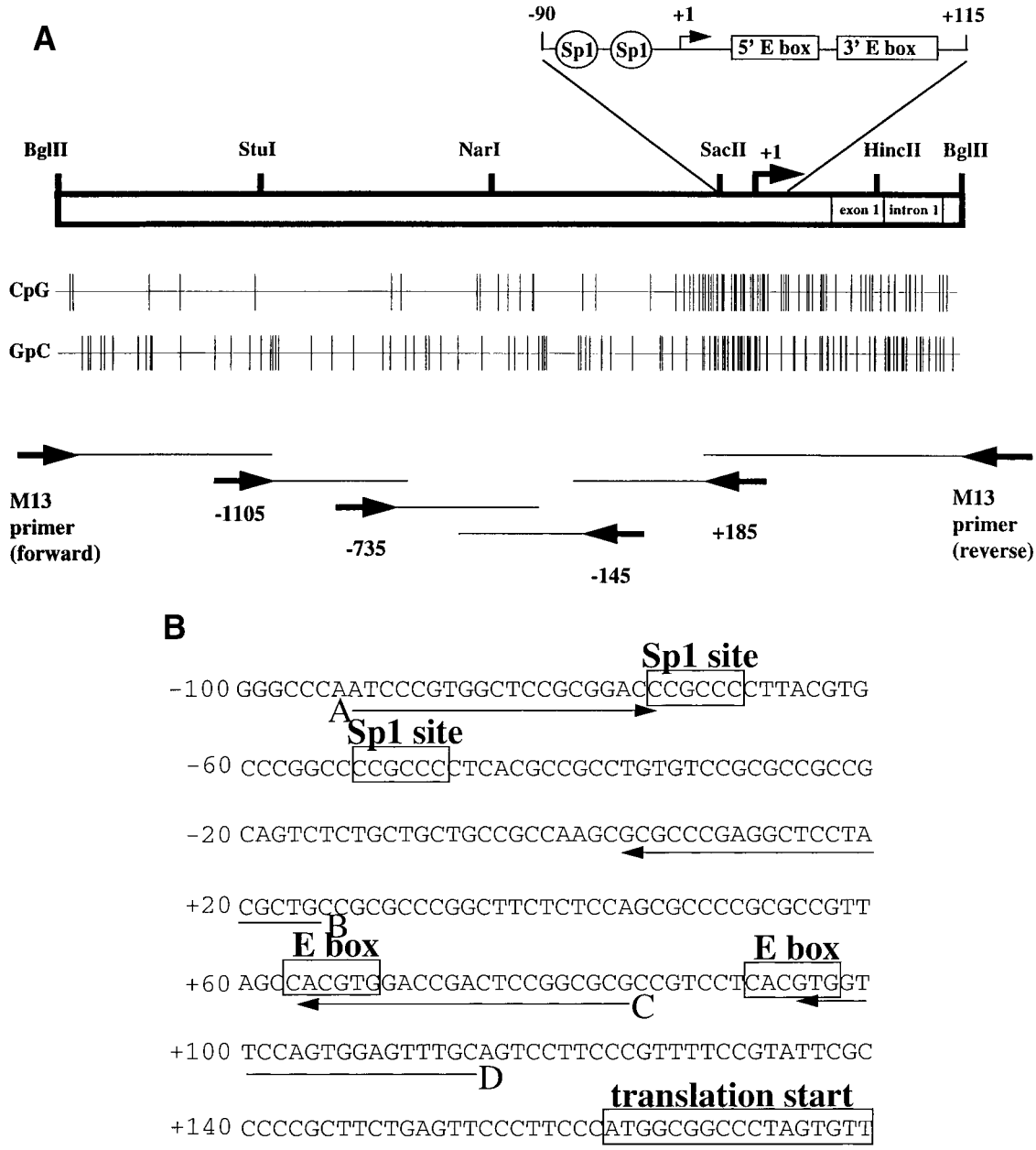

Figure 1. Cloning and sequencing of the human *cad* promoter. (A) Schematic of the 2.0-kb region containing the 5′ end of the human *cad* gene. Overlapping sequence was obtained for each junction from adjacent clones. The 2.0-kb of sequence has GenBank accession number AF015947. The box represents human *cad* sequence; location of the primers (arrows) and the extent of sequences obtained using these primers (lines extending from arrows) are indicated (bottom of page). The major transcription initiation site of *cad* mRNA is indicated by +1 (as determined by homology to the hamster major transcription initiation site and from our assays—see Figure 2). Also shown are selected restriction sites, some of which were used in subsequent cloning, an enlarged depiction of the *cad* promoter region (top of page), and the regions corresponding to exon 1, intron 1, and a small portion of exon 2. Below the box, and above the arrows, are shown the CpG and GpC frequencies, with the horizontal line representing human *cad* sequences and each vertical line representing an individual CpG or GpC dinucleotide. (B) Partial sequence of the human *cad* DNA fragment, with sequences numbered relative to +1, the major transcription initiation site. Indicated are the consensus Sp1 binding sites, the consensus E boxes located within the region required for growth regulation, and the translation start. Also shown are the primers (A–D) used in later cloning steps.

Cell Culture and Transfections

NIH/3T3 cell cultures were maintained and transfected, as described previously, using 1×10^5 cells per sample [13]. Correlation of number of hours after serum stimulation to growth-cycle phase was determined by flow-cytometric analysis of propidium iodide–stained cells [18]. For cotransfection assays, 1 μg of *myc* expression construct or empty vector, 1 μg of *CAD* reporter construct, and 15 μg of sonicated salmon sperm DNA were transfected into 1×10^5 HeLa cells. The cells were then

placed in maintenance medium (5% serum) for 36–48 h prior to harvest. Each transfection was repeated at least twice with duplicate samples and multiple DNA preparations.

Sequencing

For automated sequencing, PCR was performed with 300–500 ng of template, 3 pmol of primer, and 8 μL of the Perkin-Elmer DNA Sequencing Kit-Dye Terminator Cycle Sequencing Ready Reaction (consisting of buffer, AmpliTaq polymerase, and fluorescently labeled nucleotides, catalog number 402119), and the results were interpreted with the ABI Prism Sequencing System. The 2.0-kb BglII human *cad* fragment was initially sequenced using pBSM13 vector primers; the resulting sequence was used to design additional primers. The hCAD+25, hCAD+85, and hCAD+115 promoter constructs were sequenced on both strands using pGL2Basic vector primers by both the automated PCR and 35S dideoxy sequencing methods [19]. The complete 2-kb sequence has GenBank accession number AF015947.

In Vitro Translation and Electromobility Gel Shift Assays

In vitro–translated proteins were made as described previously [13]. The sequence for the 5′ and 3′ E box primers used for gel shift assays are listed in the Plasmid section, as 5′ E box and 3′ E box. An oligonucleotide with a mutated E box, mtE, with sequence 5′AGCGAG CCTGCA GGACCA ACT3′, was used as competitor in some reactions.

Electromobility gel shift assays were performed as previously described [17]. The probes were either end-labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase or labeled after digestion by XhoI and HindIII by filling in with [α-^{32}P] dCTP by Klenow polymerase. The gel and running buffer were 6.25 mM Tris–morpholinepropanesulfonic acid pH 7.0 and 0.125 mM EDTA. Antibodies were purchased from Santa Cruz Biotechnology (anti-Max sc-765X, anti-c-Myc sc-764X, and anti-USF sc-229X).

RNase Protection Assay

PBSCAD115 and PBSCAD3′ linearized with EcoRI were transcribed using T3 RNAP and [α-^{32}P]-GTP to make riboprobes containing human *cad* from nt −90 to +115 and 110 bases of 3′ sequence, respectively (please see Plasmid section for further description). For each sample, 1 ng of this riboprobe, with or without 40–80 μg of the appropriate RNA, was precipitated and resuspended in 10 μL of an 80% formamide, 40 mM PIPES, 0.4 M NaCl, and 1 mM EDTA solution. These samples were denatured by a 10-min incubation at 85°C and then incubated at 55°C for 3 h. The samples were next digested for 30 min at 30°C after addition of 300 μL of a 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.3 M NaCl solution containing 15 μg of RNase A and 2 U of RNase T1. After addition of 10 μL of 20% SDS and 2.5 μL of 20 mg/mL proteinase K, the samples were incubated for 15 min at 37°C. The samples were then extracted with phenol and precipitated with 150 μL of 7.5 M NH_4Ac and 900 μL of ethanol, prior to resuspension in 20 μL of a 90% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol solution. The samples were then boiled for 2 min and resolved on an 8% acrylamide, 1X TBE (Tris–boric acid–EDTA), and 7 M urea gel for 2 h in a 1X TBE running buffer. The gel was then dried and bands visualized by autoradiography.

RESULTS

Cloning and Sequencing of the Human *cad* Promoter

To clone the human *cad* promoter, we obtained a cosmid containing human genomic DNA that had been isolated by hybridization to the hamster *cad* cDNA. To determine whether the cosmid's 31-kb insert contained the human *cad* promoter, we digested this cosmid with EcoRI, BglII, or HindIII and performed Southern blot analysis on the digests, using as probe a fragment spanning from nt −333 to +222 of the hamster *cad* promoter. We detected strong hybridization to bands in each digest (data not shown), indicating that the human *cad* promoter was contained in the cosmid. Sequence analysis revealed that a 2.0-kb BglII fragment contained a region that was 75% identical to the hamster *cad* promoter (Figure 1A and B).

As a first step in characterizing the human *cad* promoter, the 2.0-kb BglII fragment was evaluated using the Neural Network Promoter Prediction program. This program, which relies heavily on the presence of a TATA box at nt −30, could not identify a promoter region. However, we noted that the human *cad* genomic DNA had an extremely high G+C content. Promoter regions of many housekeeping genes have been found to be GC-rich and to contain a high CpG frequency [20]. In the genome, the cytosine of CpG dinucleotides is often methylated at the 5′ position. Subsequent deamination of the m^5C results in a C to T transition and causes a genome-wide CpG depletion, such that there are only approximately 20% of the CpGs that would be predicted from the C and G content [21,22]. However, the CpGs located in promoter regions are often protected from methylation. These CpG-rich regions have been called CpG islands (defined as a region of genomic DNA having greater than 50% G+C content, an observed/expected CpG frequency of greater than 60%, and being greater than 200 bp in length) [23]. Therefore, we analyzed human *cad* sequences for the presence of a CpG island. The locations of the CpG dinucleotides along the 2-kb fragment are indicated in Figure 1A. For comparison, the GpC frequency is also shown to demonstrate, by contrast, the degree of CpG deple-

Table 1. Size of the GC-Rich Region, the % G+C, and the Ratio of Observed to Expected CpG Frequency in *cad* Sequences Spanning the Transcription Initiation Site from Different Species*

Species	Size (bp)	%G+C	Obs/exp CpG
Human	691	68	0.88
Hamster	457	70	0.93
Mouse	853	58	0.81
Rat	829	60	0.83

*CpG island criteria are % G+C greater than 50, an observed/expected ratio of CpG greater than 0.6, and larger than 200 bp [21]. The observed/expected CpG ratio was calculated as follows: [(Number of CpGs) × (Size in Nucleotides)] / [(Number of Cs) × (Number of Gs)].

tion that begins to appear in sequences farther away from the transcription start site. We found that the human, hamster, mouse, and rat *cad* genes all contained CpG islands (Table 1) that span the transcription initiation site; we suggest that identification of a CpG island is a more accurate predictor of a housekeeping gene promoter than current computer models.

A comparison of the human and hamster *cad* promoters is shown in Figure 2A. There was a strong conservation of the sequence and relative position of the two consensus Sp1 sites and of the consensus E box required for growth regulation of the hamster *cad* promoter. However, the human *cad* promoter contains a second E box 30 bp downstream of the first. Although the promoter region was in general highly conserved between human and hamster *cad* genes, the sequences immediately surrounding the transcription initiation site varied considerably. Instead of the 93% match to a consensus initiator that is found in the hamster *cad* sequence, at the homologous position in the human sequence there was only a 62% match to the consensus initiator (Table 2). However, three other possible initiator sites, at the −45, −20, and +45 positions relative to the site homologous to the major start site in the hamster *cad* promoter, were 71%, 95%, and 86% matches to a consensus initiator, respectively. Besides matching to consensus initiator, another factor that influences start site selection was distance from Sp1 sites. It has been proposed that RNA polymerase II scans sequences 50–100 bp downstream of a Sp1 site and, if it finds a good match to a consensus initiator, will start transcription at that point. In fact, an analysis of promoter elements indicated that the frequency of Sp1 site position peaks at −50 relative to the transcription start site in vertebrate gene promoters [24]. Studies of the hamster *cad* promoter have confirmed the importance of both Sp1 sites and initiator elements [25]. In the human *cad* promoter, the best matches to a consensus initiator were not the optimal distance downstream from an Sp1 site. Thus, all four initiator sequences were possible transcription initiation sites. To determine which of these sites, if any, serves as the major start site for human *cad*, we mapped human *cad*'s transcription initiation site using RNase protection assays.

We first utilized a riboprobe complementary to 110 bases near the 3′ end of the human *cad* coding sequences, PBSCAD3′. This probe had previously been used to detect *cad* mRNA and served as a control for our reagents and technique. As shown in Figure 2B, lane 2, a protected band was clearly detected. To map the start sites, we utilized a riboprobe complementary to human *cad* sequences spanning from −93 to +115. Our experiments revealed transcripts whose lengths—approximately 125 and 145 nucleotides (Figure 2B, lanes 5 and 6, indicated by B and C)—corresponded to transcription initiating from both the +1 and −20 sites. These sites correspond to the major and minor start sites of the hamster *cad* mRNA. Additionally, we observed a protected band of 210 nt, corresponding to the length of the entire region of the probe that contains human *cad* sequences (Figure 2B, lanes 5 and 6, indicated by A). This band represents transcription initiating either at or upstream of nt −93. Phosphorimage analysis and normalization to the number of guanines in the protected probe indicated that the total contribution of these transcripts was less than 25% of the amount attributed to the two downstream start sites. Although the sequences upstream of nt −93 did not contain any consensus Sp1 sites, a GT box existed at −174; Sp1, Sp2, and Sp3 have been shown to bind to similar GT boxes [26,27]. There were several candidate initiator sequences located 50–100 bp downstream of this GT box, including a 76% match at −137, an 87% match at −118, and an 85% match at −96 [24]. The GT box may position transcription to initiate at one or several of these sites, producing transcripts that would result in the protected band of 210 nt. As a second method for determining the transcription initiation site, we also performed in vitro transcription assays by using HeLa nuclear extract. These in vitro experiments identified the same start sites as did the RNAse protection assays of cellular RNA (data not shown). We believe that the presence of multiple start sites for the human *cad* promoter is due to the fact that none of the sites are both highly homologous to a consensus initiator and positioned optimally downstream of the Sp1 sites. We have designated the 3′ most site to be +1 to agree with the numbering system used in the hamster *cad* promoter.

Characterization of the G_1/S Phase–Specific Elements in the Human *CAD* Promoter

Our lab has previously shown that the transcription rate of the hamster *cad* promoter is linked to the proliferative state of the cell [17]. In the first cell

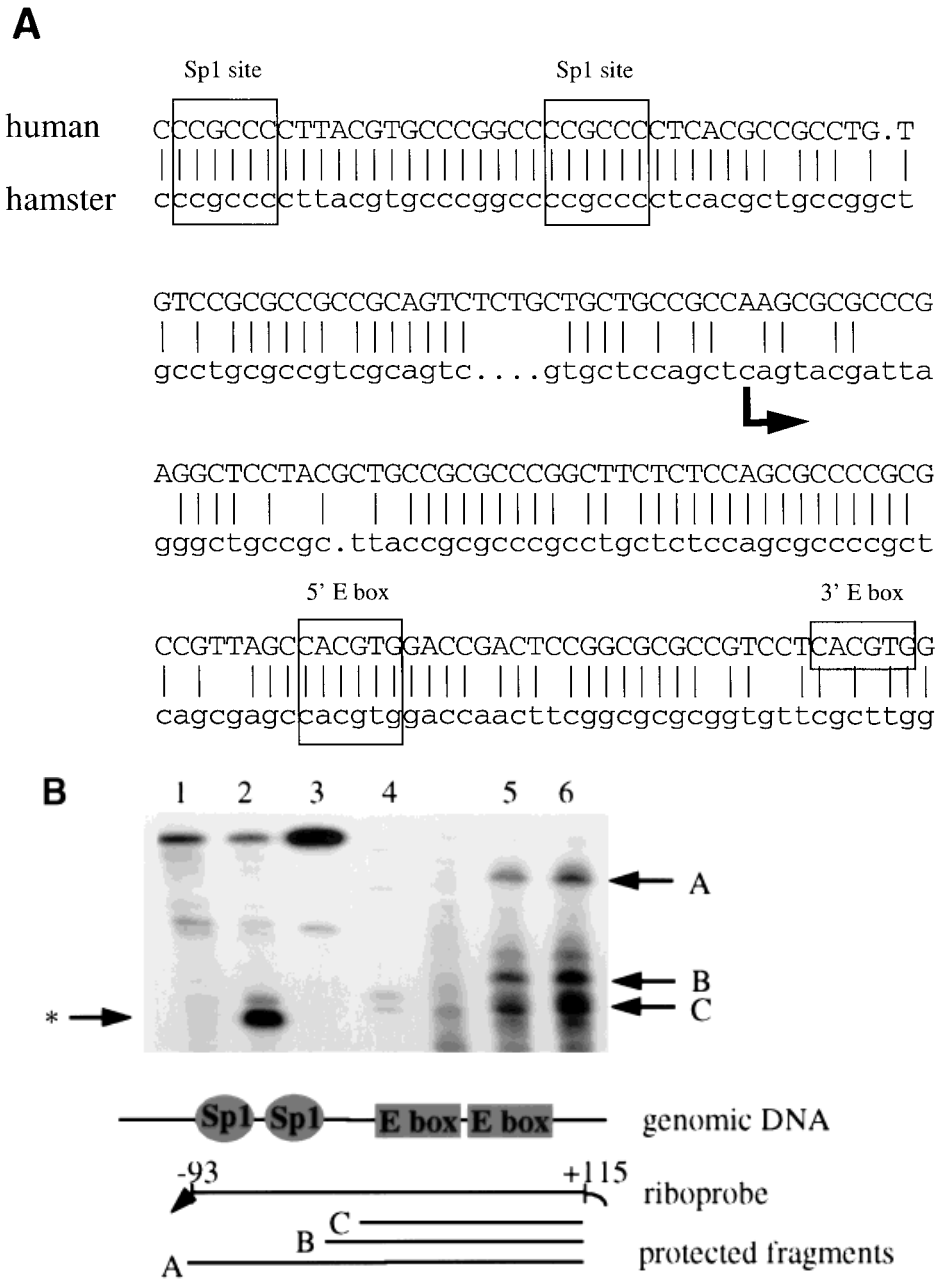

Figure 2. Characterization of the human *cad* promoter. (A) Sequence comparison of human and hamster DNA. Shown is the region of the hamster *cad* promoter known to be sufficient for growth-regulated transcription. The overall identity of this region between the human and hamster DNA was 75%. The Sp1 sites and E boxes are boxed. The major transcription initiation site for hamster *cad* is indicated by the arrow. The dots indicate gaps in the alignment; sequences separated by a series of dots are contiguous. (B) RNase protection assay of human *cad*. Riboprobes containing sequences spanning from −93 to +115 of 5′ human *cad* sequence or 110 bases of 3′ human *cad* sequence were hybridized to 200-2 or HeLa cytoplasmic RNA, respectively, for 3 h prior to digestion by RNAses A and T1. The resulting products were separated on a 8% acrylamide/ 7 M urea gel. Arrows labeled A, B, and C indicate bands protected by the 5′ probe, the arrow labeled * indicates the major band protected by the 3′ probe. Below the gel are indicated the sequences that the 5′ riboprobe spans as well as the transcripts corresponding to bands A–C. Lanes 1–3: the 3′ probe was hybridized to tRNA (lane 1), 40 µg of HeLa RNA (lane 2), or run as intact probe without any digestion (lane 3). Lane 4: size marker. Lanes 5 and 6: the 5′ probe was hybridized to 40 µg of 200-2 RNA.

cycle following serum stimulation, also referred to as the growth cycle, a fifteenfold increase in transcription from the hamster *cad* promoter is mediated through c-Myc binding to an E box at +65 [17]. An extensive time course analysis has shown that although transcription begins in mid G1 phase, maximal activation of hamster *cad* transcription occurs near the G1/S phase boundary [17]. The E box that is required for regulation of the hamster *cad* promoter was conserved in the human *cad* promoter, suggesting that the human *cad* promoter may also be growth regulated. To investigate this possi-

Table 2. Some Possible Initiators Found in the Human *cad* Promoter, Their Match to the Consensus Initiator Sequence, and Their Distance From the Two Sp1 Sites*

Potential initiator	% Match to consensus initiator	Distance (bp) from the two Sp1 sites
Hamster +1 TCAGTACG	93	49, 70
Human −45 TCACGCCG	71	7, 28
−20 GCAGTCTC	95	32, 53
+1 CCAAGCGC	62	51, 72
+45 CCAGCGCC	86	95, 116

*% Match to consensus was calculated according to the formula in Bucher [24]. All numbering is based on +1 being the site homologous to the hamster *Cad* transcription start site.

bility, we cloned sequences from the human *cad* promoter spanning from −90 to +115 upstream of the luciferase cDNA (Figure 3A). This construct was transiently transfected into NIH-3T3 cells, and promoter activity was measured in quiescent serum-starved cells as well as in cells stimulated by serum for various lengths of time. These timepoints correlate to different phases of the growth cycle, as indicated by flow-cytometric analysis of propidium iodide–stained cells [18]. As shown in a time course from a typical experiment (Figure 3B), transcription from the human *cad* promoter was increased after serum stimulation, showing a peak of activity in early S-phase.

The human *cad* promoter contains two E boxes, located 65 and 94 bp downstream of the transcription start site. Several other E box–containing promoters have two or more E boxes, such as ornithine decarboxylase (*Odc*) [28,29], *eIF-4E* [30], *cdc 25A* [31], and prothymosin α [12]. In prothymosin α, a second, nonconsensus E box determines specificity for activation of a nearby consensus E box, allowing c-Myc to transactivate the promoter, but preventing TFE3 from doing so [12]. The E boxes in the murine *Odc* gene appear to cooperate, providing an additive effect in transactivation assays [28]. We wanted to examine the roles of the two E boxes in growth regulation of transcription of the human *cad* promoter. To do this, we compared growth-regulated transcription activity from luciferase reporter constructs which contained two Sp1 sites, plus either no, one, or two E boxes (Figure 3A). The constructs containing one or both E boxes showed significant G1/S-phase inductions, while transcription from the construct lacking an E box increased very little at the G1/S-phase boundary (Figure 3B). Averaging peak G1/S-phase inductions from 12 experiments, transcription at the G1/S-phase boundary increased 10–12-fold from constructs containing at least one E box and only threefold from the construct lacking an E box (Figure 3C).

The deletion experiments suggested that the 5′, but not the 3′, E box was required for growth regulation of *cad* transcription. However, it remained possible that the two E boxes had redundant functions (i.e., the 3′ E box may be able to function in the absence of the 5′ E box). To test this hypothesis, we placed a short oligonucleotide, containing only 20 bp of human *cad* sequence surrounding either the 5′ or 3′ E box immediately downstream of the minimal promoter (Figure 4A). In a typical serum-starvation and -stimulation assay (Figure 4B), expression from the construct containing the 5′ E box was induced ten-fold at the G1/S-phase boundary, while expression from the construct containing the 3′ E box was induced fourfold, a level comparable to that of the minimal promoter. Averaging peak G1/S-phase inductions from six experiments, we found that expression increased over sevenfold from the construct containing the 5′ E box, compared to threefold to fourfold from the other constructs (Figure 4C). From this, we conclude that the 5′ E box, but not the 3′ E box, is sufficient to confer growth regulation of transcription to the minimal *cad* promoter. Although these two E boxes are very similar, CCACGTGG versus TCACGTGG, they have different functions in the transcriptional assay. Investigations to link function and specific protein binding are discussed below.

Protein Binding to the E Boxes

Although both of the E boxes in the *cad* promoter contain consensus CACGTG elements, they clearly function differently with respect to the ability to confer growth regulation. We wanted to determine what is responsible for this differential function. A previous study found that distance of the E box from the start site can convey some specificity, as USF could not transcriptionally activate from E boxes located more than 400 bp away from the start site [12]. However, since we placed both human *cad* E boxes at +25, differences in distance cannot explain our results.

Flanking sequences have also been shown to influence protein binding to E boxes [32]. A T in the −4 position (i.e., **T**CACGTG) or greater than two AT pairs in the three flanking nucleotides (i.e., XXXCACGTGXXX, where three or more of the X = A or T), will predispose a given E box to be bound by USF but not by c-Myc/Max [30]. The natural sequences surrounding the hamster *cad* E box allow binding by both c-Myc and USF [13]. The 5′ E box in the human *cad* promoter contains the same flanking sequences as found in the hamster *cad* promoter, suggesting that this element may bind both c-Myc/Max and USF, whereas the sequences flanking the 3′ E box are AT rich, suggesting that this may bind only USF.

To explore possible differential protein binding by the two human *cad* E boxes, we performed electromobility gel shift assays, incubating either HeLa

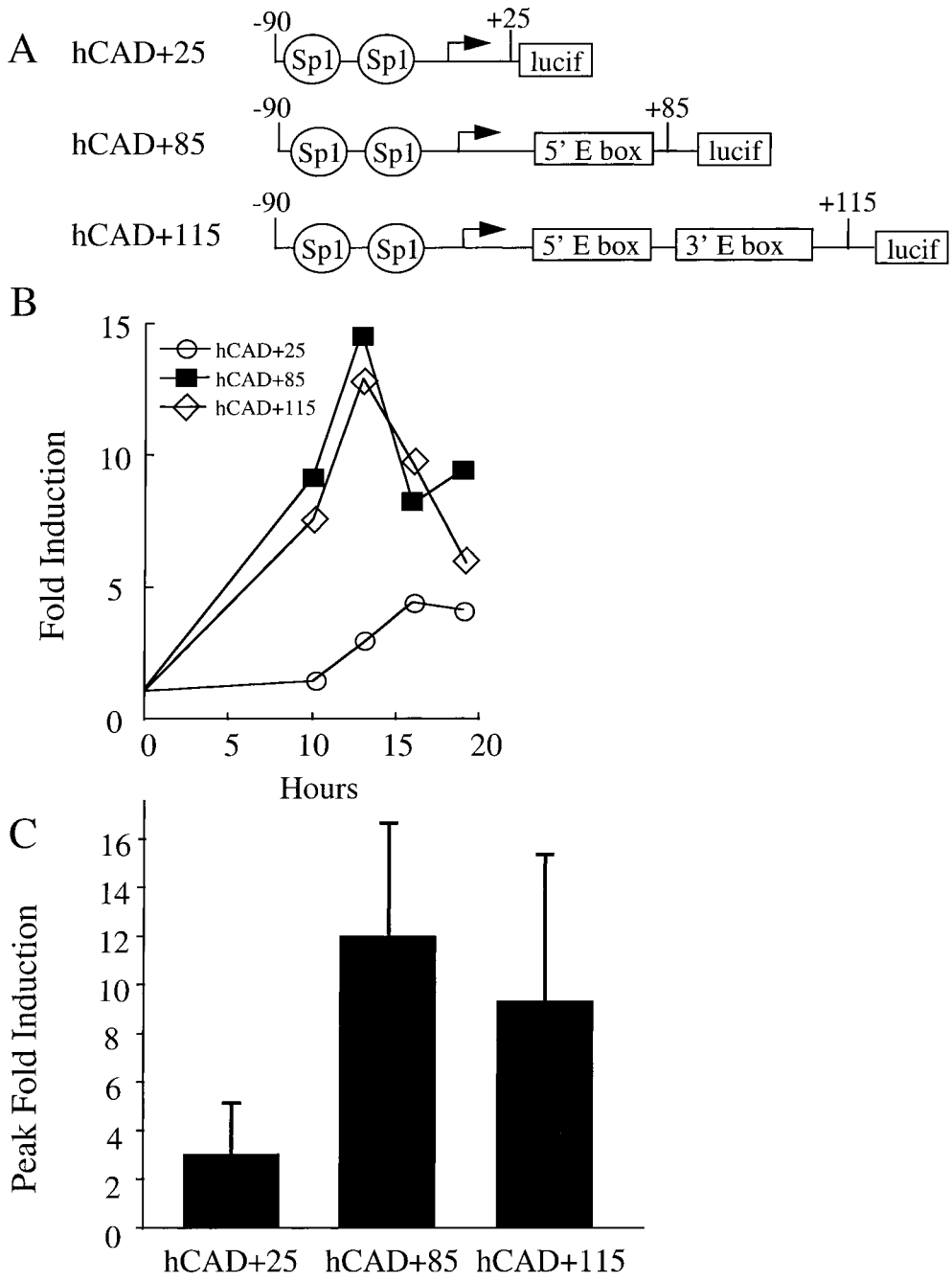

Figure 3. The human *cad* promoter display growth-regulated transcriptional activity. (A) Schematic representation of *cad* reporter constructs. Human *cad* promoter constructs were fused to the luciferase cDNA, as indicated. The major transcription start site is indicated by the arrow. Locations of the Sp1 binding sites and the E box(es) are approximated by open symbols. (B) Graph of induction of *cad* reporter activity from a representative experiment. Cells were transiently transfected with the indicated reporter constructs and serum starved for 48 h and then serum stimulated. Promoter induction is reported as the ratio of luciferase activity from a given promoter construct measured in cells harvested at timepoints following serum stimulation relative to activity from the same promoter construct in serum starved cells. (C) Graphical representation of the average induction of *cad* reporter activity through the growth cycle of NIH-3T3 cells. Results are representative of 12 experiments, using four different DNA preparations; error bars indicate standard error.

nuclear extract, as a source of USF (Figure 5), or in vitro–translated c-Myc/Max (Figure 6), with radioactively labeled oligonucleotides containing either one or the other E box, plus their respective flanking sequences. In vitro–translated c-Myc/Max was used since functional c-Myc is not easily extracted by nuclear extract preparation. We found that protein from HeLa nuclear extract (Figure 5, lanes 2 and 6) bound to both E boxes. Binding was specific, since the addition of a 100-fold excess of unlabeled wild-

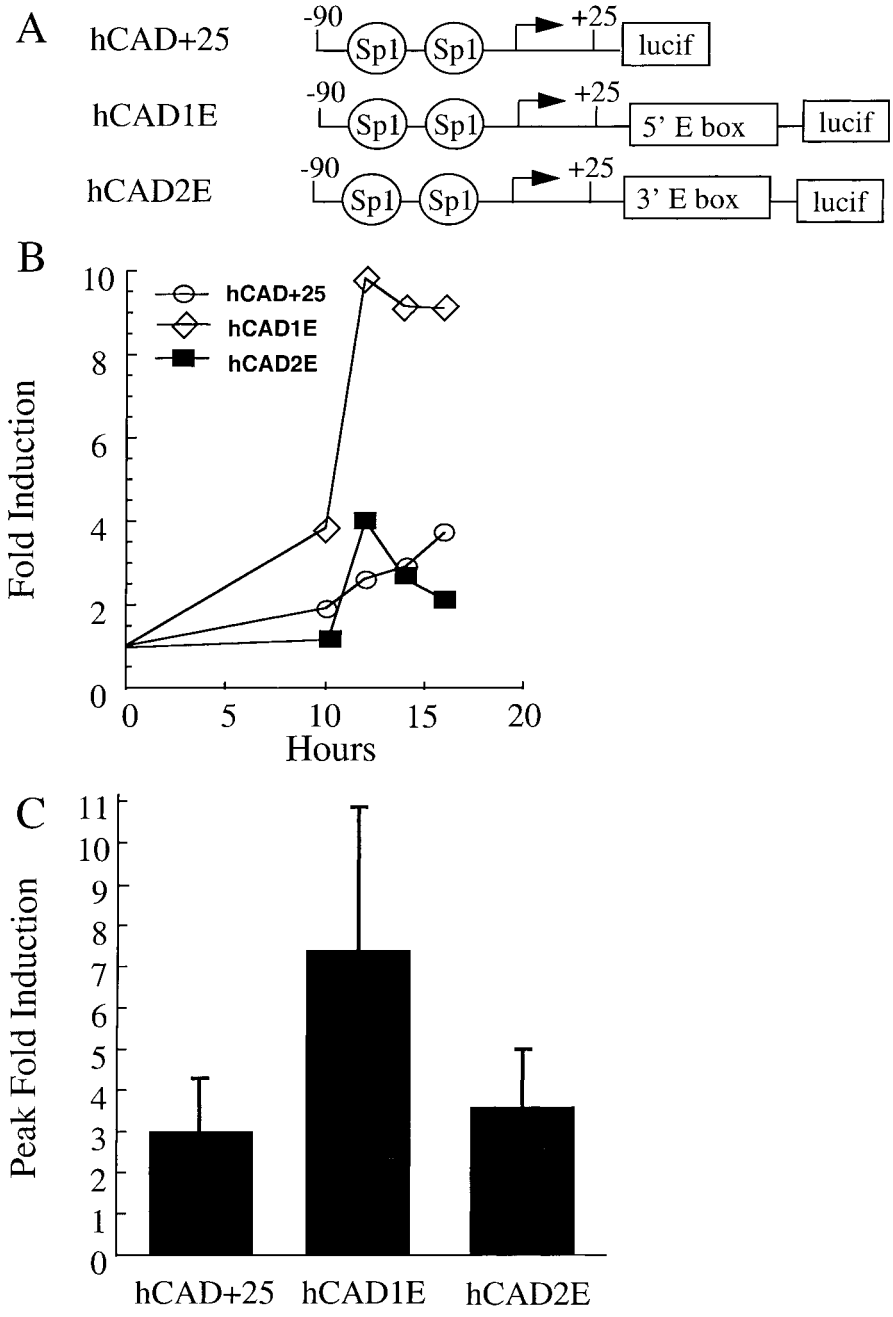

Figure 4. The 5′ but not the 3′ E box can confer growth regulation (A). Schematic representation of synthetic *cad* reporter constructs. Human *cad* promoter constructs were fused to the luciferase cDNA, as indicated. The major transcription start site is depicted by the arrow. Locations of the Sp1 binding sites and the E box(es) are approximated by open symbols (B). Graph of induction of *cad* synthetic reporter activity from a representative experiment. Promoter induction is reported as the ratio of luciferase activity from a given promoter construct measured in cells harvested at timepoints following serum stimulation relative to activity from the same promoter construct in serum starved cells (C). Graphical representation of the average induction of *cad* reporter activity through the growth cycle of NIH 3T3 cells. Results are representative of six experiments, using four different DNA preparations; error bars indicate standard error.

type oligonucleotide reduced binding (lanes 3 and 7), while addition of a 100-fold excess of unlabeled mutant oligonucleotide had no such effect (data not shown). Incubation with antibody specific to USF disrupted the DNA-protein complex and resulted in a partial supershift (lanes 4 and 8). These results indicated that USF binds to both E boxes.

In vitro–translated c-Myc/Max formed three complexes with the 5′ E box (Figure 6, lane 2), and two with the 3′ E box (lane 9), in our gel shift assays.

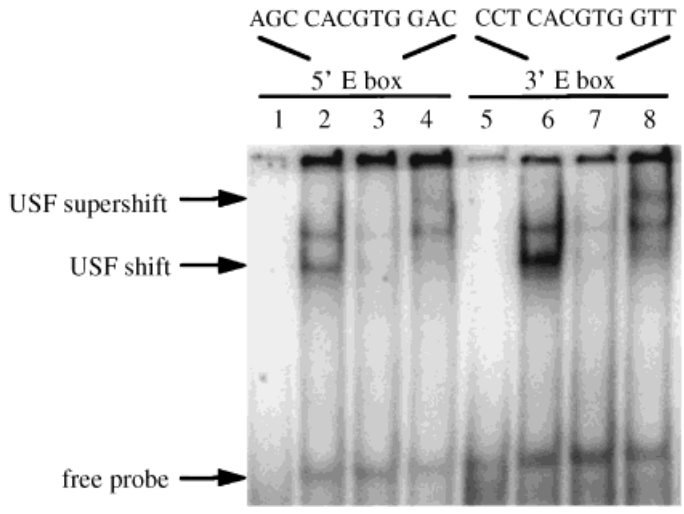

Figure 5. Protein binding to E box sequences from the human *cad* promoter. Gel mobility shift assays were performed using fragments that contained either the 5′ E box (lanes 1–4) or the 3′ E box (lanes 5–8), their respective flanking sequences, and XhoI and HindIII restriction enzyme sites. Please see Materials and Methods for full oligonucleotide sequences. Lanes 1 and 5: probe alone; lanes 2 and 6: with 10 μg of HeLa nuclear extract, lanes 3 and 7: with 10 μg of HeLa nuclear extract and 100-fold excess of unlabeled wild-type oligonucleotide, lanes 4 and 8: with 10 μg of HeLa nuclear extract and 2 μg of anti-USF antibody.

Binding was specific, since the addition of a 300-fold excess of unlabeled wild type oligonucleotide disrupted complex formation (lanes 3 and 10), while addition of a 300-fold excess of unlabeled mutated oligonucleotide had no such effect (lanes 4 and 11). Incubation with antibody specific to c-Myc disrupted the intermediate complex seen with the 5′ E box but did not alter any complexes seen with the 3′ E box (lanes 5 and 12), indicating that c-Myc was present in the intermediate complex only. Incubation with antibody specific to Max disrupted two of the complexes (lanes 6 and 13), indicating that both complexes contained Max. The largest complex, seen with both E boxes, has not been identified, as it was not disrupted by antibodies to c-Myc, Max, or USF (lanes 5, 6, and 7 and 12, 13, and 14). However, this binding was specific to E box sequences, as it was competed by excess wild-type, but not mutated, oligonucleotide. The complex is likely to be due to binding of a protein present in the rabbit reticulocyte lysate. Our results indicateed that Max/Max homodimers bound to both E boxes, but binding to the 5′ E box was favored. Importantly, c-Myc/Max bound to the 5′ E box, but not the 3′ E box, demonstrating a link between c-Myc binding and transcriptional activity of the *cad* promoter.

Based on our findings that binding of c-Myc/Max to human *cad* E boxes correlated with those E boxes' ability to confer growth regulation, we next postulated that c-Myc would be able to induce expression from a human *cad* promoter construct containing the 5′, but not the 3′, E box. Therefore, we

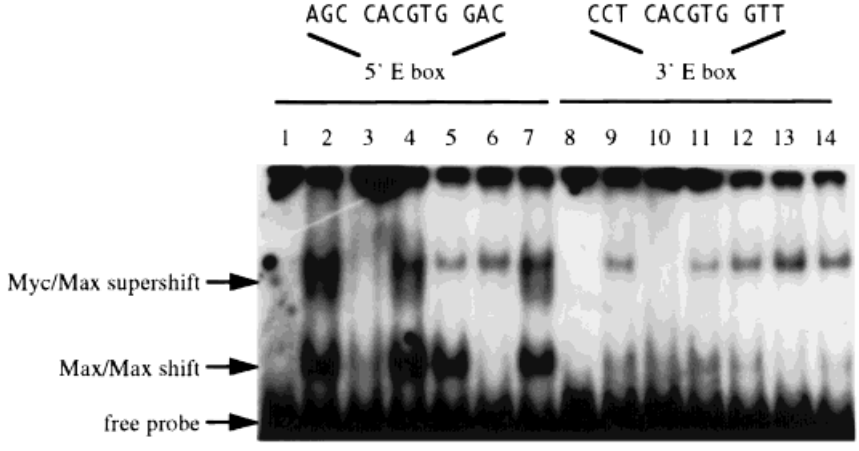

Figure 6. Protein binding to E box sequences from the human *cad* promoter. Gel mobility shift assays were performed using fragments that contained either the 5′ E box (lanes 1–7) or the 3′ E box (lanes 8–14), their respective flanking sequences, and XhoI and HindIII restriction enzyme sites. Please see Materials and Methods for full oligonucleotide sequences. Lanes 1 and 8: probe alone, lanes 2 and 9: with c-Myc/Max; lanes 3 and 10: c-Myc/Max and 300-fold excess of unlabeled wild-type oligonucleotide, lanes 4 and 11: c-Myc/Max and a 300-fold excess unlabeled mutant oligonucleotide, lanes 5 and 12: c-Myc/Max and 2 μg of antibody specific to c-Myc; lanes 6 and 13: c-Myc/Max and 2 μg of antibody specific to Max, lanes 7 and 14: c-Myc/Max and 2 μg of antibody specific to USF.

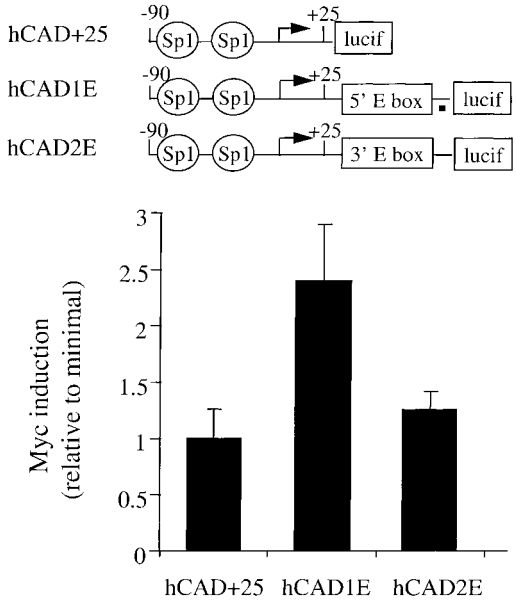

Figure 7. The human *cad* promoter responded modestly to transfected c-Myc. A graphical representation of the average induction of *cad* reporter activity by a c-Myc expression vector after transfection into HeLa cells, as compared to activity seen when the *cad* reporter plasmid was cotransfected with empty vector, is shown. The induction of the minimal promoter (hCAD+25), which lacks an E box, was set equal to one. Results are representative of eight experiments, using multiple different DNA preparations; error bars indicate standard error. Shown at the top is a schematic representation of synthetic *cad* reporter constructs. Human *cad* promoter constructs were fused to the luciferase cDNA, as indicated. The transcription start site is depicted by the arrow. Locations of the Sp1 binding sites and the E boxes are approximated by open symbols.

performed a cotransfection experiment in which a c-Myc expression plasmid was introduced into HeLa cells in combination with a *cad* promoter construct lacking an E box (hCAD+25), containing the 5′ E box (hCAD1E), or containing the 3′ E box (hCAD2E). Analysis of luciferase activity indicated that while the 5′ E box was able to confer induction by c-Myc, the 3′ E box was not (Figure 7). Although these results are modest, they are comparable to those obtained using other c-myc target genes [33]. In summary, we showed that the 5′ E box is a c-Myc binding site (Figure 6), can confer growth regulation (Figure 4), and c-Myc responsiveness (Figure 7).

CAD Expression in Burkitt's Lymphoma Cell Lines

To further explore c-Myc regulation of the *cad* promoter, we measured expression of *cad* mRNA in Burkitt's lymphoma cell lines derived from tumors having deregulated c-Myc. Using western blot analysis, we confirmed previously published reports that the Burkitt's lymphoma lines we chose, Daudi, DG75, and Raji, have elevated c-Myc levels (data not shown). As another control, we also assayed glyceraldehyde-3 phosphate dehydrogenase RNA from these cell lines and found the RNA to be of equivalent quality and quantity (data not shown). In RNAse protection assays, we could not detect any

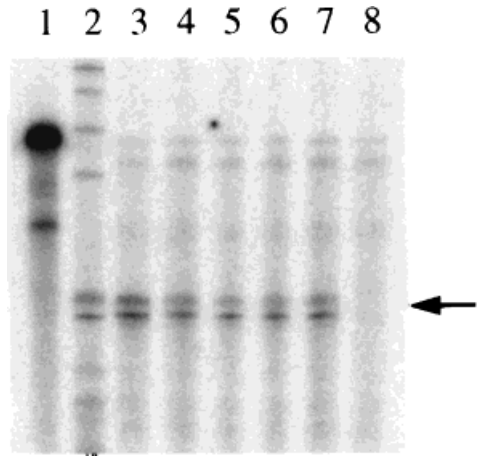

Figure 8. *CAD* mRNA levels were not elevated in Burkitt's lymphoma lines with increased c-myc. A riboprobe containing 110 bases of 3′ human *cad* sequence was hybridized to cytoplasmic RNA from Burkitt's lymphoma cell lines containing increased c-myc (Daudi, DG75, and Raji) as well as from HeLa cells and LCL 721, an Epstein Barr virus–positive human lymphoblastoid cell line without translocated or otherwise deregulated c-myc, prior to digestion by RNAses A and T1. The resulting products were separated on an 8% acrylamide/ 7 M urea gel. The arrow to the right of the gel indicates the *cad* products, which run as a doublet of identical sizes as 121- and 127-base markers. Lane 1: undigested probe; lane 2: size-marker; lane 3: 40 μg of HeLa RNA; lane 4: 80 μg of LCL 721 RNA; lane 5: 80 μg of DG75 RNA; lane 6: 80 μg of Raji RNA; lane 7: 80 μg of Daudi RNA; lane 8: tRNA alone.

significant difference in *cad* mRNA levels between the various Burkitt's lymphoma lines and LCL721, our negative control, an Epstein Barr virus–positive human lymphoblastoid cell line without translocated or otherwise deregulated c-Myc (Figure 8).

DISCUSSION

Deregulation of Myc proteins is strongly correlated with many human tumors; however the mechanism by which these proteins cause neoplasia is still unknown. Although expression of several genes has been shown to be positively or negatively regulated by c-Myc protein, identification of c-Myc effector genes relevant to neoplasia has thus far progressed slowly. A true c-Myc effector gene should meet three criteria. First, the protein encoded by a c-Myc target gene should perform a function that promotes proliferation, if transactivated by c-Myc, or blocks proliferation, if repressed by c-Myc. Second, a c-Myc target gene should contain one or more E boxes to which c-Myc/Max binds. Third, expression of a c-Myc target gene should closely correlate with the expression of c-Myc proteins, if activated by c-Myc, or be inversely correlated, if repressed by c-Myc.

The CAD protein performs the first three rate-limiting steps in pyrimidine biosynthesis and thus clearly promotes proliferation. We have shown that the human (this manuscript), mouse [34], and hamster [17] promoters all show E box–dependent growth regulation and that c-Myc/Max heterodimers bind to the E boxes. Finally, *cad* expression

levels correlate with the abundance of c-Myc protein in two systems: (i) In the growth cycle, c-Myc protein levels increase as the cell progresses from quiescent to S-phase, in a pattern similar to *cad* mRNA levels. (ii) In c-Myc knockout cells, loss of c-Myc protein abrogates the increase in *CAD* mRNA normally seen in S phase [16]. However, our results using Burkitt's lymphoma cell lines, suggest that *cad* cannot be further activated by increasing levels of c-Myc protein over the amount normally found in cells. We suggest that levels of c-Myc required for maximal activity of the *cad* promoter are already present in the cell. Preliminary data in support of this hypothesis come from chromatin immunoprecipitation experiments. We have shown that c-Myc binds to the human *cad* promoter in vivo [34]. However, a comparison of protein binding to cad in B cells having low versus high amounts of c-Myc indicates that increasing the amount of c-Myc over normal levels does not result in more c-Myc bound to *cad* in vivo (unpublished data). The complete abrogation of cad growth regulation seen in c-Myc knockout cells, however, does indicate an absolute requirement for c-Myc under normal physiological conditions.

In summary, we propose that although the human *cad* gene is regulated by c-Myc in normal cells, increased levels of CAD protein are not critical for the ability of c-Myc to induce neoplasia. However, the *cad* gene was the only gene whose normal growth regulation is abrogated in c-Myc knockout cells. We suggest that the genes whose deregulation is correlated with c-Myc–mediated neoplasia may constitute a different set of target genes than those controlled by c-Myc under normal physiological conditions. Our current studies are focused on the identification of genes specifically deregulated by c-Myc in tumor cells.

ACKNOWLEDGMENTS

We thank Kathy Boyd, Chris Fry, Stephanie Bartley, and Mark Sandberg for guidance and technical assistance and George Stark, Bob Eisenman, Olga Chernova, Thea Tlsty, and Bill Sugden for sharing reagents and cell lines. This work was supported by grants CA45240, CA07175, and CA09135.

REFERENCES

1. Davis AC, Wims M, Spotts GD, Hann SR, Bradley A. A null c-myc mutation causes lethality before 10.5 days of gestation in homozygotes and reduced fertility in heterozygous female mice. Genes Dev 1993;7:671–682.
2. Evan GI, Littlewood TD. The role of c-*myc* in cell growth. Curr Opin Genet Dev 1993;3:44–49.
3. Henriksson M, Luscher B. Proteins of the *myc* network: Essential regulators of cell growth and differentiation. Adv Cancer Res 1996;68:109–182.
4. Klein G. Multistep evolution of B-cell-derived tumors in humans and rodents. Gene 1993;135:189–196.
5. Amati B, Brooks MW, Levy N, et al. Oncogenic activity of the c-*myc* protein requires dimerization with Max. Cell 1993; 72:233–245.
6. Schmidt EV. *myc* family ties. Nat Genet 1996;14:8–10.
7. Roussel MF, Ashmun RA, Sherr CJ, Eisenman RN, Ayer DA. Inhibition of cell proliferation by the Mad1 transcriptional repressor. Mol Cell Biol 1996;16:2796–2801.
8. Schreiber-Agus N, Chin L, Chen K, et al. An amino-terminal domain of Mxi1 mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor SIN3. Cell 1995;80:777–786.
9. Gu W, Cechova K, Tassi V, Dalla-Favera R. Opposite regulation of gene transcription and cell proliferation by c-Myc and Max. Proc Natl Acad Sci USA 1993;90:2935–2939.
10. Hurlin PJ, Queva C, Koskinen PJ, et al. Mad3 and Mad4: novel Max-interacting transcriptional repressors that suppress c-myc dependent transformation and are expressed during neural and epidermal differentiation. EMBO 1995;14:5646–5659.
11. Luo X, Sawadogo M. Functional domains of the transcription factor USF2: Atypical nuclear localization signals and context-dependent transcriptional activation domains. Mol Cell Biol 1996;16:1367–1375.
12. Desbarats L, Gaubatz S, Eilers M. Discrimination between different E-box-binding proteins at an endogenous target gene of c-*myc*. Genes Dev 1996;10:447–460.
13. Boyd KE, Farnham PJ. Myc versus USF: Discrimination at the *cad* gene is determined by core promoter elements. Mol Cell Biol 1997;17:2529–2537.
14. Lee LA, Dolde C, Barrett J, Wu CS, Dang CV. A link between c-Myc-mediated transcriptional repression and neoplastic transformation. J Clin Invest 1996;97: 1687–1695.
15. Li L-H, Nerlov C, Prendergast G, MacGregor D, Ziff EB. c-Myc represses transcription in vivo by a novel mechanism dependent on the initiator element and Myc box II. EMBO J 1994;13:4070–4079.
16. Bush A, Mateyka M, Dugan K, et al. c-Myc null cells misregulate *cad* and *gadd45* but not other proposed c-myc targets. Genes Dev 1998;12:3797–3802.
17. Miltenberger RJ, Sukow K, Farnham PJ. An E box-mediated increase in *Cad* transcription at the G1/S-phase boundary is suppressed by inhibitory c-Myc mutants. Mol Cell Biol 1995;15:2527–2535.
18. Slansky JE, Li Y, Kaelin WG, Farnham PJ. A protein synthesis-dependent increase in E2F1 mRNA correlates with growth regulation of the dihydrofolate reductase promoter. Mol Cell Biol 1993;13:1610–1618 [author's correction: 13: 7201].
19. Maniatis T, Fritsch EF, Sambrook J. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1982.
20. Cross SH, Bird AP. CpG islands and genes. Curr Opin Genet Dev 1995;5:309–314.
21. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol 1987;196:261–282.
22. Bestor TH, Tycko B. Creation of genomic methylation patterns. Nat Genet 1996;12:363–366.
23. Laird PW, Jaenisch R. DNA methylation and cancer. Hum Mol Genet 1994;3:1487–1495.
24. Bucher P. Weight matrix descriptions of four eucaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J Mol Biol 1990;212:563–578.
25. Kollmar R, Sukow KA, Sponagle SK, Farnham PJ. Start site selection at the TATA-less carbamoyl-phosphate synthase (glutamine-hydrolyzing)/aspartate carbamoyltransferase/dihydroorotase promoter. J Biol Chem 1994; 269:2252–2257.
26. Kingsley C,Winoto A. Cloning of GT box-binding proteins: A novel Sp1 multigene family regulating T-cell receptor gene expression. Mol Cell Biol 1992;12:4251–4261.

27. Lania L, Majello B, De Luca P. Transcriptional regulation by Sp family proteins. Intl J Biochem Cell Biol 1997;29: 1313–1323.
28. Bello-Fernandez C, Packham G, Cleveland JL. The ornithine decarboxylase gene is a transcriptional target of c-Myc. Proc Natl Acad Sci USA 1993;90:7804–7808.
29. Peña A, Reddy CD, Wu S, et al. Regulation of human ornithine decarboxylase expression by the c-Myc:c-Max protein complex. J Biol Chem 1993;268:27277–27285.
30. Jones RM, Branda J, Johnston KA, et al. An essential E box in the promoter of the gene encoding the mRNA cap-binding protein (Eukaryotic Initiation Factor 4E) is a target for activation by c-myc. Mol Cell Biol 1996;16:4754–4764.
31. Galaktionov K, Chen X, Beach D. Cdc25 cell-cycle phosphatase as a target of c-Myc. Nature 1996;382:511–517.
32. Fisher F, Crouch DH, Jayaraman P-S, et al. Transcription activation by Myc and Max: Flanking sequences target activation to a subset of CACGTG motifs *in vivo*. EMBO J 1993;12:5075–5082.
33. Benvenisty N, Leder A, Kuo A, Leder P. An embryonically expressed gene is a target for c-Myc regulation via the c-Myc-binding sequence. Genes Dev 1992;6:2513–2523.
34. Boyd KE, Wells J, Gutman J, Bartley SM, Farnham PJ. c-Myc target gene specificity is determined by a post-DNA-binding mechanism. Proc Natl Acad Sci USA 1998;95: 13887–13892.