Expression of the TAF4b gene is induced by MYC through a non-canonical, but not canonical, E-box which contributes to its specific response to MYC

KWESI TEYE1, KENGO OKAMOTO1, YUJI TANAKA1, TOSHIYUKI UMATA2, MASAAKI OHNUMA3, MASAAKI MOROI3, HIROSHI KIMURA4 and MAKOTO TSUNEOKA1

1Faculty of Pharmacy, Takasaki University of Health and Welfare, Takasaki 370-0033; 2Radioisotope Research Center, University of Occupational and Environmental Health, Kitakyushu 807-8555; 3Division of Protein Biochemistry, Institute of Life Science, Kurume University, Kurume 839-0861; 4Chiba Institute of Science, Choshi 288-0025, Japan

Received May 19, 2008; Accepted August 4, 2008

DOI: 10.3892/ijo_00000118

Abstract. Transcription factor binding sites are short DNA sequences that interact with transcription factors and the proper control of gene expression appears to require the mechanisms including the regulation through the genome context around the transcription factor binding sites. The MYC proteins are central regulators of cell growth. Many genes have been reported to be regulated by MYC through E-box sites. However, the characters of E-box that Myc selects to function are not clear and identification of additional genes controlled by MYC will provide information to completely understand the functions of MYC. Here we report that MYC directly induces TAF4b expression. We mapped the transcription start site and characterized functional promoter elements for MYC response in the TAF4b promoter. There are several E-box sequences near the transcription start site, including canonical (CACGTG) and non-canonical (CGCGTG) ones. We found that c-MYC induces TAF4b expression through one of the non-canonical E-box sites, which is in a highly conserved region of TAF4b promoters in mammals, suggesting the importance of the genome context around the target E-box. When the non-canonical E-box in the TAF4b promoter was mutated to a canonical one, MYC functioned on both E-boxes, while another E-box-binding transcription factor, USF, did so only on the canonical E-box. These results suggest that in addition to the context where the target E-box exists, a sequence within an E-box is involved in the mechanisms by which specific E-box sites are selected by Myc.

Introduction

Transcription factor binding sites are short DNA sequences that interact with transcription factors. Although there are a great number of the candidate binding sites in the genome, the transcription factors function only on limited number of sites. Furthermore, while some transcription factors share their target sequences, the genes that respond to them and their biological functions are different. Thus, the proper control of gene expression requires the mechanisms, by which transcription factors specifically select their target sites. Recent studies suggest that the specificity of protein-DNA interaction depends, at least in part, on the genome context around the target sites. For example, the veracity of a potential target site correlates with the GC content and CpG island (1), preferential spacing between transcription factor binding sites and transcription start sites (2), co-operation of cis-regulatory molecules (3) and clustering of degenerate transcription factor binding sites near the target sites (4). However, it is still unclear how the target specificity for each transcription factor is achieved.

The myc family of proto-oncogenes consists of three main genes: c-myc, N-myc and L-myc (5-9). The three genes have different expression patterns but appear to have basically the same biological activity. The accumulating data indicate that the myc family genes are central regulators of cell growth (6,7,10-14). The MYC protein encoded by myc is a member of the basic helix-loop-helix leucine zipper transcription factors (6,14). MYC functions as an E-box-binding transcription factor and activates transcription of a number of genes (15-24). Although there are a number of E-box sites in the genome, Myc selects only some E-boxes to function on.

Besides proteins in the myc proto-oncogene family, there are other classes of E-box-binding transcription factors, including USF (25) and TFE3 (26). Although these factors also transactivate gene expression through an E-box, their biological functions and the genes that respond to them are different from Myc. However, it is unclear which factors are really responsible for the E-box-binding transcription factors to select the specific target genes.
Here we report that MYC directly induces the expression of TAF4b. The sequence responsible for this induction is a non-canonical E-box and the surrounding sequence of this element is highly conserved in mammals, confirming the importance of the context where E-box exists. Although this site responded specifically to MYC, it also responded to another E-box binding transcription factor, USF, when the non-canonical E-box was mutated to the canonical one. These results suggest that the sequence within E-box have a role for the selection of the target genes by E-box transcription factors.

Materials and methods

Cell culture. Human glioblastoma cell line T98G cells and its derivative T98GmycER-2 cells, which express c-MycER chimeric protein (23), were cultured in Eagle's medium supplemented with non-essential amino acids and 10% fetal calf serum (FCS). The African green monkey kidney fibroblast line Cos-7 were cultured in Dulbecco's modified Eagle's medium supplemented with 20% FCS. Human promyelocytic leukemia HL60 cells were cultured in RPMI-1640 medium supplemented with 10% FCS. Human erythroleukemia cell line HEL (Health Science Research Resources Bank, Osaka, Japan) were cultured in RPMI-1640 medium supplemented with 10% FCS.

Polymerase chain reaction (PCR). PCR amplifications were performed in 50 μl EX Taq buffer (Takara, Shiga, Japan) containing 10 pmol of each primer, 1.2 U EX Taq DNA polymerase in 50 μl EX Taq buffer (Takara, Shiga, Japan). PCR amplifications were performed using the primer TAF4bRACE-1 (5'-GGTCTCGGCT CTTGTTACAGTTTGCTG-3'), a sequence in the 5'-terminal region of TAF4b cDNA reported previously (27) and the AP1 primer provided by the supplier for the first PCR; and the TAF4bRACE-2 primer (5'-GGATTGTCGTGGTGTTGGG GCCTTTCA-3'), which corresponds to a sequence in the upstream region of TAF4bRACE-1 and the AP2 primer provided by the supplier for the nest PCR. The amplified DNA fragment was cloned into a pGEM-T vector (Promega) and sequenced.

Reverse transcriptase-polymerase chain reaction (RT-PCR). Synthesis of single-strand cDNA of HEL cells was performed on total RNA (1 μg) using a Superscript first-strand synthesis system (Invitrogen, Carlsbad, CA). One μl (total 20 μl) of the resultant single-strand cDNA was used as the template for PCR. The RT-PCR primers for the amplification were TAF4b coding U (5'-CGGAATTCAGCTCGAGAGGT CCGCCGGTGTGC-3') and TAF4b coding L (5'-CGGAA TTTCCGCAGTAATAATGCAAGGATGTGATGGA-3'), a sequence in the 3' untranslated region plus an EcoRI site) and TAF4b coding L (5'-CGGAA TTTCCGCAGTAATAATGCAAGGATGTGATGGA-3'), a sequence in the 3' untranslated region plus an EcoRI site). The temperature profile was 35 cycles of denaturing at 98°C for 15 sec, annealing at 65°C for 1 min and extension at 72°C for 2.5 min. The amplified DNA fragment was cloned into a pGEM-T vector (Promega) to produce pT/TAF4b and sequenced. The sequence was confirmed by direct sequencing using the primer provided by the supplier. The amplified DNA fragment was cloned into a pGEM-T vector (Promega) and sequenced.

Reporter plasmids with a TAF4b genomic DNA fragment. A genomic DNA fragment of the human TAF4b gene, which extends from the promoter region to part of exon 1, was amplified by PCR with 5'-TTTATTACATAACCTCACTTG TGGAGGGGG-3' and 5'-TGTGCTGGGACTTCTAGACCTACGGCG ACTATCCTG-3'. The 2.2-kb amplified fragment was cleaved with SacI and Ncol and inserted into the 4.8-kb fragment of pG3 containing firefly luciferase (Promega) cut with SacI and Ncol to produce pTAF4b(W)luci. The 0.6 kp fragment of TAF4b(W)luci cut with Smal and Ncol was inserted into the 4.8-kb fragment of pG3 cut with Smal and Ncol to produce pTAF4b(W)luci. The 0.6 kp fragment of TAF4b(W)luci cut with XhoI and Ncol was inserted into the 4.8-kb fragment of pG3 cut with XhoI and Ncol to produce pTAF4b(Xho)luci. pTAF4b(dSmal)luci was cleaved with XhoI and Ncol, bluntend with Klenow enzyme and self-ligated to produce pTAF4b(dSmal)luci. The 2.2-kb amplified fragment was cleaved with PvuII and Ncol, bluntend with Klenow enzyme and self-ligated to produce pTAF4b(Pvu-Nco)luci.

Further deletion mutants, pTAF4b(117Pvu)luci, pTAF4b(114Pvu)luci, pTAF4b(102Pvu)luci, pTAF4b(73Pvu)luci, pTAF4b(43Pvu)luci, pTAF4b(Sma118)luci, pTAF4b(Sma87)luci, pTAF4b(Sma74)luci, pTAF4b(Sma68)luci, pTAF4b(Sma59)luci and pTAF4b(Sma45)luci were produced by PCR with appropriate primers using pTAF4b(Sma-Pvu)luci as a template and the subsequent introduction of the resultant fragments into a pG3 vector.

Mutation was introduced at E-box sites of TAF4b(W)luci using a Gene Editor™ in vitro site-directed mutagenesis system (Promega) to produce pTAF4b(mE1/2)luci, in which two E-boxes (CAGGTG elements) were mutated to CACCGT. Mutation was also introduced at non-canonical E (NCE)-box sites (CGCGTG) of pTAF4b(Sma-Pvu)luci to produce...
Sequence of the rat and monkey TAF4b promoters. A 1.2 kb fragment of the rat TAF4b promoter was isolated as follows: First, the mouse DNA sequence homologous to the human TAF4b gene was selected (GenBank Accession no. AK012135). The two regions in the mouse DNA sequence whose nucleotide sequences exactly match the human TAF4b genomic sequences were selected, and oligonucleotides rat TAFII-F (5′-GGGCACGTGTGAAGCCCTGG-3′) and rat TAFII-P1 (5′-CTGACAGAGCTGCTGAC-3′) were synthesized. With these oligonucleotides as primers, the rat TAF4b genomic DNA was amplified using rat colon cancer cell RCN-9 genomic DNA as a template. The amplified DNA fragment was sequenced and the upstream region of this sequence was further cloned by promoter find kit (Clontech) with DraI and SspI libraries as templates using primers rat TAFII-P3 (5′-GCCGGGGCGGGCAGGAGCTC-3′) and rat TAFII-P4 (5′-GGCGACTCCCTAGCCGGCTC-3′), whose sequences were in the rat TAF4b DNA sequence identified above and AP1 and AP2. The amplified rat TAF4b genomic DNA fragments were sequenced.

The monkey TAF4b promoter was amplified with TAF-primer 1 (5′-GACACAGAGGAACGGAGGAAGAGCTC-3′) and TAF-primer 2 (5′-GACACAGAGGAACGGAGGAAGAGCTC-3′) matching the human TAF4b gene using the genomic DNA of Cos-7 cells as a template and sequenced.

Other plasmids. pRL-CMV containing the Renilla reniformis luciferase gene under control of the CMV promoter was purchased from Promega. CMV-c-Myc, pCMV-USF and pCMV-TFE-3 were gracious gifts from Dr Martin Eilers (28).

RNA preparation. Total RNA was isolated from cells by the acid guanidium thiocyanate-phenol/chloroform extraction method using a DEPC-treated RNA preparation solution set (Nakalai, Kyoto, Japan).

Differential display using DNA chip (cDNA micro array). cDNA micro array was performed as described previously (23). Total RNA was isolated from untreated T98Gmycer-2 cells and from cells treated by 4-hydroxytamoxifen (OHT) for 20 h. Poly(A)+ RNA was recovered and subjected to differential display using a DNA chip (Incyte Genomics, Palo Alto, CA). About 9,000 kinds of cDNAs including expressed sequence tag (EST) clones were plated on the chip (UniGEM Human V Ver. 2).

Northern blot analysis and preparation of probe DNAs. RNA was electrophoresed in an agarose gel containing formaldehyde, transferred to Hybond-N (Amersham Biosciences) and probed with 32P-labeled cDNA. Probes were labeled with [α-32P]-dCTP using a Multiprime labeling kit (Amersham Biosciences). The results were quantified using a BAS2000 image analyzer (Fuji Photo Film, Tokyo, Japan). The DNA probe for c-myc was described previously (18). The DNA probe for TAF4b was a 0.6 kb SacI fragment from pT/fragmentTAF4b.

Transient expression assay. Transient expression reporter assay in T98Gmycer-2 cells was performed as described previously (23). The effects of MYC and other transcription factors on the TAF4b promoter were investigated by transfection of the plasmids encoding the transcription factors. T98G cells were grown in medium supplemented with 10% FCS. Cells (3x10⁴) were plated into a dish (12-well plate; well diameter 22 mm) and cultured for 20-24 h. Transfections were carried out using FuGENE 6 reagent (Roche Diagnostics, Indianapolis, IN) with 1 μg of the reporter plasmids, 20 ng of pRL-CMV as an internal transfection marker and various amounts of plasmids expressing the transcription factors. The total amount of DNA in each transfection was kept constant by adding the empty expression vector. One day later, cells were collected and the luciferase activities were determined as described previously (18).

Chromatin immunoprecipitation. Chromatin immunoprecipitation assay was performed basically as described previously (23). Immunoprecipitated DNA fragments were detected by PCR. The PCR primers were 5′-GACACAAAGGAGGAACGGGATC-3′ (TAF4b-primer 1) and 5′-GGCCGGGGGAGGGAGCGTAC-3′ (TAF4b-primer 2), which amplify a 521-bp fragment containing the transcription start site of the human TAF4b gene (TAF4b primers) and 5′-TTACAGGTAAGCCCTCCAATGACC-3′ and 5′-GCCGGGCTGCCCCCGGAGCGA CAC-3′, which matching the human TAF4b gene using the genomic DNA of Cos-7 cells as a template and sequenced.

Antibodies. Antibody-1 against c-MYC was described previously (18). Anti-c-MYC antibody-2 (N262) (Santa Cruz Biotechnology, Santa Cruz, CA) was purchased.

The sequence information. The sequence data in this study have been submitted to the DDBJ databases under accession No. AB234096 for the DNA sequence of full-length human TAF4b, accession No. AB234097 for the DNA sequence of the human TAF4b promoter, accession No. AB234098 for the DNA sequence of the monkey TAF4b promoter and accession No. AB234099 for the DNA sequence of the TAF4b promoter (Cos-7 cells).

Results

c-MYC induces the expression of TAF4b. To conditionally induce c-MYC activity, the estrogen-inducible MYC system (23,29) was used. The chimeric protein c-MycER consists of human c-MYC and the estrogen-binding domain of the human estrogen receptor. c-MycER anchors to cytoskeletal components of cells in the absence of estrogen. When estrogen or its analogous molecule 4-hydroxytamoxifen (OHT) binds to the chimeric protein, it becomes free to function as c-MYC. The human glioblastoma cell line T98G was used as the
containing a chromosome 22 E-box (control primers) (20). (TAF4b primers) or primers to specifically amplify a DNA fragment designed to amplify a DNA fragment of the TAF4b promoter region. Following DNA purification, samples were subjected to PCR with primers which no nuclear extract was added, was also performed (nuclear extract -). expression was greatly reduced (Fig. 2A). Mock immunoprecipitation, in control antibody, or no antibody (no antibody). In the presence of TPA, c-anti-c-MYC antibody-2 reacting with the C-terminal half of c-MYC), (anti-c-MYC antibody-1 reacting with the N-terminal half of c-MYC and chromatins were immunoprecipitated using antibodies against c-MYC (+TPA) were fixed with 1% formaldehyde, nuclear extracts were collected, proliferating phase (TPA) and those treated with 10 nM TPA for 24 h (CAGCTG), and a TATA box sequence (shaded), a non-canonical E (NCE)-box sites, NCE-box 1 and NCE-box 2 (underlined), E-box sites, E-box 1 and E-box 2 (surrounded by dotted lines), translation start sites are indicated by an arrowhead and an asterisk, sequence of the N-terminal end, which had not been reported, is also shown sequence of the human TAF4b promoter and part of exon 1. The amino acid sequence was deposited in GenBank as accession no. AB234096). The cDNA clone encodes a protein of 862 amino acids with a predicted molecular weight of 91090.52 Da and PI of 9.59.

Expression of TAF4b mRNA. Human promyelocytic leukemia HL60 cells are terminally differentiated by 12-O-tetradecanoylphorbol-13-acetate (TPA), in which the c-myc expression level decreases (30,31), after a transient stimulation (32). This experimental system was used to investigate whether MYC-target genes are affected during the shut-off of MYC (33,34). As shown in Fig. 2A, the level of c-myc mRNA had doubled at 1 h after addition of TPA, then started to decline at 3 h, and reached one-fifth at 7 h. The level of TAF4b mRNA had increased to one and a half-fold at 3 h after addition of TPA. At 6 h it started to decline, and reached one-tenth at 9 h. These results show that the expression of TAF4b mRNA closely followed that of c-myc.

Next, the effect of c-MYC activation on TAF4b mRNA was investigated (Fig. 2B) using the MycER system. The level of TAF4b mRNA had increased to 2- and 3-fold at 3 and 7 h, respectively, in OHT-treated T98Gmycer-2 cells. The OHT treatment of T98G parent cells did not stimulate the TAF4b mRNA level. As shown in Fig. 2C, the level of c-myc mRNA had increased by 2.6-, 1.6- and 1.9-fold, respectively, with c-MycER activation. These results suggest that MYC target genes could be detected in this experimental system.

We found that the signal for EST clone AA287145 was stimulated by 2.3-fold with c-MYC activation. cDNA encoding the 5' upstream part of EST clone AA287145 was isolated using the 5'-RACE protocol from a library of human erythroid leukemia (HEL) cells, and it was found that AA287145 was part of the 3' non-coding region of the TAF4b gene (previously called TAF2.105) (27). Because the translation and transcription start sites of the human TAF4b gene had not been reported, we isolated cDNA encoding the 5'-end of the transcript by 5'-RACE experiments. The transcription start site was mapped on the TAF4b genomic sequence, which is indicated by an arrowhead in Fig. 1A. Using primers consistent with the 5' and 3' sequences of TAF4b mRNA, a 3.2-kb length cDNA was amplified from total RNA of HEL cells by RT-PCR protocol and sequenced. Although the entire nucleotide sequence was almost the same as the previously reported one (27), we found an additional 66-amino acid sequence extending from the N-terminal amino acid glycine of TAF4b reported previously (Fig. 1A), and several amino acids were different (the sequence was deposited in GenBank as accession no. AB234096). The cDNA clone encodes a protein of 862 amino acids with a predicted molecular weight of 91090.52 Da and PI of 9.59.

Figure 1. The genomic DNA sequence of the human TAF4b promoter and chromatin immunoprecipitation experiments. (A) The genomic DNA sequence of the human TAF4b promoter and part of exon 1. The amino acid sequence of the N-terminal end, which had not been reported, is also shown under the DNA sequence with one letter notation. The transcription and translation start sites are indicated by an arrowhead and an asterisk, respectively. E-box sites, E-box 1 and E-box 2 (surrounded by dotted lines), non-canonical E (NCE) box sites, NCE box 1 and NCE box 2 (underlined), TATA box sequence (shaded), a Smal site (CCCGGG), a PvuII site (CAGCTG), and a XhoI site (CTCGAG) are also shown. (B) Chromatin immunoprecipitation using anti-c-MYC antibody. HL60 cells in the proliferating phase (TPA) and those treated with 10 nM TPA for 24 h (+TPA) were fixed with 1% formaldehyde, nuclear extracts were collected, and chromatins were immunoprecipitated using antibodies against c-MYC (anti-c-MYC antibody-1 reacting with the N-terminal half of c-MYC and anti-c-MYC antibody-2 reacting with the C-terminal half of c-MYC), control antibody, or no antibody (no antibody). In the presence of TPA, c-myc expression was greatly reduced (Fig. 2A). Mock immunoprecipitation, in which no nuclear extract was added, was also performed (nuclear extract -). Following DNA purification, samples were subjected to PCR with primers designed to amplify a DNA fragment of the TAF4b promoter region (TAF4b primers) or primers to specifically amplify a DNA fragment containing a chromosome 22 E-box (control primers) (20).

parental cell into which ectopic c-MYC activity was introduced. A T98G cell line expressing c-MycER protein (T98Gmycer-2 cells) was established previously (23). Total RNAs from T98Gmycer-2 cells in an exponentially growing phase in the presence or absence of OHT for 20 h were subjected to cDNA micro array analysis. Specific signals for MYC target genes, ornithine decarboxylase (17), nucleolin (19) and mina53 (23) were increased by 2.6-, 1.6- and 1.9-fold, respectively, with c-MycER activation. These results suggest that MYC target genes could be detected in this experimental system.

My Gene of Interest...
Figure 2. Expression of TAF4b mRNA. (A) The changes of TAF4b and c-myc mRNA in human promyelocytic leukemia HL60 cells after TPA treatment. RNA was isolated at the indicated time points after exposure to 10 nM TPA and analyzed by Northern blotting to detect TAF4b and c-myc mRNAs (left). 28S and 18S ribosomal RNAs are also shown. The results were quantified and plotted (right). (B) Increase of TAF4b mRNA levels in T98G cells by activation of MycER protein. After T98Gmycer-2 cells (T98GMyER) and parental cells (T98G) were cultured in the medium supplemented with 0.25% serum for 40 h, cells were cultured in the presence (+) or absence (-) of 200 nM OHT. RNA was isolated at the indicated time points and analyzed by Northern blotting to detect TAF4b mRNA (left). 28S and 18S ribosomal RNAs are also shown. The results were quantified and expressed as a bar graph (right). (C) The increase of TAF4b mRNA levels due to activation of MycER protein was maintained in the presence of the protein synthesis inhibitor. After T98Gmycer-2 cells were cultured in the medium supplemented with 0.25% serum for 40 h, cells were treated with 200 nM OHT for 8 h. RNA was isolated and analyzed by Northern blotting to detect TAF4b mRNA (left). When indicated as + in the row of CHX, 20 μg/ml cycloheximide was added 20 min before the addition of OHT. 28S and 18S ribosomal RNAs are also shown. The results were quantified and expressed as a bar graph (right).
endogenous TAF4b gene fragment in each sample was monitored by PCR using primers that specifically amplify DNA around the transcription initiation site in the TAF4b gene. As shown in Fig. 1B, two different anti-c-MYC antibodies immunoprecipitated the TAF4b DNA fragment from HL60 cells in the proliferating phase, while the same antibodies did not immunoprecipitate detectable levels of DNA fragments from cells treated with TPA, in which c-myc expression became very low. Enrichment of TAF4b genomic DNA fragments is dependent on c-MYC binding to the TAF4b gene because a nonspecific antibody did not immunoprecipitate TAF4b DNA fragments. Additionally, binding of c-MYC detected in the TAF4b gene is specific, since antibodies against c-MYC did not enrich the genomic DNA fragment containing an E-box that is located in a chromosomal region without any detectable gene in chromosome 22. These results demonstrate that the TAF4b gene around the transcription initiation site is bound by c-MYC specifically at the proliferating phase of HL60 cells.

**Figure 3.** Transient expression from the TAF4b promoter. (A) The reporter plasmid [pTAF4b(W)luci] was constructed to detect the TAF4b promoter activity. In pTAF4b(mE1/2)luci, two E-boxes (CACGTG elements) of the TAF4b promoter were mutated to CACCTG (see Fig. 1A). Deletion mutants of pTAF4b(W)luci, pTAF4b(dSma)luci, pTAF4b(dXho)luci, pTAF4b(Sma-Xho)luci, and pTAF4b(Sma-Pvu)luci were also produced. T98Gmycer-2 cells were transfected with the reporter plasmids. One day later, OHT was added at the final concentration of 200 nM to activate the MycER chimeric protein, and cells were further cultured for 17 h. Cell extracts were analyzed for activities of firefly luciferase. The promoter activities were determined by an arrowhead in Fig. 3A. The transcription initiation site was at -460 bp, which is indicated by an arrowhead in Fig. 1A. The TATA box sequence was found ~30 bp upstream of the transcription initiation site.

To investigate the activity of the TAF4b promoter, the genomic DNA fragment including a 1.6-kb region from the upstream region of exon 1 to part of exon 1 was ligated to firefly luciferase cDNA to construct a reporter plasmid, pTAF4b(W)luci (Fig. 3A). Transient expression assays in T98Gmycer-2 cells indicated that the DNA fragment has promoter activity. After c-MycER activation by OHT, the luciferase activity was increased up to 2.5-fold (Fig. 3A). This stimulation is comparable to that observed for an increase in the TAF4b mRNA level by OHT in T98Gmycer-2 cells (Fig. 2B).

There are two canonical E-box sites (CACGTG elements) near the transcription initiation site. We designated the first and second CACGTG elements as E-box 1 and E-box 2, respectively. Since it had been reported that CACGTG elements function as MYC-responsive elements in many MYC-responsive genes, we examined whether the E-box sites were involved in activation of the promoter activity by c-MYC. When E-boxes 1 and 2 were mutated, however, luciferase activity was still increased by MYC as in the wild-type reporter plasmid (Fig. 3A, lane for pTAF4b(mE1/2)luci).

Next, the region responsible for MYC stimulation was investigated by using deletion mutants (Fig. 3A). While deletion of the region between -1575 and -535 [pTAF4b(dSma)luci ] did not decrease stimulation of the promoter by MYC, deletion of the DNA fragment between -1575 and -367 [pTAF4b(dXho)luci] abolished MYC stimulation, suggesting the importance of the region between -534 and -367. Then we constructed two additional reporters, which contained the sequences between -534 and -367 [pTAF4b(Sma-Xho)luci] and -534 and -377 [pTAF4b(Sma-Pvu)luci], respectively. The expression of these two reporters was well stimulated by MYC.
To find the responsible element for MYC-activation, the DNA fragment in the reporter \([pTAF4b(Sma-Pvu)luci]\) was deleted sequentially from the 5’-end of the TAF4b promoter (Fig. 3B). While the deletion of the region between -534 and -492 \([pTAF4b(117Pvu)luci]\) marginally reduced the elevation by MYC, deletion of -534 to -490 \([pTAF4b(114Pvu)luci]\) abolished the elevation by MYC. The expression by the reporters with further deletions \([pTAF4b(102Pvu)luci, pTAF4b(73Pvu)luci, pTAF4b(43Pvu)luci]\) was not elevated by MYC. These results suggest the importance of the region between -534 and -490, especially between -492 and -490.

The region between -492 and -490 is in a non-canonical E-box element CGCGTG (-492 to -487 in Fig. 1A). It has been reported that c-MYC could bind this sequence (6). Then we tested whether CGCGTG elements were involved in activation by MYC. There are two CGCGTG elements very near the TATA box sequence (Fig. 1A). We designated the first and second CGCGTG elements as non-canonical E (NCE)-box 1 (-509 to -504) and NCE-box 2 (-492 to -487), respectively. We constructed the reporter plasmids \(pTAF4b(Sma-Pvu, mNCE1)luci, pTAF4b(Sma-Pvu, mNCE2)luci\) and \(pTAF4b(Sma-Pvu, mNCE1/2)luci\), whose one or both CGCGTGs were mutated to CGCCTG (Fig. 3C). While mutation of NCE-box 1 had little effect on the stimulation of promoter activity by c-MYC, mutation of NCE-box 2 or of both NCE-box 1 and NCE-box 2 completely abolished the MYC-stimulation. These results indicate that c-MYC transactivates the TAF4b promoter through NCE-box 2.

The response of CGCGTG and CACGTG elements to E-box binding transcription factors. In many genes, a sequence of MYC responsive E-box is CACGTG, rather than CGCGTG. To test whether the sequence CGCGTG has any biological significance, NCE-box 2 (CGCGTG) was mutated to the canonical E-box (CACGTG). The promoter with CACGTG element was stimulated by c-MYC and the elevation was higher than that of the wild-type promoter (Fig. 4A).

Besides proteins in the \(myc\) proto-oncogene family, there are other E-box-binding transcription factors, including USF (25) and TFE3 (26). Then we examined how these E-box-binding transcription factors responded to CACGTG and CGCGTG elements. As shown in Fig. 4B and C, the wild-type TAF4b promoter was hardly stimulated by TFE3 and USF. These results suggest that the TAF4b promoter specifically responds to MYC. However, when the CGCGTG element was mutated to the canonical CACGTG, the expression from the TAF4b promoter was stimulated by USF. TFE3 still hardly stimulated the expression from the reporter.

The MYC-responsive E-box element and the surrounding sequence in the TAF4b promoter is conserved in mammals. Since the sequence between -493 and -461 appears to be critically important for MYC activation, it was curious to test whether this sequence also exists in other species. We isolated a 1.2 kb fragment of the rat TAF4b genomic DNA containing the promoter region, and sequenced it. Fig. 5A shows the difference in the nucleotide sequences between the human and rat TAF4b promoter regions. The nucleotide diversity was calculated using the average number of nucleotide differences in sliding and overlapping windows of 100 bp, placed at 10 bp intervals. In this analysis, an insertion/deletion (indel) was treated as a single mutation. The minimum values of nucleotide diversity, namely the strongest homology, were observed around NCE-box 2 (location 792). Low values of nucleotide diversity were also observed around location 1000. This may be attributed to the fact that, in those locations, there are several large indels, which caused the nucleotide diversity to be underestimated. When the human TAF4b promoter sequence was compared to that of mouse, whose DNA sequence was recently deposited in the GenBank (accession no. AC133172.4), similar results were obtained (data not shown).

We isolated the TAF4b promoter DNA fragment of monkey from African green monkey kidney fibroblast Cos-7 cells and sequenced it. The 29 bp sequence containing NCE-box 2 and the TATA box element are completely conserved among the four species studied.
The MYC responsive E-box in the TAF4b promoter is a non-canonical (CGCGTG) element. When the CGCGTG element was mutated to the canonical CACGTG, c-MYC still up-regulated the promoter activity (Fig. 4A). While USF, another E-box-binding transcription factor, hardly stimulated the wild-type TAF4b promoter, it up-regulated the activity of the promoter with the CACGTG element (Fig. 4C). These results suggest that in addition to the context where the target E-box exists, a sequence within an E-box is involved in the mechanisms by which the E-box site is selected specifically by MYC.

When the CGCGTG element was mutated to the canonical CACGTG, c-MYC more strongly up-regulated the promoter activity (Fig. 4A). It was shown that c-MYC binds to CACGTG E-box element with higher affinity than other E-boxes (37). The decreased sensitivity of the non-canonical E-box with lower affinity to MYC may contribute to the proper gene activation by MYC. This idea is consistent with a recent study, which indicates that the affinity of binding sites is a primary determinant of the level of physiological stimulus necessary for substantial gene activation, contributing to fine-tuning the responses (38). TFE3, another E-box-binding transcription factor, activated the TAF4b promoter with neither CACGTG element nor CGCGTG element (Fig. 4B). Together, these results indicate that the sequences within E-box differently affect the activities of E-box-binding transcription factors and function as a factor to determine the spectrum of genes stimulated by the transcription factors.

**Discussion**

c-MYC induces TAF4b expression through an E-box site, which is in a highly conserved region of TAF4b promoters in mammals. We found evidence that MYC directly induces the expression of TAF4b. After treatment with TPA, the expression of TAF4b mRNA closely followed that of c-myc in HL60 cells (Fig. 2A). Stimulation of the expression of TAF4b mRNA was observed when c-MYC was activated in the c-MycER chimeric protein (Fig. 2C), indicating that c-MYC directly controls the expression of TAF4b. c-MYC protein bound the TAF4b promoter in vivo in proliferating HL60 cells (Fig. 1B). The TAF4b promoter activity was elevated by c-MYC (Fig. 3). This elevation was mediated through an E-box (NCE-box 2; a TAF4b promoter activity was elevated by c-MYC (Fig. 3).

All E-boxes are not always MYC-responsive elements. The MYC-responsive E-box in the TAF4b promoter exists in the most conserved region between human and rat (Fig. 5A) where GC content is high (Fig. 1A). The 29 bp-nucleotide sequence covering the MYC-responsive element and TATA box is completely conserved in human, monkey, rat, and mouse (Fig. 5B), suggesting that the distance from the E-box site to the transcription start site is conserved between species. The high conservation of the DNA sequence suggests that the interacting proteins on the DNA region are conserved. There are several E-boxes besides the MYC-responsive E-box near the promoter region. Thus, the characteristics proposed to be used for specific protein-DNA interaction, including high GC content (1), preferential spacing between transcription factor binding sites and transcription start sites (2), co-operativity of cis-regulatory molecules (3), and clustering of transcription factor binding sites near the target sites (4), are observed around this E-box site in the TAF4b promoter and may be involved in the mechanisms by which NCE-box 2 specifically functions as a MYC-target site.**

**TAF4b is a Myc-target gene.** In eukaryotes, protein-encoding genes are transcribed by RNA polymerase II (RNA Pol II). The molecular machinery that guides RNA Pol II to initiate transcription of a specific gene is composed of multiple classes of regulatory proteins (39). TFIIID is a core component of the RNA Pol II regulatory machinery, which is conserved from *Saccharomyces cerevisiae* to humans and is a large multi-protein complex composed of the TATA box-binding protein (TBP) and several TBP-associated factors (TAFs). Although TFIIID was initially thought to be ubiquitous in expression and function, identification of tissue-specific TAFs suggested that a specialized TFIIID complex could play a direct role in regulating tissue-specific programs of gene expression. The first cell type-specific TFIIID subunit identified was TAF4b...
(previously called TAF4b(10S) (27), which is a homologous protein of a ubiquitously expressed factor, TAF4 (previously called TAF4b(30)). The comparison between the three dimensional structure of the TFIIID complex with TAF4b and that with original TAF4 suggest that distinct localized structural changes in TFIIID may cause a difference in their activity (40). Indeed, it was reported that TAF4b and TAF4 proteins share a common non-redundant function in regulating gene expression control (41,42). These findings show that the expression control of TAF4b is a physiologically important issue. However, there have been no studies on the mechanisms of the TAF4b expression control.

It is well established that MYC is a central regulator of cell growth. There are studies describing functions and expression of TAF4b, which suggest the functional and expression control between TAF4b and MYC. In adult mouse testes, TAF4b was required for germ cell proliferation (43), and the expression of TAF4b and c-myc was detected in premeiotic spermatogonia (43-45). When primary B lymphocytes were activated by bacterial lipopolysaccharide (LPS), expression of TAF4b (46) and c-myc (47) was increased. TAF4b can induce a number of genes involved in the control of cell growth (41). It was also reported that TAF4b elevates the expression of VEGF (42), while vasculogenesis and angiogenesis were induced by MYC, which can indirectly induce the expression of VEGF (48). Therefore, TAF4b may be involved in the induction of VEGF expression by MYC (49). Together, the above studies suggest that the induction of TAF4b induced by MYC may contribute to cell growth control and tumorigenesis.

Acknowledgements

We thank Martin Eilers (Philips-Universität Marburg) for providing pCMV-USF and pCMV-TFE3. We thank Kathryn E. Boyd and Peggy J. Farnham (University of Wisconsin Medical School) for sharing their protocols for chromatin immunoprecipitation analysis. This study was supported by Fukuoka Gann, Sankyo-Foundation and Kathryn E. Boyd and Peggy J. Farnham (University of Wisconsin Medical School) for sharing their protocols for chromatin immunoprecipitation analysis. We thank Martin Eilers (Philips-Universität Marburg) for providing pCMV-USF and pCMV-TFE3. We thank Kathryn E. Boyd and Peggy J. Farnham (University of Wisconsin Medical School) for sharing their protocols for chromatin immunoprecipitation analysis. This study was supported by Fukuoka Gann, Sankyo-Foundation and Grants-in-Aid from the Ministry of Education, Science and Culture of Japan.

References


