TFEC Can Function as a Transcriptional Activator of the Nonmuscle Myosin II Heavy Chain-A Gene in Transfected Cells

Myung-Chul Chung,[‡] Hyung-Kwoun Kim,[‡,§] and Sachiyo Kawamoto*

Laboratory of Molecular Cardiology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-1762

Received December 15, 2000; Revised Manuscript Received April 27, 2001

ABSTRACT: Transcription of the human nonmuscle myosin II heavy chain-A (NMHC-A) gene is regulated via multiple elements located in intron 1, including element F which contains an E-box. In this study we have identified and characterized the factors that are capable of binding to element F. Yeast one-hybrid screening using element F allowed isolation of cDNAs encoding transcriptional factors TFEC, TFE3, and USF2, each of which contains basic helix−loop−helix and leucine zipper motifs. Furthermore, cDNA cloning by polymerase chain reaction yielded cDNAs for two TFEC isoforms, designated TFEC-l and TFEC-s, which are generated by alternative pre-mRNA splicing. In addition to these four factors, USF1, which is known to share the same DNA binding elements with USF2, was isolated for comparison. Electrophoretic mobility shift assays and cotransfection studies of the expression constructs with reporter gene constructs revealed that the above five factors have different binding activities for element F with different transactivation potencies. USF1 and USF2 demonstrate the highest binding activity to element F, yet show the lowest element F-dependent transactivation. TFE3 has a high transactivation potency but the lowest binding activity. TFEC-l demonstrates a high binding activity with the highest transactivation potency, whereas TFEC-s has the same binding activity as TFEC-l with intermediate transactivation. We also demonstrate that an N-terminal activation domain exists only in TFEC-l, whereas a C-terminal activation domain is common to both the l and s isoforms. This study provides the first evidence of TFEC being an activator of transcription, with two separate activation domains.

Myosin is a family of mechanochemical proteins that contain a conserved ∼80 kDa motor domain which can bind to actin, hydrolyze ATP, and translocate along actin filaments (*1*, *2*). Conventional myosin (myosin II in the classification of the myosin superfamily) consists of a pair of heavy chains (∼200 kDa) and two pairs of light chains (15−20 kDa). Myosin II is present in all eukaryotic cells and appears to be involved in diverse cellular contractile and motile processes, such as muscle contraction, cytokinesis, cell migration, and cell attachment (*3*). In higher organisms, different types of cells contain different isoforms of myosin II that contain different myosin II heavy chains (MHCs).[1] In vertebrates, over 10 genes encode MHCs, and they are divided into two subgroups, i.e., sarcomeric (skeletal and cardiac muscles) and nonsarcomeric (smooth muscle and nonmuscle) MHC genes. There are at least eight genes for sarcomeric MHCs, and expression of these genes is regulated developmentally, hormonally, and in a muscle fiber type-specific manner. This regulation occurs mainly at a transcriptional level, and a number of muscle-specific enhancers and transcriptional factors have been identified (*4*−*7*).

For nonmuscle MHCs, we and others demonstrated the existence of at least two genes, referred to as nonmuscle MHC-A (NMHC-A) and NMHC-B (*8*−*10*). In the case of NMHC-B, an alternative pre-mRNA splicing mechanism is also utilized to generate additional NMHC-B isoforms (*10*, *11*). The two NMHC mRNAs are expressed in a variety of tissues, but the relative amounts of the two mRNAs vary among different tissues (*12*−*14*). NMHC-A is the dominant isoform in spleen, thymus, and epithelium, whereas NMHC-B is dominant in brain. Lung and kidney contain approximately equal amounts of each NMHC mRNA. In contrast, both NMHC-A and -B mRNAs are barely detected in fully differentiated skeletal muscles where the sarcomeric MHC is dominant. Serum and other mitotic stimulants change the expression of these genes differently (*12*, *15*). Accumulating evidence shows that the expression of specific NMHC isoforms is dependent on cell type and is linked to cell proliferation and differentiation. However, the regulatory mechanisms controlling NMHC gene expression are poorly understood.

We have previously cloned and characterized the promoter region of the human NMHC-A gene (*16*). The structure of this region shows many features typical of a housekeeping gene. There is no TATA element and the GC content is high, with multiple GC boxes. We have also reported that the region just downstream from the transcriptional start sites (between +62 and +257, where +1 is a transcriptional start

* Address correspondence to this author at the National Institutes of Health, Building 10, Room 8N202, 10 Center Drive, MSC 1762, Bethesda, MD 20892-1762. Telephone: 301-496-1912. Fax: 301-402-1542. E-mail: KawamotS@NHLBI.NIH.GOV.

[‡] Both of these authors made major contributions to this study.

[§] Present address: Environmental Bioresources Lab, KRIBB, P.O. Box 115, Yusong, Taejon 305-600, Korea.

[1] Abbreviations: NMHC, nonmuscle myosin heavy chain; bHLH, basic helix−loop−helix; LZ, leucine zipper; AD, activation domain; PCR, polymerase chain reaction; DBD, DNA binding domain; EMSA, electrophoretic mobility shift assay; CAT, chloramphenicol acetyltransferase; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; aa, amino acid(s).

10.1021/bi002847d
Published on Web 06/28/2001

site) is involved in cell type-specific activation of NMHC-A gene expression via both pretranslational (transcription and/or mRNA stability) and translational mechanisms. Furthermore, we have recently identified three clustered cis elements (C, A, and F) in intron 1 located 23 kb downstream from the transcriptional start sites, which modulate transcription in a cell type- and differentiation state-dependent manner (*17*). Element C is recognized specifically by Sp1 and Sp3 transcriptional factors. However, the factors, which bind to the other two elements, remain to be identified. Of note, however, is that element F includes an E-box sequence.

The E-box sequence consists of the hexanucleotide CANNTG, where N is A, T, G, or C. The E-box sequence is typically recognized by a large family of transcriptional factors with a common structural feature, termed the basic helix−loop−helix (bHLH) motif (*18*, *19*). Many genes include an E-box-containing element in their promoters and enhancers, and an E-box-containing element in any specific gene is generally recognized by a few specific bHLH proteins. Some of the bHLH proteins such as the MyoD family or the myc family have been previously demonstrated to play important roles in cell lineage determination, cell proliferation, and differentiation (*20*, *21*). Some bHLH proteins have been identified by homology of their common bHLH motif; however, their target genes or their functional properties have not been well characterized.

In this study, we have focused on element F of the NMHC-A gene. We have identified and characterized the factors which are capable of binding to element F. The basic helix−loop−helix leucine zipper (bHLH-LZ) proteins, TFEC-l and -s, which are alternatively spliced isoforms, TFE3, USF1, and USF2 have all been found to bind to element F with different binding activities and with different transcriptional activation potencies. Among these factors, TFEC-l showed the highest transactivation potency for element F-dependent transcription. We further demonstrate that TFEC-l contains two activation domains, one of which is absent in TFEC-s. This study provides the first evidence that TFEC is a sequence-specific transcriptional activator and that NMHC-A is its target gene.

EXPERIMENTAL PROCEDURES

Yeast One-Hybrid Screening. A target reporter construct and a reporter yeast strain were prepared by using a matchmaker one-hybrid system (Clontech). Three tandem copies of element F with flanking sequences (5′-AATGTGTCAGGTGATT-3′) were prepared by annealing two complementary strands of synthetic oligonucleotides, which include the cohesive ends of *Eco*RI and *Xba*I, and inserted upstream of a reporter gene *HIS3* in pHISi. The resulting target-reporter construct was integrated at the *HIS3* locus of yeast strain YM4271, yielding a reporter yeast strain. This reporter yeast strain was used as a host strain for the library screening. Human leukocyte and HeLa cell cDNA libraries fused to the GAL4 activation domain (AD), which were constructed in yeast expression vectors pGAD10 and pGADGH (Clontech), were screened according to the manufacturer's instructions. Selection media included 10 mM 3-amino-1,2,4-triazole to inhibit the background *HIS3* expression.

Plasmid Construction. Plasmid pCITE-4a(+) (Novagen), which contains the T7 promoter, a cap-independent translation enhancer sequence, and S-tag and His-tag coding sequences, was used for in vitro transcription and translation. pCS3+MT (*22*), which contains the CMV promoter, six copies of myc-tag coding sequence, and an SV40 polyA transcription termination signal, was used for eukaryotic expression. The full-length coding regions of the cDNAs for TFEC-l, TFEC-s, TFE3, USF1, and USF2 were obtained from human leukocyte or HeLa cell cDNA libraries (Clontech) by polymerase chain reaction (PCR) using appropriate primers and inserted into the pCITE vector. The full-length coding regions with a His-tag coding sequence were then transferred from pCITE to pCS3+MT. TFEC-l, TFEC-s, USF1, and USF2 include five copies of myc-tag, and TFE3 includes six copies. A plasmid pM (Clontech), which contains an SV40 promoter, a GAL4 DNA binding domain (DBD) coding sequence (aa 1−147), and an SV40 polyA transcription termination signal, was used as a host vector to make constructs for the GAL4DBD-TFEC hybrid proteins. Defined fragments of TFEC were generated by PCR and inserted 3′ to the GAL4DBD. The fidelity of all constructs was verified by DNA sequencing.

Luciferase reporter genes were constructed in a promoterless and enhancerless reporter vector pGL2 basic (Promega). The core promoter of the human NMHC-A gene used for reporter constructs corresponds to the region between −112 and +61, where +1 is a major transcriptional start site (*16*). The 120 bp fragment of the intronic enhancer region, which contains three binding elements, A, C, and F, with the wild-type or mutated elements, was inserted upstream of the core promoter. Construction and sequences of these constructs have been described (*17*). For this study, the construct containing the 120 bp fragment with wild-type element F and the mutated elements A and C and that with mutations in all three elements are used as reporters. A reporter plasmid pG5CAT, which contains five copies of the GAL4 binding site and the adenovirus E1b minimal promoter upstream of the chloramphenicol acetyltransferase (CAT) gene, and a positive control plasmid pM3-VP16, which expresses a fusion of GAL4DBD to the VP16AD, are from Clontech. pCMV β-galactosidase (pCMV β-Gal) and pCS3+MT were gifts from Dr. Yongsok Kim (NHLBI).

In Vitro Synthesis of Proteins and Electrophoretic Mobility Shift Assay (EMSA). The full-length proteins with S- and His-tags were synthesized in vitro from the pCITE constructs using a TNT quick-coupled transcription/translation system (Promega) in the presence or absence of L-[^{35}S]methionine (Redivue, 1000 Ci/mmol, Amersham). The unlabeled proteins in the reticulocyte lysate reaction mixture were directly used for EMSA.

The wild-type element F, mutant element F, and element μE3 target DNAs used for EMSA were prepared by annealing two complementary strands of oligonucleotides. The upper strand of sequence of each oligonucleotide is as follows: wild-type element F, 5′-ccggaattcAATGTGTCAGGTGATTctagagc-3′; mutant element F, 5′-ccggaattcAATGTGTCAGG**CACG**Tctagagc-3′; element μE3, 5′-GATCTGGTCATGTGGCAAGGCTATTTGGG-3′. Lower case, underlined, and bold letters represent adapter sequences including restriction enzyme sites, E-box, and mutated sequences, respectively. DNA probes were 5′ end labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP (7000 Ci/mmol, ICN Radiochemicals).

Binding reactions were carried out in a 10 μL mixture that contains 0.2−3 μL of in vitro translation reaction mixture, (0.5−5) × 10^5 cpm of DNA probe, 2 μg of sheared DNA, 0.5 μg of poly(dI·dC), 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, and 4% glycerol. For competition experiments, the indicated unlabeled DNAs were preincubated for 15 min at room temperature before addition of the probe. For antibody supershift experiments, indicated antibodies were preincubated for 15−20 min at room temperature prior to addition of the probe. Following incubation of the reaction mixture for 20 min at room temperature after addition of the probe, the samples were subjected to electrophoresis in a 6% polyacrylamide gel in 0.5× Tris−borate−EDTA buffer. The dried gels were autoradiographed, and the relative ^{32}P radioactivities of the DNA−protein complexes were quantitated by a phosphorimager (Molecular Dynamics). Antibodies against USF1 and USF2 were obtained from Santa Cruz Biotechnology, Inc. and a mouse anti-His antibody was from Qiagen.

Cell Culture, Transfection of DNA, and Detection of Reporter Gene Products. HeLa cells were obtained from the American Type Culture Collection and maintained in MEM medium supplemented with nonessential amino acids, 10% fetal bovine serum, and antibiotics in a 5% CO_2/air environment at 37 °C. Transfection of DNAs was performed by the calcium−phosphate−DNA coprecipitation method. Typically, 0.5 × 10^5 cells in a six-well plate were transfected with indicated amounts of the pCS3+MT expression constructs or empty vector, 2 μg of luciferase reporter gene construct, and 0.2 μg of pCMV β-Gal. For the experiments in Figure 5A, total amounts of the pCS3+MT test construct and empty vector were 1 μg. For the experiments in Figure 8, 1 μg of the GC construct, 2 μg of pG5CAT reporter construct, and 0.2 μg of pCMV β-GAL were transfected. Transfected cells were harvested in a reporter lysis buffer (Promega), and the activities of luciferase and β-galactosidase were measured using the substrate mixtures from Promega. CAT was detected by a CAT ELISA kit (Boehringer Mannheim).

Sodium Dodecyl Sulfate−Polyacrylamide Gel Electrophoresis (SDS−PAGE) and Protein Blot Analysis. To evaluate in vitro translated proteins, reaction mixtures were subjected to SDS−PAGE using 10% or 12% polyacrylamide gels. The ^{35}S-labeled proteins were detected by autoradiography and quantitated by a phosphorimager (Molecular Dynamics). The unlabeled proteins were blotted and detected by an S-tag HRP chemiluminescent Western blot kit (Novagen). To detect exogenously expressed myc-tagged proteins from the pCS3+MT constructs in the transfected cells, the total protein extracts were prepared by addition of SDS−sample buffer directly to the transfected cells. Protein samples were electrophoresed in SDS−polyacrylamide (10% or 12%) gels and transferred to Immobilon-P membranes (Millipore Corp.). A monoclonal antibody specific to the myc-epitope (Invitrogen) was used for detection with the SuperSignal System (Pierce Chemical Co.).

RNA Blot Analysis. PolyA$^+$ RNA blots of human tissues were obtained from Clontech Laboratories, Inc. The full-length coding region of human TFEC-l, TFE3, USF1, and USF2 cDNAs and 1.8 kb of the 3′ coding region of human NMHC-A cDNA (*23*) were labeled using random priming and [α-^{32}P]dCTP (3000 Ci/mmol, NEN). The blots were hybridized with ^{32}P-labeled probes in a hybridization solution, Hybrisol I (Intergen), at 42 °C and washed in a solution containing 0.1−0.2× SSC (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0) plus 0.5% SDS at 60−68 °C.

RESULTS

Identification of TFEC, TFE3, USF1, and USF2 as Element F-Binding Proteins. We previously demonstrated that the 10 bp element F (GTGTCAGGTG) regulates NMHC-A gene transcription. Furthermore, EMSA with antibodies and native nuclear extracts revealed that USF2 is one of the factors that can bind to this element (*17*). However, the major binding factor(s) remain(s) unidentified. In an effort to identify transacting proteins that bind to element F, we undertook yeast one-hybrid screening.

Three tandem copies of element F were inserted upstream of the minimal promoter of *HIS3*. This reporter gene was integrated in the genome of a yeast strain, YM4271, generating a reporter host strain for cDNA library screening. Yeast expression libraries consisting of human leukocyte and HeLa cell cDNAs fused to the GAL4 activation domain (AD) were then screened to isolate cDNA clones which bind to element F and activate *HIS3* transcription from the reporter gene. A leukocyte cDNA library was chosen, since NMHC-A is highly expressed in these cells. From an initial screening of approximately 1 × 10^6 colony forming units of library plasmid DNA, 62 transformed yeast colonies were isolated. Plasmids recovered from these initial transformants were then tested by retransformation of the reporter strain. Only four were found to reproducibly induce *HIS3* expression in this secondary screen. Sequence analysis of the four positive cDNA clones followed by a search in the GeneBank using the BLAST program revealed that two clones include the partial coding sequence of TFEC and one each includes the coding sequence for TFE3 and USF2. The cDNA region encoded by each clone, relative to the full-length coding region, is illustrated in Figure 1. In the case of clones L25 (TFEC) and L14 (USF2), an open reading frame that is in-frame with GAL4AD is present within the insert sequence. The reading frame of the insert of clone L16 (TFE3) shifts to that of GAL4AD; however, this cDNA clone includes its initiating methionine codon and its own transcriptional activation domain characterized previously (*24*, *25*). Clone L8-1 includes a part of the coding region and 3′ untranslated region of TFEC in a reverse orientation. However, the TFEC transcript could be generated from a cryptic promoter within the ADH1 terminator and could express the bHLH-LZ DNA binding domain along with the C-terminal activation domain (AD2) that is identified in the study below. TFEC, TFE3, and USF2 belong to the bHLH-LZ family of transcriptional factors. All of the isolated clones include at least the bHLH region which has been shown to be responsible for binding to target DNAs containing an E-box sequence, CANNTG (*18*, *19*). The F-element includes the E-box core sequence, indicating consistency of the isolation of the above three factors as F-element binding proteins. In addition, the isolation of USF2 by this screening is also consistent with the previous EMSA results obtained by using native nuclear extracts (*17*).

To examine whether the full-length proteins can bind to element F and to examine the specificity of their DNA

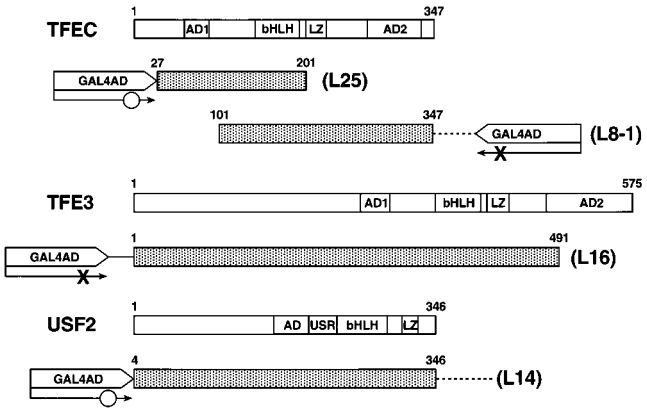

FIGURE 1: Isolation of cDNA clones which bind to element F using the yeast one-hybrid system. Opened bars represent the full-length proteins with functional domains indicated. Isolated cDNA clones (stippled bars) encode the corresponding region of each of the proteins shown above. The solid line in clone L16 and the broken lines in clones L8-1 and L14 represent the 5′ and 3′ untranslated regions, respectively, but the 3′ untranslated regions have not been sequenced. Arrows with ○ indicate that the reading frame of the GAL4 activation domain (AD) and the following cDNAs are the same. Arrows with X indicate that the reading frame shifts between GAL4AD and the following cDNAs. The AD1 and AD2 activation domains in TFEC are identified in this study. The locations of the ADs for TFE3 and USF2 are from refs *25* and *34*, respectively. The numbers above the bars indicate amino acids. USR: USF-specific region.

binding in vitro, the cDNA fragments that encode the full-length coding sequences were amplified by PCR and cloned downstream of the T7 promoter for in vitro transcription and translation. Amplification of TFEC cDNAs from a human leukocyte cDNA library yielded two products with different sizes. The longer product encoding 1041 nts corresponds to TFEC-l (*26*), whereas the shorter product encoding 954 nts corresponds to the human homologue of rat TFEC (*27*), designated TFEC-s in this study. The nucleotide sequences of the two products are identical except that TFEC-s is missing 87 nts between 181 and 267 of TFEC-l. Therefore, TFEC-l and -s share identical sequences in the bHLH-LZ region. We also cloned USF1, since it is known that USF1 and USF2 share the same target DNA binding sites and often form a heterodimer (*28*, *29*). The in vitro transcribed and translated proteins contained a His-tag at their C-terminus as well as an S-tag at their N-terminus. EMSAs were carried out using ^{32}P-labeled element F as a probe. Each of the five different proteins forms a major single complex as shown in Figure 2 (lanes 2, 5, 15, 19, 23). The multiple diffuse complexes formed with TFE3 may be due to multimers of TFE3 in addition to dimers (lane 23) (*30*). The specificity of the DNA sequence for complex formation was confirmed by competition with unlabeled wild-type element F and E-box mutated DNA. Each of the complexes is efficiently competed with the wild-type DNA (lanes 3, 6, 16, 20, 24) but not the mutant DNA (lanes 4, 7, 17, 21, 25). In addition, each complex is recognized by antibodies specific to the expressed proteins, resulting in supershifts of the complexes (lanes 9, 13, 18, 22, 26). These results demonstrate that each of the five different full-length proteins can form a DNA−protein complex with element F in an E-box sequence-specific manner.

Differential DNA Binding Activities of Element F-Binding Proteins. The high degree of conservation of the bHLH-LZ region of USF1 and USF2 (*28*, *31*) as well as TFEC and TFE3 (*24*, *26*, *27*) prompted the question as to whether all these proteins recognize the same DNA element with a similar affinity. We, therefore, compared the binding activities of the five different in vitro translated proteins to element F. First, the protein concentration of each translated protein was estimated in parallel using in vitro translation reactions that included [^{35}S]methionine. Figure 3A shows an autoradiogram of the ^{35}S-labeled proteins following SDS−PAGE. The major band in each lane is the full-length species of each protein with the expected molecular weight. The relative amount of ^{35}S was quantitated by a phosphorimager, and the protein amounts were then normalized according to the methionine content of each protein. Unlabeled proteins were also verified by protein blot using S-protein (data not shown), which confirmed that the quality and relative quantity of each reaction were similar to those of the labeled proteins. EMSAs were performed using the same molar amounts of the unlabeled proteins with a number of different concentrations. Two representative concentrations for each protein are shown in Figure 3B. The complexes formed with the ^{32}P-labeled element F probes were quantitated using a phosphorimager. Relative binding activities are expressed as relative band

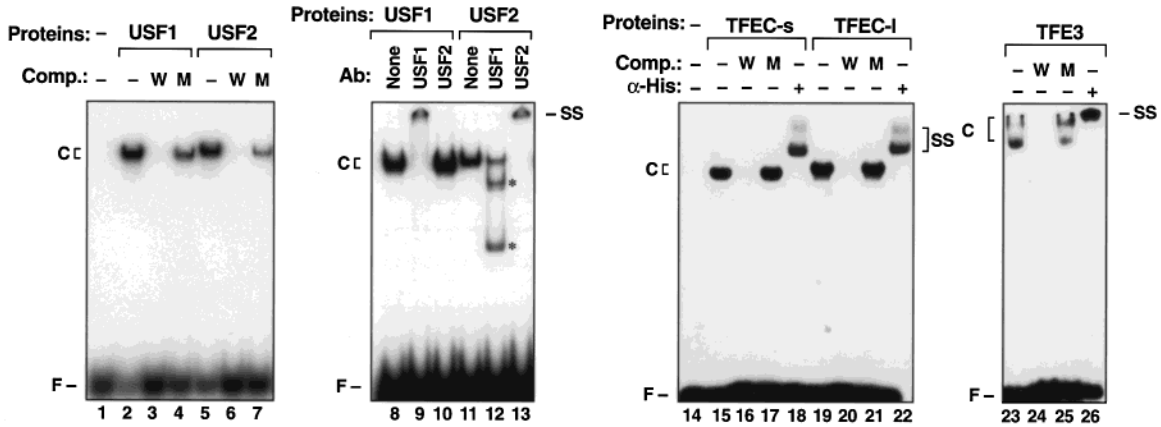

FIGURE 2: Specific DNA−protein complex formation of element F-binding proteins. EMSAs were carried out using ^{32}P-labeled element F as probe and the in vitro translated and His-tagged proteins indicated. Competition experiments (Comp) were performed with a 200-fold molar excess of the unlabeled wild-type (W) and mutant (M) element F. Free probe (F), probe−protein complexes (C), and supershifts (SS) of the complexes by antibodies are indicated. The bands marked by * may be due to degradation of USF2 protein by a contaminated protease in the USF1 antibody preparation. − indicates no addition.

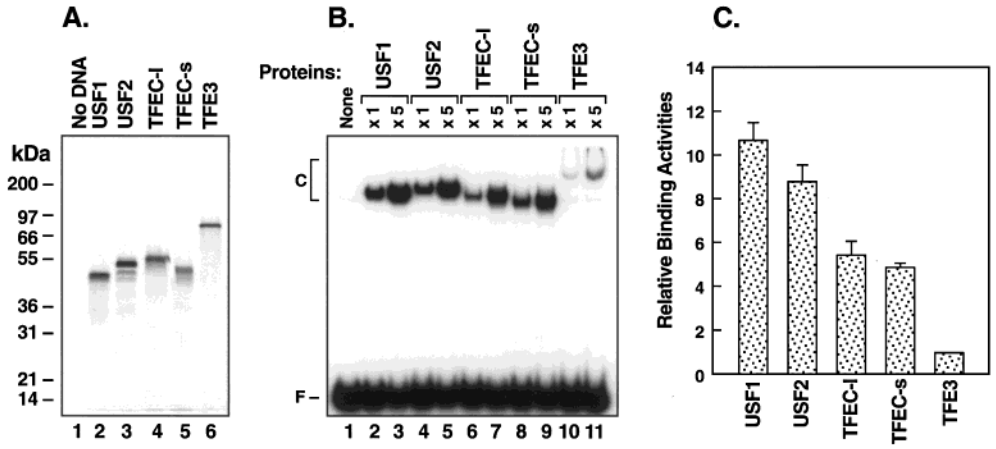

FIGURE 3: Relative DNA binding activities of element F-binding proteins. (A) In vitro translated proteins. The indicated proteins were synthesized by in vitro transcription and translation in a reaction mixture containing [^{35}S]methionine. The autoradiogram following SDS−PAGE is shown. (B) EMSA. EMSAs were carried out using the same molar amounts of each protein indicated with two different concentrations. Element F was used as a probe. (C) Relative binding activities. The complexes formed with the ^{32}P-labeled element F probe were quantitated using a phosphorimager. Relative binding activities of indicated proteins are expressed as relative radioactivities of complexes formed by the indicated proteins to that of the complex formed by TFE3.

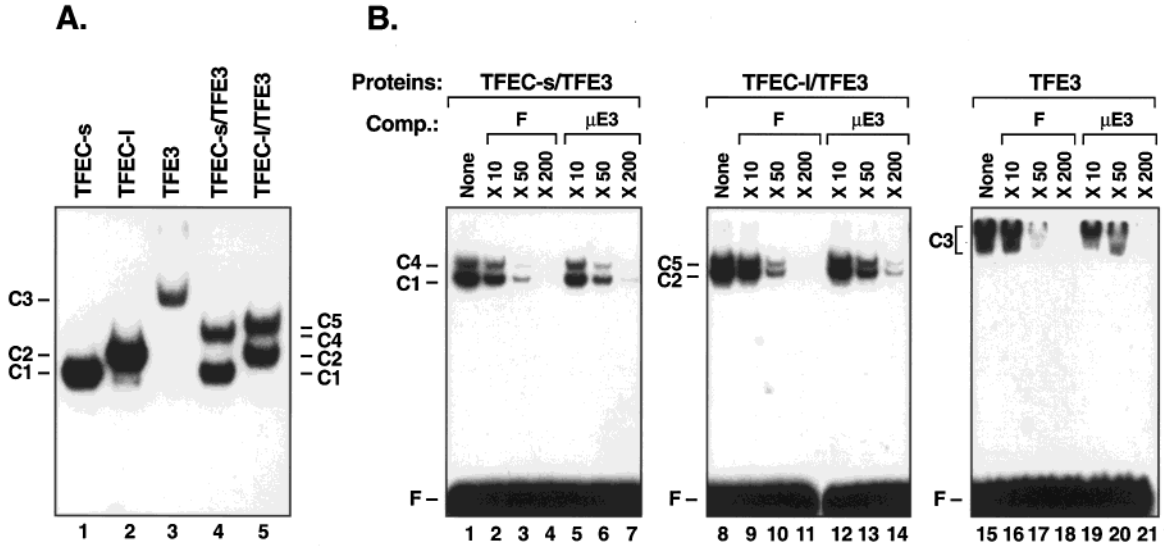

FIGURE 4: Binding activities of homo- and heterodimers of TFEC and TFE3 to target DNAs. (A) Complex formation of TFEC-TFE3 heterodimer and element F. The in vitro translated proteins indicated were used for EMSA using element F as a probe. Cotranslated protein mixtures were used for lanes 4 and 5. C1, homodimeric complex of TFEC-s (lanes 1 and 4); C2, homodimeric complex of TFEC-l (lanes 2 and 5); C3, homodimeric complex of TFE3 (lane 3); C4, heterodimeric complex of TFEC-s and TFE3 (lane 4); C5, heterodimeric complex of TFEC-l and TFE3 (lane 5). (B) Competition for element F−protein complex formation with element μE3. Concentrations of unlabeled competitor DNAs are shown as molar excess ($\times$10, 50, 200) relative to the labeled probe (element F) concentration.

intensity compared to TFE3 (Figure 3C). USF1 and USF2 bind to element F around 10 times stronger than TFE3, and TFEC-l and TFEC-s bind around 5 times stronger.

The bHLH-LZ proteins including USF1, USF2, and TFE3 are known to bind to a target DNA as a dimer and/or tetramer (*30*, *32*). The dimerization properties are mainly determined by the HLH-LZ sequences. USF1 and USF2 have been described as forming heterodimers with a number of target DNA sequences (*28*, *29*). The bHLH-LZ sequences of TFEC and TFE3 are 81% identical, and both are classified in the same subfamily of bHLH-LZ proteins. We, therefore, tested the possibility that heterodimers of TFEC and TFE3 would bind to element F. As shown in Figure 4A (lanes 1−3), each of the translated proteins forms a single complex with different electrophoretic mobility. The mobilities of the TFEC-s− and TFEC-l−DNA complexes (C1 and C2 in lanes 1 and 2) are faster than that of the TFE3−DNA complex (C3 in lane 3), presumably reflecting the differences in molecular weights of these proteins. However, when TFEC-s and TFE3 or TFEC-l and TFE3 were cotranslated and used for EMSA, the unique complexes C4 and C5 with intermediate mobilities are detected along with complexes C1 and C2 (lanes 4 and 5). These intermediate complexes are consistent with their identities as heterodimers made up of TFEC-s and TFE3 and TFEC-l and TFE3. Of note is that the TFE3 homodimer−DNA complex is no longer detected, although the complex of the TFEC homodimer with DNA can still be detected in the cotranslation products. Since the amounts of TFEC (s or l) and TFE3 polypeptides present in the cotranslation mixture were approximately equal (data not shown), these results indicate that the products of TFEC and TFE3 cotranslation form complexes with DNA in which TFE3 is mainly present as a heterodimer with TFEC. The results also suggest that TFEC and TFE3 preferentially form

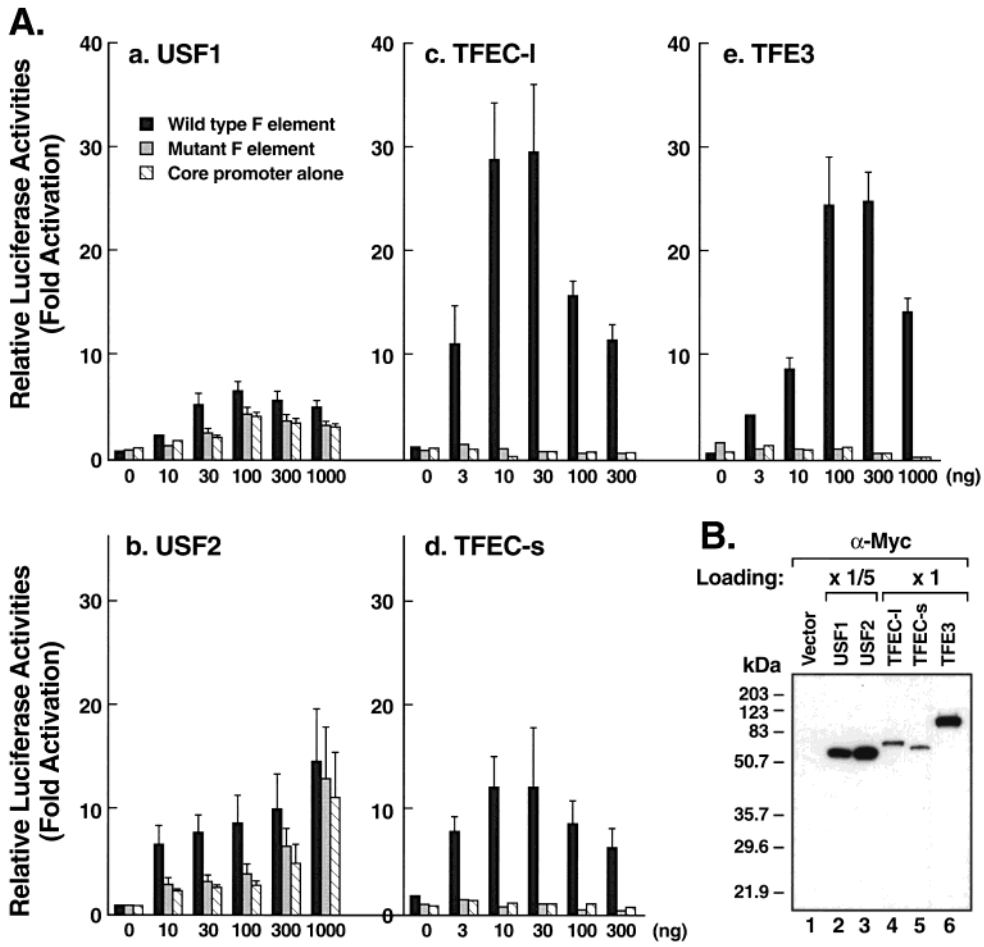

FIGURE 5: Differential transactivation potencies of element F-binding proteins. (A) Reporter gene activation by exogenous expression of element F-binding proteins. Indicated amounts of the expression constructs were cotransfected with various luciferase reporter gene constructs and pCMV β-Gal in HeLa cells. Relative luciferase activities normalized by β-galactosidase activities are shown as fold activation (mean $\pm$ SD, $n = 4$). The luciferase activity due to the core promoter alone cotransfected with the empty vector is represented as 1. (B) Immunoblots of the expressed proteins in transfected cells. Whole cell extracts transfected with 250 ng of the indicated constructs were subjected to immunoblot analysis using an anti-myc antibody following SDS−PAGE. One-fifth of the cell extracts transfected by USF1 and USF2 expression constructs were loaded compared to those transfected by TFEC-l, TFEC-s, and TFE3.

a heterodimer or that the heterodimer binds to element F more efficiently than the TFE3 homodimer in vitro.

To address the question of whether the homo- and heterodimers of TFEC and TFE3 have different specificities for target DNAs, we carried out competition experiments for EMSAs using the F element from the NMHC-A gene as well as the μE3 element from the immunoglobin heavy chain gene. The μE3 element includes a core E-box sequence CATGTG with different flanking sequences from those of element F (*24*; for sequence, see Experimental Procedures). Element μE3 has been shown to be a target for homo- and heterodimers of TFE3 and TFEC-s (*27*). The same increasing amounts of elements F and μE3 were used as competitors in EMSAs with labeled element F as a probe. As shown in Figure 4B, both unlabeled elements F and μE3 can compete for formation of all complexes C1−C5 with similar efficiency, indicating that the binding specificities of homo- and heterodimers of TFEC-s, TFEC-l, and TFE3 for element F are comparable to those of the corresponding complexes for element μE3.

Differential Transactivation Potencies of Element F-Binding Proteins. To assess potential roles of each transcription factor described above in the regulation of NMHC-A gene transcription, we analyzed the transactivation potency of each factor. For this, transient coexpression experiments using the expression constructs for the transcription factors and the reporter gene constructs were carried out in a HeLa cell background. The promoter region of the luciferase reporter gene constructs consists of the core promoter of the NMHC-A gene, which is a TATA-less and initiator-less promoter, and one copy of wild-type or mutant element F with the flanking sequences. The expression of the transcription factors are driven by the CMV promoter, and the expressed proteins contain five or six copies of a myc epitope at the N-terminus for antibody detection. As can be seen in Figure 5, when each of the USF1, USF2, TFEC-l, TFEC-s, and TFE3 constructs was cotransfected with the reporter gene containing wild-type element F, all of these factors activate the expression of luciferase in a dose-dependent manner. However, when USF1 and USF2 are cotransfected with the reporter genes containing mutant element F or the core promoter alone, they still show significant activation of the reporter gene (Figure 5A, a and b) although the core promoter, as well as mutant element F, lacks an E-box sequence. These results suggest that USF1 and USF2 can transactivate the NMHC-A gene in an element F-dependent as well as an independent manner. The element F-dependent activation by USF1 and USF2 is only 2- and 5-fold,

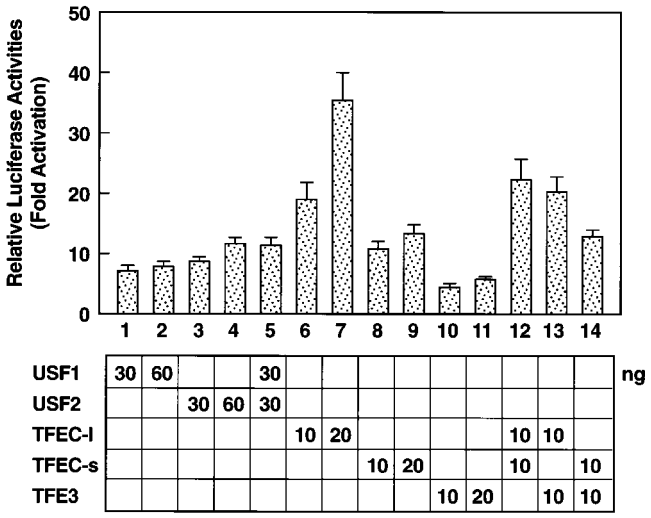

FIGURE 6: Effects of cotransfection of USF1 and USF2 or TFEC and TFE3 on reporter gene transactivation. Indicated amounts of the expression constructs were cotransfected with the luciferase reporter gene containing wild-type element F and pCMV β-Gal. Relative luciferase activities normalized by β-galactosidase activities are shown (mean $\pm$ SD, $n = 3$). The luciferase activity due to the core promoter alone cotransfected with the empty vector was taken as 1 (not shown).

respectively, at most. In contrast, TFEC-l, TFEC-s, and TFE3 activate the reporter gene expression 30-, 10-, and 25-fold, respectively, solely in an element F-dependent manner (Figure 5A, c−e). Moreover, TFEC-l and TFEC-s show the highest transactivation even though the cells were transfected with less plasmid. The decreased activation at the higher concentrations of the transfected plasmids seen with TFEC and TFE3 may be due to a squelching phenomenon. The protein amounts expressed were also verified by immunoblot analysis using an anti-myc antibody (Figure 5B). We see reproducible differences in the expressed protein levels among different proteins. These differences are not due to transfection efficiency nor due to the constructs themselves, since the transfection efficiency has been normalized by the cotransfected β-galactosidase activities and the same constructs can produce the same amounts of proteins in an in vitro transcription and translation system. It is also not likely due to the differential effects of the expressed proteins on the CMV promoter activity, since the cotransfected β-galactosidase activities driven by the same CMV promoter do not vary more than 2-fold. The steady-state levels of the expressed TFEC-l and TFEC-s are lower than those of the other proteins, presumably due to the short lives of these proteins. Indeed, the C-terminal region of TFEC contains two PEST-like sequences which have been demonstrated to serve as signals for regulated and constitutive proteolysis (*33*). Despite the fact that TFEC protein amounts that have accumulated in the cells are the lowest, the plasmid amounts required to obtain maximal activation are also the lowest. Therefore, TFEC-l has the highest potency for transactivation among the factors analyzed.

As shown above, since TFEC and TFE3 can form heterodimers and bind to element F in vitro (see Figure 4), and USF1 and USF2 are known to form a heterodimer, we examined the effects of the cotransfection of these two factors on reporter gene expression. As shown in Figure 6, transfection of both USF1 and USF2 does not show any increase

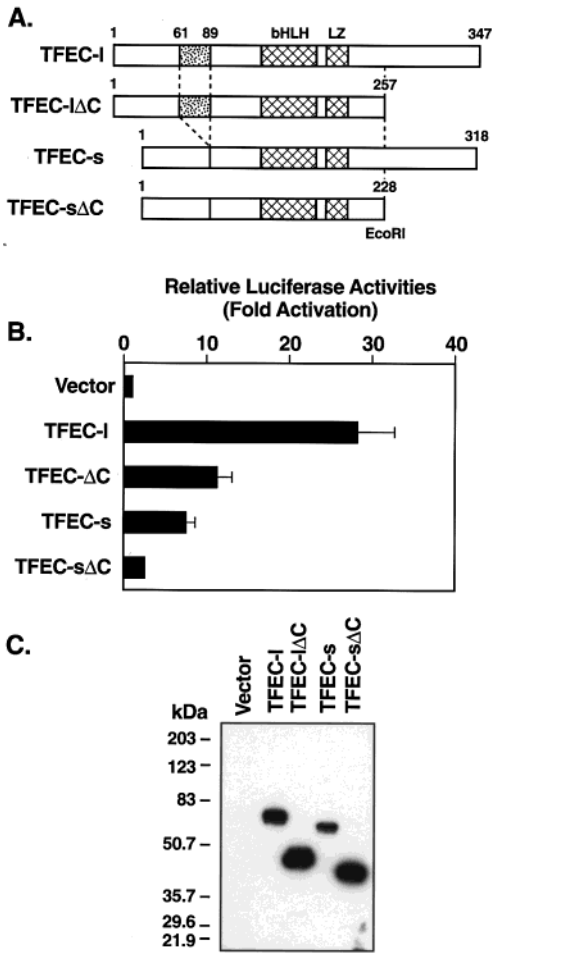

FIGURE 7: Two activation domains are present in TFEC. (A) Schematic representation of the full-length and truncated TFECs. (B) Transactivation of the reporter gene. Thirty nanograms of the indicated constructs was cotransfected with the luciferase reporter gene containing wild-type element F and pCMV β-Gal into HeLa cells. Relative luciferase activities normalized by β-galactosidase activities are shown (mean $\pm$ SD, $n = 4$). The luciferase activity due to the core promoter alone cotransfected with the empty vector is represented as 1. (C) Immunoblots of the expressed proteins. Cell extracts, normalized by the β-galactosidase activities, were loaded in each lane. Expressed proteins were detected by using an anti-myc antibody.

in activation compared to transfection of either USF1 or USF2 alone. Similarly, transfection of both TFEC-l and TFE3 or both TFEC-s and TFE3 shows no increased activation than that due to TFEC-l and TFE3 alone or due to TFEC-s and TFE3 alone. Although the extent of heterodimer formation was not determined, we could not detect functional cooperativity between USF1 and USF2 or between TFEC and TFE3.

Taken together, TFEC-l is a major factor, together with TFE3, for activation of the NMHC-A transcription via element F. Of particular importance is that TFEC-l and TFEC-s both function as transcriptional activators.

Two Activation Domains of TFEC. The different transactivation potencies between TFEC-l and TFEC-s suggest the possibility for the existence of more than one region being responsible for transactivation. The only difference in the primary amino acid sequences between TFEC-l and TFEC-s is a 29 aa region (aa 61−89 in TFEC-l) located at the N-terminal side from the bHLH-LZ region. This region could

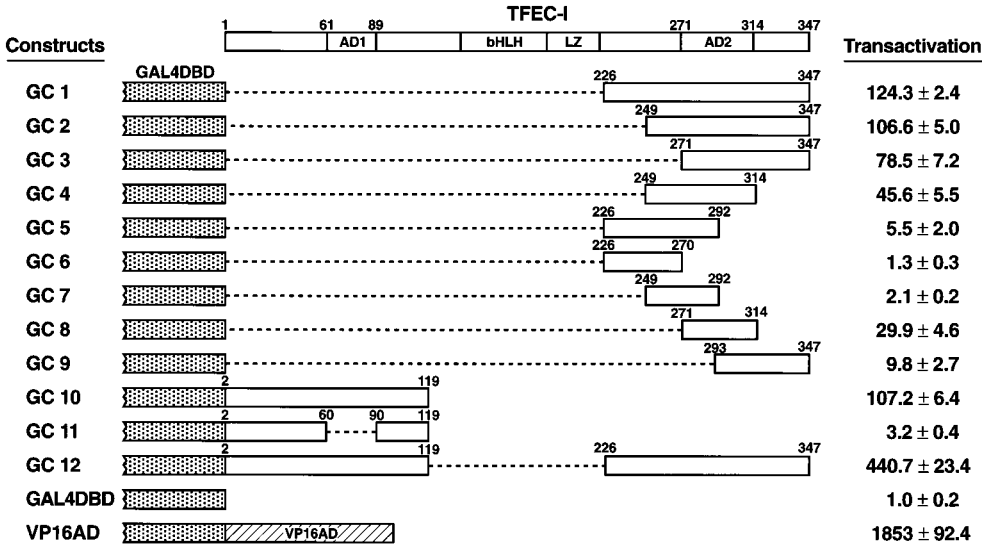

FIGURE 8: Localization of activation domains of TFEC using a heterologous GAL4 DNA binding domain. The top diagram indicates functional domains of TFEC-l. Numbers represent amino acids. Each construct includes the region indicated by the boxes, and deletions are indicated by the broken lines. Transactivation of various GAL4DBD-TFEC hybrid proteins was determined by cotransfection with pG5CAT which contains five copies of GAL4 DNA binding sites and pCMV β-Gal. Relative CAT activities normalized by β-galactosidase activities are shown (mean ± SD, $n = 3$). The CAT activity due to the GAL4DBD alone is represented as 1. Construct VP16AD is a positive control plasmid, pM3-VP16, which expresses a hybrid protein of the GAL4DBD and the VP16AD.

serve as an activation domain, since TFEC-s, which lacks this region, is a weaker transactivator than TFEC-l. On the other hand, TFEC-s also retains significant transactivation potency. This observation prompted us to undertake characterization of the activation domains in this protein.

First, the C-terminal 122 aa region was deleted in both the TFEC-l and TFEC-s contexts. The deletion mutants TFEC-lΔC and TFEC-sΔC as well as the wild type were cotransfected with the luciferase reporter gene that includes element F. The mutant proteins were expressed at higher levels compared to the wild types in both cases (Figure 7C). These truncated mutant proteins are capable of localizing to the nucleus and binding to the target sequence since they contain an intact bHLH-LZ domain, which has been shown to be responsible for nuclear localization and DNA binding of a number of bHLH proteins (*34*−*37*). Despite the higher expression, as shown in Figure 7B, the mutant proteins have reduced transactivation activities compared to their wild-type counterparts. TFEC-sΔC shows only residual activation. These results indicate that a second activation domain exists in the C-terminal half of the molecule and this activation domain functions in both the l and s forms of TFEC.

To further characterize the activation domains of TFEC, we made use of the heterologous GAL4 DNA binding domain (DBD) to create hybrid proteins. Various segments of TFEC were fused to the GAL4DBD by using the 1−147 aa region of GAL4, which is known to be sufficient for dimerization, DNA binding, and nuclear localization but does not, by itself, activate transcription. Expression constructs for the GAL4DBD-TFEC hybrids were cotransfected into HeLa cells with a CAT reporter gene driven by a minimal E1b promoter with five upstream copies of the GAL4-binding elements. Fusion of GAL4DBD with the C-terminal 122 aa region of TFEC results in a hybrid protein that stimulates transcription from the reporter gene 124-fold (GC1) over that obtained with the GAL4DBD alone (Figure 8). To define the critical region for this transactivation, various segments of the C-terminal region were further analyzed as GAL4DBD-hybrid proteins, and the results are summarized in Figure 8. Although deletion of any segment causes a decrease in activation of the reporter gene to some extent, the most critical sequences responsible for transactivation lie between aa 271−314 and may extend to aa 347. Indeed, only the hybrids containing this region show significant levels of activation (over 20-fold, GC1-4,8), and all hybrids lacking it are essentially inactive (GC5-7,9). Furthermore, the construct which contains only this region (GC8) displays significant activity. The sequences between aa 226−270 are dispensable. We also constructed a fusion of GAL4DBD with the entire N-terminal region of TFEC-l (aa 1−119, GC10). This hybrid protein activates the reporter gene transcription 107-fold. In contrast, the entire N-terminal region of TFEC-s (GC11) does not show significant activation of the reporter. These results strengthen the previous conclusion using the native protein context that aa 61−89 of TFEC-l are critical for transactivation, in addition to the C-terminal activation domain described above.

We next addressed the question of whether the two activation domains have any functional interaction using the same GAL4DBD-hybrid system. The middle bHLH-LZ domain was excluded from this hybrid to avoid possible interference with the DNA binding and dimerization properties of GAL4DBD. The hybrid containing both the N-terminal and C-terminal regions activates the reporter gene 440-fold (GC12). This activation potency is much higher than the simple sum of the activation potency of the two activation domains. Thus, the two activation domains appear to function synergistically with one another. However, we could not rule out the possibility that a difference in protein expression of the constructs could confound the interpretation of our results since we were unable to detect the expressed hybrid proteins to determine the levels of expression by immunoblots, using a number of different available antibodies against the GAL4DBD.

On the basis of the results obtained in the context of both native protein and GAL4DBD-hybrid protein, we conclude

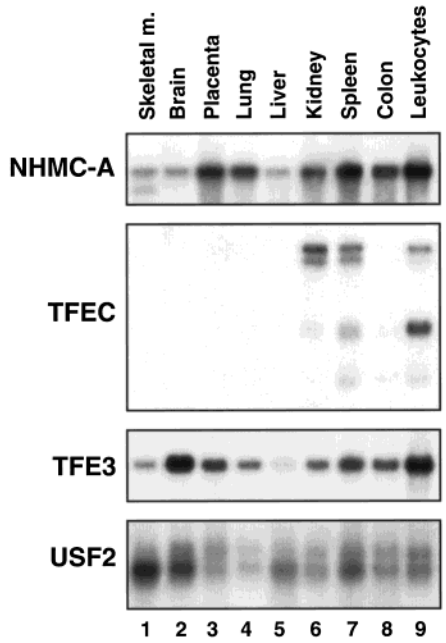

FIGURE 9: mRNA expression of element F-binding proteins and NMHC-A in human tissues. RNA blots, which include polyA$^+$ RNA isolated from the tissues indicated at the top, were hybridized by the probe indicated at the left. Sizes of the mRNAs are as follows: NMHC-A, 7.5 kb; TFEC, 8.3, 7.2, 3.5, and 2.0 kb; TFE3, 3.7 kb; and USF2, 2.4 and 2.0 kb.

that TFEC has two activation domains, one (AD1) at the N-terminus and the other (AD2) at the C-terminus. These two domains are separated in the primary aa sequence by a large bHLH-LZ domain. Whereas TFEC-s lacks the N-terminal activation domain (AD1), the C-terminal activation domain (AD2) functions in both the l and s forms.

mRNA Expression of Element F-Binding Proteins. Finally, in an initial effort to investigate the biological relevance of element F-binding proteins for NMHC-A gene regulation, we analyzed mRNA expression of these transcriptional factors as well as NMHC-A. Figure 9 shows RNA blots of various human tissues. The TFEC probe detects four transcripts with different sizes (8.3, 7.2, 3.5, and 2.0 kb), even at the highest stringency (0.1× SSC, 68 °C). Although we do not know the differences in sequence among these four transcripts, these are not due to cross-hybridization to the other members of the closely related family, such as TFE3. There is no previous report for human TFEC mRNA using RNA blot analysis, although a previous study using rat chondrosarcoma cells detected two TFEC mRNAs, 3.0 and 1.6 kb in size (*27*), and a study of mouse macrophage cell lines detected a 1.8 kb mRNA for TFEC (*38*). The same TFEC probe used here for human mRNA analysis also hybridizes to a single band of 1.7 kb in mouse tissues (data not shown). Therefore, it seems that the sizes of TFEC transcripts differ among different species. The TFEC transcripts are detected only in a number of limited tissues or cells, such as kidney, spleen, and leukocytes (lanes 6, 7, and 9 in Figure 9; *27*, *38*). In contrast, TFE3 and USF2 are expressed ubiquitously, although the relative abundance varies (Figure 9; *27*, *39*, *40*). Of note is that skeletal muscle and brain express USF2 mRNA at a higher level (lanes 1 and 2). In agreement with previous studies (*12*−*14*), NMHC-A mRNA is expressed abundantly in placenta, lung, kidney, spleen, colon, and leukocytes (lanes 3, 4, and 6−9 in Figure 9). Skeletal muscle and brain express much less (lanes 1 and 2 in Figure 9). A comparison of mRNA expression shows that all of the tissues where TFEC is present express NMHC-A at a high level. This is consistent with the observation that TFEC has the highest transcription potency. On the other hand, the high level of expression of NMHC-A mRNA in placenta, lung, and colon is not related to TFEC expression. This figure also shows that, in the absence of TFEC expression, regardless of the presence of TFE3, the tissues where USF2 expression is higher tend to express NMHC-A less abundantly. This is consistent with the observation that USF2 shows a low transcription potency with 10 times higher binding activity than TFE3. Although NMHC-A transcription is regulated via multiple elements and not only by element F, we see a better correlation between NMHC-A expression and TFEC expression compared to that of the other factors.

DISCUSSION

In an effort to understand the molecular mechanism underlying the regulation of NMHC-A gene expression, we have identified and characterized transcriptional factors which are capable of binding to element F of the NMHC-A gene. Element F includes a core E-box hexanucleotide sequence, CANNTG. The E-box sequence is known to be recognized by a large family of transcriptional factors which contain the bHLH motif (*18*, *19*). Among this large family of proteins, TFEC, TFE3, USF1, and USF2 were found to bind to element F by using the yeast one-hybrid system and/or EMSA. All four of these proteins also contain an LZ structure following the bHLH region. TFEC and TFE3 belong to the same subfamily whereas USF1 and USF2 belong to another subfamily. We also found two isoforms, l and s, for TFEC, which are generated, most likely, by alternative pre-mRNA splicing. Using quantative reverse transcription−PCR, we estimate that approximately 70−80% of total TFEC is the l form and 20−30% is the s form in human promyelocytic leukemia HL60 cells and approximately 90% is the l form in peripheral leukocytes.[2] In vitro binding studies using EMSAs and cotransfection studies of the transcription factor expression constructs with the promoter-reporter constructs showed that these factors have different binding activities to element F with different transactivation potencies. TFEC-l shows high binding activity with the highest transactivation potency whereas TFEC-s has the same binding activity as TFEC-l with an intermediate transactivation. TFE3 has a high transactivation potency but the lowest binding activity. In contrast, USF1 and USF2 bind to element F with the strongest binding activities, yet element F-dependent transactivation of the reporter gene by these factors is the lowest. This low transactivation property of USF may be due to the fact that transactivation by USF1 and USF2 is highly dependent on the context of the promoter sequences as well as on the nature of the particular cell background (*34*, *41*, *42*). It has been reported that USF activates transcription from promoters which contain an initiator and a TATA box to a high extent, but only to a small extent when the promoter lacks an initiator. The NMHC-A core promoter used in this study lacks an initiator as well as a TATA box. Therefore,

[2] S. Sukenick and S. Kawamoto, unpublished observation.

although element F itself can recruit USF1 and/or USF2, transcription of the NMHC-A-luciferase gene is not efficiently activated by these factors.

Knowing the properties of the element F-binding proteins described above, how might these proteins be implicated in the regulation of NMHC-A transcription in vivo? A comparison of mRNA levels of each transcriptional factor with those of NMHC-A among different tissues suggests the following: in tissues such as muscle or brain where TFEC is not expressed, USF1 and USF2, which have high binding activities but low transactivation potencies, might occupy element F, preventing it from binding to TFE3, which has low binding activity but a high transactivation potency resulting in low expression of NMHC-A. Thus, USF functions as a negative regulator. On the other hand, in tissues such as spleen and kidney, TFEC, which has comparable binding activity to USF and a high transactivation potency, would occupy element F, resulting in high transcription of the NMHC-A gene. It is noteworthy that a TFEC-TFE3 heterodimer has a higher binding activity than a homodimer of TFE3 and this heterodimer has a high transactivation potency. Since TFE3 is expressed in most of the tissues and cells, expression of TFEC would become the determinant for heterodimer formation. The data presented here are consistent with the idea that TFEC can function as a transcriptional activator for the NMHC-A gene. Elucidation of the physiological roles of the element F-binding proteins in the regulation of NMHC-A transcription is an important subject for future studies.

In addition to the finding that NMHC-A is a target gene for TFEC, another important finding in this study is that TFEC is, indeed, a transcriptional activator. TFEC-s was originally isolated by homology of the bHLH region to TFE3 using rat chondrosarcoma mRNA (*27*). TFEC-l was subsequently isolated using the bHLH-LZ region cDNA fragment of MITF from a human THP-1 monocytic leukemia cell cDNA library (*26*). MITF belongs to the same subfamily of bHLH-LZ proteins as TFEC and TFE3. In both cases, however, the transactivation potency of TFEC has not been previously demonstrated. TFEC-s has been described as being unable to transactivate a reporter gene linked to a promoter containing tandem copies of the immunoglobin μE3 enhancer, although TFEC-s can bind to element μE3 (*27*). The effects of TFEC-l on transcription were previously examined using a melanocyte-specific tyrosinase gene and a ubiquitously expressed heme oxygenase-1 gene (*26*). Each promoter contains a cis-acting E-box sequence. Although slightly higher reporter gene activation compared to the basal promoter was detected, this activation corresponds to only 1−10% of the activation due to TFE3, indicating essentially no activation of the tyrosinase and heme oxygenase-1 genes by TFEC-l. Moreover, there is no evidence that TFEC is capable of binding to the E-box sequences of these two genes. In contrast, TFEC-l can activate the NMHC-A promoter-reporter gene 30-fold in an element F-dependent manner, and this activation is higher than that due to TFE3. TFEC-s also causes 10-fold activation of the reporter gene. In this study, we have demonstrated that TFEC has two separate activation domains. One is located at the N-terminal side of the bHLH-LZ region, and the other is at the C-terminal side. The presence of the C-terminal activation domain in TFEC is consistent with the original isolation of yeast clone L8-1, the protein product of which is predicted to include no GAL4AD nor the N-terminal activation domain, yet it activates the reporter *HIS3* gene in yeast screening. The most critical region of the N-terminal activation domain (AD1) was mapped between aa 61−89 of TFEC-l. These 29 aa consist of eight acidic amino acids and 10 bulky hydrophobic amino acids. Thus, this domain is classified as an acidic activation domain (*43*, *44*). This acidic activation domain is conserved among all four members of the subfamily, TFEC, TFE3, TFEB, and MITF (*45*−*47*). Specifically, the region of aa 61−74 is highly conserved in the subfamily and is predicted to form an amphipathic α-helical conformation (*24*, *48*). This region of MITF has been shown to interact with the coactivator, CBP/p300 (*48*).

The most critical region of the C-terminal activation domain (AD2) of TFEC is located between aa 271−314 of TFEC-l. Amino acid composition of this 44 aa region shows 14 bulky hydrophobic aa, 9 acidic aa, 8 serine or threonine, 3 prolines, and 2 glutamines. On the basis of this aa composition, the C-terminal activation domain does not fit into the usual classification for activation domains (*43*, *44*). Among four members of the subfamily, the existence of the C-terminal activation domain has been described only for TFE3, which is rich in proline (*25*). However, the C-terminal regions of TFEC and TFE3 are quite different in aa sequences and composition and show sequence similarity only at the very C-terminal 29 aa. Deletion of this region causes only a 2.5-fold reduction of the activation using the GAL4DBD-hybrid system. A large portion of the activation still remains in the region lacking homology to TFE3. How the two activation domains of TFEC can interact with other factors, such as coactivators, and how the two domains functionally interact will be the subject of future studies. As far as we are aware, this study is the first characterization of TFEC as a transcriptional activator and of NMHC-A as its target gene.

ACKNOWLEDGMENT

We thank Dr. Robert S. Adelstein (NHLBI) for continued encouragement and critical reading of the manuscript. We also thank Scott Sukenick (Duke University) for useful contributions and Dr. Yongsok Kim (NHLBI) for helpful discussions. We acknowledge the DNA synthesis and DNA sequencing of George Poy (NIDDK) and the excellent editorial assistance of Catherine S. Magruder.

REFERENCES

1. Mooseker, M. S., and Cheney, R. E. (1995) *Annu. Rev. Cell Dev. Biol. 11*, 633−675.
2. Sellers, J. R., and Goodson, H. V. (1995) *Protein Profile 2*, 1323−1423.
3. Spudich, J. A. (1989) *Cell Regul. 1*, 1−11.
4. Weintraub, H. (1993) *Cell 75*, 1241−1244.
5. Olson, E. N. (1993) *Circ. Res. 72*, 1−6.
6. Yu, Y.-T., Breitbart, R. E., Smoot, L. B., Lee, Y., Mahdavi, V., and Nadal-Ginard, B. (1992) *Genes Dev. 6*, 1783−1798.
7. Robbins, J., Palermo, J., and Rindt, H. (1995) *Ann. N.Y. Acad. Sci. 752*, 492−505.
8. Katsuragawa, Y., Yanagisawa, M., Inoue, A., and Masaki, T. (1989) *Eur. J. Biochem. 184*, 611−616.
9. Shohet, R. V., Conti, M. A., Kawamoto, S., Preston, Y. A., Brill, D. A., and Adelstein, R. S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7726−7730.

10. Takahashi, M., Kawamoto, S., and Adelstein, R. S. (1992) *J. Biol. Chem. 267*, 17864−17871.
11. Kawamoto, S. (1996) *J. Biol. Chem. 271*, 17613−17616.
12. Kawamoto, S., and Adelstein, R. S. (1991) *J. Cell Biol. 112*, 915−924.
13. Simons, M., Wang, M., McBride, O. W., Kawamoto, S., Yamakawa, K., Gdula, D., Adelstein, R. S., and Weir, L. (1991) *Circ. Res. 69*, 530−539.
14. Phillips, C. L., Yamakawa, K., and Adelstein, R. S. (1995) *J. Muscle Res. Cell Motil. 16*, 379−389.
15. Toothaker, L. E., Gonzalez, D. A., Tung, N., Lemons, R. S., Le Beau, M. M., Arnaout, M. A., Clayton, L. K., and Tenen, D. G. (1991) *Blood 78*, 1826−1833.
16. Kawamoto, S. (1994) *J. Biol. Chem. 269*, 15101−15110.
17. Beohar, N., and Kawamoto, S. (1998) *J. Biol. Chem. 273*, 9168−9178.
18. Atchley, W. R., and Fitch, W. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5172−5176.
19. Massari, M. E., and Murre, C. (2000) *Mol. Cell. Biol. 20*, 429−440.
20. Puri, P. L., and Sartorelli, V. (2000) *J. Cell Physiol. 185*, 155−173.
21. Grandori, C., Cowley, S. M., James L. P., and Eisenman, R. N. (2000) *Annu. Rev. Cell Dev. Biol. 16*, 653−699.
22. Roth, M. B., Zahler, A. M., and Stolk, J. A. (1991) *J. Cell Biol. 115*, 587−596.
23. Saez, C. G., Myers, J. C., Shows, T. B., and Leinwand, L. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1164−1168.
24. Beckmann, H., Su, L.-K., and Kadesch, T. (1990) *Genes Dev. 4*, 167−179.
25. Artandi, S. E., Merrell, K., Avitahl, N., Wong, K.-K., and Calame, K. (1995) *Nucleic Acids Res. 23*, 3865−3871.
26. Yasumoto, K.-i., and Shibahara, S. (1997) *Biochim. Biophys. Acta 1353*, 23−31.
27. Zhao, G.-Q., Zhao, Q., Zhou, X., Mattei, M.-G., and de Crombrugghe, B. (1993) *Mol. Cell. Biol. 13*, 4505−4512.
28. Sirito, M., Walker, S., Lin, Q., Kozlowski, M. T., Klein, W. H., and Sawadogo, M. (1992) *Gene Exp. 2*, 231−240.
29. Viollet, B., Lefrancois-Martinez, A.-M., Henrion, A., Kahn, A., Raymondjean, M., and Martinez, A. (1996) *J. Biol. Chem. 271*, 1405−1415.
30. Artandi, S. E., Cooper, C., Shrivastava, A., and Calame, K. (1994) *Mol. Cell. Biol. 14*, 7704−7716.
31. Gregor, P. D., Sawadogo, M., and Roeder, R. G. (1990) *Genes Dev. 4*, 1730−1740.
32. Ferre-D'Amare, A. R., Pognonec, P., Roeder, R. G., and Burley, S. K. (1994) *EMBO J. 13*, 180−189.
33. Rechsteiner, M., and Rogers, S. W. (1996) *Trends Biochem. Sci. 21*, 261−271.
34. Luo, X., and Sawadogo, M. (1996) *Mol. Cell. Biol. 16*, 1367−1375.
35. Goldfarb, A. N., and Lewandowska, K. (1994) *Exp. Cell Res. 214*, 481−485.
36. Deed, R. W., Armitage, S., and Norton, J. D. (1996) *J. Biol. Chem. 271*, 23603−23606.
37. Nagoshi, E., Imamoto, N., Sato, R., and Yoneda, Y. (1999) *Mol. Biol. Cell 10*, 2221−2233.
38. Rehli, M., Lichanska, A., Cassady, A. I., Ostrowski, M. C., and Hume, D. A. (1999) *J. Immunol. 162*, 1559−1565.
39. Sirito, M., Lin, Q., Maity, T., and Sawadogo, M. (1994) *Nucleic Acids Res. 22*, 427−433.
40. Roman, C., Cohn, L., and Calame, K. (1991) *Science 254*, 94−97.
41. Du, H., Roy, A. L., and Roeder, R. G. (1993) *EMBO J. 12*, 501−511.
42. Qyang, Y., Luo, X., Lu, T., Ismail, P. M., Krylov, D., Vinson, C., and Sawadogo, M. (1999) *Mol. Cell. Biol. 19*, 1508−1517.
43. Johnson, P. F., Sterneck, E., and Williams, S. C. (1993) *J. Nutr. Biochem. 4*, 386-398.
44. Triezenberg, S. J. (1995) *Curr. Opin. Genet. Dev. 5*, 190−196.
45. Fisher, D. E., Carr, C. S., Parent, L. A., and Sharp, P. A. (1991) *Genes Dev. 5*, 2342−2352.
46. Hemesath, T. J., Steingrimsson, E., McGill, G., Hansen, M. J., Vaught, J., Hodgkinson, C. A., Arnheiter, H., Copeland, N. G., Jenkins, N. A., and Fisher, D. E. (1994) *Genes Dev. 8*, 2770−2780.
47. Rehli, M., Den Elzen, N., Cassady, A. I., Ostrowski, M. C., and Hume, D. A. (1999) *Genomics 56*, 111−120.
48. Sato, S., Roberts, K., Gambino, G., Cook, A., Kouzarides, T., and Goding, C. R. (1997) *Oncogene 14*, 3083−3092.

BI002847D