Divergent hTAFII31-binding Motifs Hidden in Activation Domains

Yongmunn Choi, Shinichi Asada, and Motonari Uesugi

Activation domains are functional modules that enable DNA-binding proteins to stimulate transcription. Characterization of these essential modules in transcription factors has been hampered by their low sequence homology. Here we delineate the peptide sequences that are required for transactivation and interaction with hTAFII31, a classical target of the acidic class of activation domains. Our analyses indicate that hTAFII31 recognizes a diverse set of sequences in transcription for approximately 67% of the actively expressed transcription factors: NFAT1, ALL1, NF-IL6, ESX, and HSF-1. The interaction surfaces are localized in short peptide segments of activation domains. The brevity and heterogeneity of the motifs may explain the low sequence homology among acidic activation domains.

It has been reported that hTAFII31 makes direct contacts with the activation domains of VP16, p53, and NF-κB p65 and that the strength of the interactions correlates with the ability to activate transcription (5–9, 17, 18). NMR and biochemical studies have shown that the activation domains of VP16 and p53 undergo an induced transition from random coil to α-helix upon interaction with hTAFII31, with key hydrophobic residues along one face of the nascent helix (17, 18). The pattern of such hydrophobic residues, FXXΦΦ (where X represents any residue and Φ represents any hydrophobic residue) is conserved among the activation domains of VP16, p53, and NF-κB p65, suggesting that this sequence represents a recognition element for hTAFII31.

Here we delineate the peptide sequences that are required for transactivation and for interaction with hTAFII31. Our analyses indicate that hTAFII31 recognizes a more divergent set of peptide sequences than FXXΦΦ for the transmission of activation signals. This sequence characterization enabled the identification of hTAFII31-binding sequences hidden in the activation domains of human transcription factors: NFAT1, ALL1, NF-IL6, ESX, and HSF-1. A combination of mutational studies and NMR analyses indicated that the interaction surfaces comprise short peptide regions containing signature α-helical motifs. Furthermore, the strength of the interactions between these activators and hTAFII31 correlates with the ability to activate transcription in human cells, supporting the notion that hTAFII31 and its homologs are important targets of eukaryotic transactivators.

**EXPERIMENTAL PROCEDURES**

*Mutant Library Screening*—We constructed four small libraries of a mammalian expression vector, each encoding a 17-amino acid peptide from the VP16 activation domain (VP16-(469–485)) fused with the GAL4 DNA-binding domain. Each of the four libraries consists of random point mutants at one of the four positions within the FXXΦΦ sequence of VP16-(469–485) (Phe<sup>479</sup>Thr<sup>480</sup>Arg<sup>481</sup>Asp<sup>482</sup>Asp<sup>483</sup>). 192 clones from each library were miniprepped and individually transfected into human Jurkat Tag cells in two 96-well plates (100 ng/well of expression plasmid and ~10<sup>4</sup> cells/well). The reporter construct we used (120 ng/well) was pG5IL2SX in which the secreted alkaline phosphatase (SEAP) gene is controlled by five GAL4-binding sites. After a 48-h incubation, each well was assayed for SEAP activity through fluorescence change of 4-methylumbelliferyl phosphate as described (19). Fluorescence measurements were carried out by a microplate reader, Fluoroskan II (Labsystems). Positive clones were characterized by DNA sequencing, and their activities were quantitatively estimated through repeated transfection experiments in a larger volume.

*In Vitro Protein-Protein Interaction Assay*—The protein hTAFII31-(1–140) was purified as described (18). Glutathione S-transferase fusion proteins of activation domains were expressed in BL21(DE3)pLysS and purified by affinity chromatography using glutathione-agarose beads. The beads binding either glutathione S-transferase (GST) fusion protein or GST only were incubated with hTAFII31-(1–140) in 200 μl of binding buffer containing 25 mM NaCl, 0.005% Nonidet P-40, 10% glycerol, 20 mM Tris–HCl, pH 7.5, 1 mM MgCl<sub>2</sub>, and 2 mM dithiothreitol.
for 1 h at 4 °C. After extensive washing with the same buffer, the bound hTAFII31-(1–140) was analyzed by SDS-polyacrylamide gel electrophoresis.

Transcription Assay—The DNA encoding the GAL4 DNA-binding domain (residues 1–94) was subcloned into the HindIII/HpaI site of pCDNA3 (Invitrogen) and the resultant pcDNA/GAL plasmid was used to construct mammalian expression vectors for GAL4 fusions of ALL1-(2829–2883), NFIL6-(24–124), ESX-(129–159), HSF-1-(371–430) and NFAT1-(1–96). Jurkat Tag cells (2 × 10^6) were transfected with 500 ng of each GAL4 fusion construct along with 2 μg of pG5IL25X. After a 48-h incubation, an aliquot of the culture was removed and assayed for SEAP activity as described (19). The expression levels of the GAL4 fusion proteins were comparable, as judged by Western blot analyses using an antibody against the GAL4 DNA-binding domain.

NMR Studies—Peptides were synthesized using Rink Amide MBHA resin, purified by high performance liquid chromatography, and characterized by NMR, amino acid analyses, and electrospray ionization mass spectroscopy. The peptide was dissolved in 95% H_2O plus 5% D_2O containing 130 mM KCl, 5 mM perdeuterated dithiothreitol, 20 mM perdeuterated Tris-AcOH (pH 6.2), and 10 μM EDTA, and then the pH of the solution was adjusted to ~6.2 by adding dilute KOH. The final concentrations of peptides were determined by UV absorption or amino acid analyses to be ~4 mM. NMR experiments were performed in the absence or presence of hTAFII31-(1–140) (240 μM) on a Bruker AMX600 spectrometer. The sequential assignment of the peptide signals was obtained by using a combination of total correlation spectroscopy, DQF-COSY, and nuclear Overhauser effect spectroscopy (NOESY) data sets of a free peptide sample. Sequential d_ii(i, i + 1) NOEs, although weak, were observed in the NOESY spectra of the free peptide, which served as a basis for the complete sequential assignment. In the NOESY spectra, 512 free induction decays were recorded at 290 or 300 K with mixing times of 350 ms. The data were processed with the Felix 98.0 software ( Biosym Technologies) with appropriate apodization and zero-filling.

RESULTS

Divergent Motif for hTAFII31—To delineate the peptide sequences required for transactivation through interaction with hTAFII31, we first constructed four small libraries of a mammalian expression vector, each encoding a 17-amino acid peptide from the VP16 activation domain (VP16-(469–485)) fused with the GAL4 DNA-binding domain. VP16-(469–485) was chosen because it is the minimal VP16 peptide that binds a fragment of hTAFII31 (hTAFII31-(1–140)) and activates transcription (18). Each of the four libraries consists of random point mutants at one of the four positions within the FXXΦΦ sequence of VP16-(469–485) (Phe^{479}, Thr^{480}, Asp^{481}, Ala^{482}, Leu^{483}). 192 clones from each library were miniprepped and individually transfected into human Jurkat T cells in 96-well plates. We cotransfected the cells with a reporter containing the SEAP gene under the control of five GAL4-binding sites. This permits convenient detection of transcriptional activation through fluorescence change of a phosphatase substrate.

The DNA sequence analyses of positive clones (>50% SEAP activity of the wild type) revealed that the ability of VP16-(469–485) to activate transcription can endure a variety of amino acid substitutions (Fig. 1). The screen identified Trp, Ile, and Leu at the conserved position of Phe^{479}; substitution of Phe^{479} with any one of the three residues exhibited substantial loss in its transcriptional activity. The mutation of Phe^{479} with Val abrogated much of its activation potential, thus validating our screen and indicating the presence of a clear boundary between Leu and Val for the activity. At the positions of Thr^{480} and Ala^{482}, we obtained many clones that activate transcription more than 50% of the wild type. Although hydrophobic residues are favored at these two positions, clones with some of hydrophilic residues such as Tyr also had significant activity. The identity of Leu^{483} is more tightly controlled, since we isolated only Trp, Phe, and Leu as positive clones. Clones with the other bulky hydrophobic residues, i.e., Ile and Val, at this position exhibited less than 20% activity of the wild type. However, the simultaneous substitution of Ala^{482} with a bulky hydrophobic residue (Trp, Phe, Ile, or Leu) restored the transcriptional level of the Ile and Val mutants to >50% of the wild type, indicative of the complementarity between these two adjacent positions.

The peptide sequences examined above were then fused with GST and tested for the ability to bind hTAFII31-(1–140). As shown in Fig. 1, the activation-positive peptides bound hTAFII31-(1–140) to the same extent as the wild type, whereas the binding of activation-deficient peptides was significantly impaired. Thus the strength of the interaction with hTAFII31-(1–140) in vitro correlates with the ability to activate transcription in transfected cells.
Search for hTAFII31-binding Sequences—Guided by the information obtained from the screen, we searched for potential hTAFII31-binding sequences in human activation domains. A series of selection steps was carried out on the activation domains of 65 distinct human transcription factors. The amino acid sequences of these 65 activation domains and their original references are available on the World Wide Web and in the supplemental materials. In an initial step, we selected for any activation domains that contain a signature P1-P2-X-P3-P4 sequence (where P1 represents Phe, Trp, Ile, or Leu; P2 represents Ile, Leu, Tyr, Trp, Met, Asn, Ala, Thr, Val, Ser, Glu, or Gln; X represents any amino acid; and P3-P4 represents Trp/Phe/Ile/Leu-Trp/Phe/Leu/Ile/Val or Ala/Tyr/Val/Cys/Met/Trp/Phe/Leu) and found 26 that did so. Those activation domains whose signature sequences were not conserved among species and subtypes were eliminated in a second step, and a third elimination step was then run on the remaining candidates by calculating the probability of α-helix formation of their signature sequences. 15 candidates were eliminated by the second and third steps. For example, the activation domain of cell cycle regulator E2F1 was eliminated because its signature sequence, FSGLL, was not conserved in its chicken homolog (FPFGFL) and because these sequences had little α-helix probability. The activation domain of MSG1 was also eliminated because its signature sequence, LMSLV, was not conserved in a subtype of its mouse homolog (LTSLE). As expected, the activation domains of E2F1 and MSG1 had no detectable affinity to hTAFII31, validating our elimination steps (data not shown). 11 activation domains survived all three elimination steps. These are the activation domains of ALL1, NF-IL6, NFAT1, Sox-4, MyoD, c-Jun, HIF-1α, TEF-1, HSF-1, TREB5, and ESX (Fig. 2A). The peptide segments that correspond to VP16-(469–485) were fused to GST and assayed for the ability to bind hTAFII31. Only the peptides of ALL1, NF-IL6, ESX, HSF-1, and NFAT1 bound hTAFII31 as tightly as VP16-(469–485) (Fig. 2B).

The full-length activation domains of these factors were further analyzed. NFAT1 (nuclear factor of activated T cells 1) belongs to the NFAT family of transcription factors and plays a central role in inducible gene transcription during the immune response. Whereas the full-length activation domain of NFAT1 (amino acids 1–96) bound hTAFII31-(1–140) to the same extent as its peptide version, mutation of Phe30 and Phe34 with Ala impaired its interaction (Fig. 3A), indicative of direct involvement of the signature sequence in the interaction with hTAFII31. ALL1 (acute lymphoblastic leukemia gene product; also referred as HRX or MILL) is a human transcription factor that is involved in acute lymphoblastic leukemia (20, 21), and its transcriptional activity is considered to be responsible for malignant transformation (22). We identified the Ile2849-Met2850-Asp2851-Phe2852-Val2853 sequence in its activation domain as a hTAFII31-binding motif. This was in good agreement with the previous mutational studies showing the importance of these residues in transactivation (22). As shown in Fig. 3A, the full-length activation domain of ALL1 (amino acids 2829–2883) bound hTAFII31-(1–140) as tightly as its peptide version. Substitution of Ile2849 and Val2853 by Ala greatly impaired the interaction, indicative of the involvement of the IMDPFV sequence in the interaction. NF-IL6 (nuclear factor interleukin-6; also referred as C/EBPβ or LAP) induces cytokine genes and has been implicated as a master regulator of the acute-phase response (23). Its full-length activation domain (amino acids 24–124) (24) bound hTAFII31-(1–140) to the same extent as its peptide version. This interaction appears to be mediated by the LSDLF sequence, since substitution of Leu118 and Phe122 by Ala impairs the interaction. The (L/F)(S/A)DLF sequence is conserved among the activation domains of C/EBPs, and amino acid substitutions in the conserved region in rC/EBPα adversely affect its transactivation potential (25, 26). HSF-1 (heat shock factor 1) responds to a multitude of stress conditions and plays an important role in the molecular response to nonnative proteins (27), and its activation domain is known to be highly potent (28). We found a hTAFII31-binding motif in its COOH-terminal half, which was consistent with the previous transactivation studies (29). Whereas the full-length activation domain of HSF-1 (amino acids 371–430) bound hTAFII31-(1–140) to the same extent as its peptide version, mutation of Leu414 and Phe418 with Ala impaired its interaction (Fig. 3A), indicative of direct involvement of the signature sequence in the interaction with hTAFII31. Last, we identified a hTAFII31-binding sequence in the activation domain of ESX (an epithelium-restricted Ets factor) that regulates the expression of the HER2/neu (c-erbB2) oncogene in human breast cancer (30) and has been found to be overexpressed at an early stage of human breast cancer development (31). Once again, the ESX activation domain (amino acids 129–159) bound hTAFII31-(1–140) as tightly as its peptide version. Substitution of Ile138 and Leu143 by Ala compromised this interaction, indicating that the interaction is mediated at least in part by the IELLEL sequence.

To analyze the ability to activate transcription, each activator was fused with the GAL4 DNA-binding domain, and its expression plasmid was transfected into human Jurkat T cells
along with the reporter plasmid driven by five copies of the GAL4-binding element. As shown in Fig. 3B, the interaction-deficient mutants of the activation domains were correspondingly unable to activate transcription of the reporter gene, whereas the wild type proteins activated transcription, similar to the VP16 activation domain (VP16452–490). This functional reduction of the mutants is not the result of differences in their expression levels as judged by Western analyses. Thus, the hTAFII31-binding sequences of NFAT1, ALL1, NF-IL6, HSF-1, and ESX are critical for their ability to activate transcription; perhaps hTAFII31 directly mediates the transcriptional activation by these human factors. However, it is not impossible to imagine that the same amino acids involved in the interaction with hTAFII31 may also interact with surfaces in some other co-activators.

Comparison of hTAFII31-binding Sequences—including p53 and p65, we have now obtained a total of seven human activation domains whose activities are critically dependent on their hTAFII31-binding sequences. Their amino acid sequences are compared in Fig. 2A. Apart from the COOH-terminal signature sequences, the hTAFII31-binding peptides have no sequence similarity among themselves. Nonetheless, the NH2-terminal nonhomology region is necessary for binding because truncation of the NH2-terminal five residues in ALL1-(2839–2855) and VP16-(469–485) abolish hTAFII31 binding (18) (Fig. 4A).

To compare the conformations of the hTAFII31-binding peptides upon binding to hTAFII31, each peptide was chemically synthesized and analyzed by transferred nuclear Overhauser effects (TRNOE), an NMR technique that provides conformational information of a small ligand interacting weakly with its macromolecular receptor (18, 32). Only ALL12839–2855 and NF-IL6108–124 showed good physical properties under NMR conditions and were amenable to TRNOE analyses. Amide regions of 350-ms NOESY spectra of ALL1-(2839–2855) (B) and NF-IL6-(108–124) (~4 mM) (C) in the presence of hTAFII31-(1–140) (240 μM). The identities of residues that exhibit NOE cross-peaks are indicated. Sequential NOEs characteristic of a helix formation are summarized in the lower panels. For the NMR analysis of ALL1, the peptide in which Cys3841 was substituted with Ser was used for technical convenience. The GST pull-down experiment independently verified that this substitution had no effect on the interaction with hTAFII31-(1–140).

**Fig. 3.** A, in vitro binding assay. It is evident that GST fusions of the activation domains of ALL1 (lane 3), NF-IL6 (lane 5), ESX (lane 7), HSF-1 (lane 9), and NFAT1 (lane 11) bind hTAFII31-(1–140), whereas the glutamine-rich activation domain of Sp1 (lane 13) and GST alone (lane 14) have no affinity to hTAFII31-(1–140). Substitutions of key hydrophobic residues in the signature motifs greatly impair interactions (lanes 4, 6, 8, 10, and 12). The position of hTAFII31-(1–140) is indicated. B, activities of the activation domains of ALL1, NF-IL6, ESX, HSF-1, NFAT1, and their mutants in human cells when fused with the DNA-binding domain of GAL4. The expression construct of each GAL4 fusion was transiently transfected into human Jurkat T cells along with a reporter gene that expresses SEAP under the control of five GAL4-binding sites. SEAP activities were monitored by fluorescence.

**Fig. 4.** A, truncation study. Deletion of the NH2-terminal five residues from ALL1-(2839–2855) impaired the interaction with hTAFII31-(1–140) (compare lanes 2 and 3). B and C, transferred NOE NMR experiments. Amide regions of 350-ms NOESY spectra of ALL1-(2839–2855) (B) and NF-IL6-(108–124) (~4 mM) (C) in the presence of hTAFII31-(1–140) (240 μM). The identities of residues that exhibit NOE cross-peaks are indicated. Sequential NOEs characteristic of a helix formation are summarized in the lower panels. For the NMR analysis of ALL1, the peptide in which Cys3841 was substituted with Ser was used for technical convenience. The GST pull-down experiment independently verified that this substitution had no effect on the interaction with hTAFII31-(1–140).
Divergent Motifs in Activation Domains

these cross-peaks and the presence of long range $d_{35}(i, i + 3)$ and $d_{36}(i, i + 3)$ NOEs suggest the formation of short $\alpha$-helices encompassing the signature motifs in ALL1-(2839–2855) and NF-IL6-(108–124). No TRNOEs were observed with a control 17-amino acid peptide that has a similar acidity/hydrophobicity profile but no affinity to hTAF$_{31}$. indicating that the interaction with hTAF$_{31}$-(1–140) under the NMR condition is specific (data not shown). These results support the notion that the interactions between activation domains and hTAF$_{31}$ are mediated generally by short $\alpha$-helices in the activators.

In ALL1-(2839–2855) and NF-IL6-(108–124), the NH$_2$-terminal halves appear to be in an extended conformation, and we failed to detect any NOEs that suggest the formation of a folded structure in this region. These nonhomology segments thus may make variable contributions to the association, possibly by lowering the energetic barrier for helix formation or by making additional contacts with the surface of hTAF$_{31}$, perhaps including those between main chain amide groups in the peptides and chemically complementary functional groups in hTAF$_{31}$.

**DISCUSSION**

Our analyses indicate that hTAF$_{31}$ recognizes a diverse set of peptide sequences in activation domains. There are two advantages for the cells in using such a promiscuous interaction for transactivation. One is the weakness of the interaction; the dissociation constant of the interaction between hTAF$_{31}$ and the VP16 activation domain is in the high micromolar range, and the weakness of the interactions is often translated to the diversity of binding sequences. Synergism of such weak interactions between activators and co-activators makes transactivation signals diverse and steep enough to emulate a binary switch (3, 33, 34). Low affinity interactions also permit dynamic modulation in response to the alteration of signals that high affinity interactions would be unable to generate (35). Therefore, the coupling of weak interactions with transcriptional activation may be ideal for eukaryotic cells that respond to various signals in a highly tuned manner.

Another advantage is the fact that each one of the binding sequences can be unique enough to be recognized specifically by its regulatory proteins. This permits specific modulation of activity of particular transcription factors in response to the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognizes and masks the alteration of signals. For example, the MDM2 protein, a cellular attenuator of p53, specifically recognize...