

The activation domain of the enhancer binding protein p45NF-E2 interacts with TAF_{II}130 and mediates long-range activation of the α - and β -globin gene loci in an erythroid cell line

PERSIS J. AMROLIA*, LAKSHMAN RAMAMURTHY*, DAMAN SALUJA†, NAOKO TANESE†, STEPHEN M. JANE‡§, AND JOHN M. CUNNINGHAM*§¶

*Division of Experimental Hematology, St. Jude Children's Research Hospital, Memphis, TN 38101; †Department of Microbiology, New York University Medical Center, New York, NY 10016; and ‡Rotary Bone Marrow Research Laboratory, Royal Melbourne Hospital, Parkville, Victoria 3050, Australia

Communicated by Robert Tjian, University of California, Berkeley, CA, July 14, 1997 (received for review March 7, 1997)

ABSTRACT We have used the interaction between the erythroid-specific enhancer in hypersensitivity site 2 of the human β -globin locus control region and the globin gene promoters as a paradigm to examine the mechanisms governing promoter/enhancer interactions in this locus. We have demonstrated that enhancer-dependent activation of the globin promoters is dependent on the presence of both a TATA box in the proximal promoter and the binding site for the erythroid-specific heteromeric transcription factor NF-E2 in the enhancer. Mutational analysis of the transcriptionally active component of NF-E2, p45NF-E2, localizes the critical region for this function to a proline-rich transcriptional activation domain in the NH₂-terminal 80 amino acids of the protein. In contrast to the wild-type protein, expression of p45 NF-E2 lacking this activation domain in an NF-E2 null cell line fails to support enhancer-dependent transcription in transient assays. More significantly, the mutated protein also fails to reactivate expression of the endogenous β - or α -globin loci in this cell line. Protein-protein interaction studies reveal that this domain of p45 NF-E2 binds specifically to a component of the transcription initiation complex, TATA binding protein associated factor TAF_{II}130. These findings suggest one potential mechanism for direct recruitment of distal regulatory regions of the globin loci to the individual promoters.

Tissue and developmental specificity of eukaryotic gene expression is influenced by regulatory sequences in core promoter regions and distal enhancer elements (1). The activity of these sequences is dependent on the binding of ubiquitous and tissue-specific transcription factors (1, 2). In the context of proximal promoter sequences, numerous protein-protein interactions have been demonstrated between the activation domains of promoter-bound factors and components of the transcription initiation complex (2–4). These interactions are essential for high levels of gene expression. The general transcription factor complex TFIID, which binds to the TATA box, and an associated component TFIIB, have been identified as the major targets of these upstream transactivators (4, 5). TFIID consists of a core TATA-binding protein (TBP), which is sufficient for promoter recognition and basal transcription, and TBP-associated factors (TAFs) that are required for activated transcription. TAFs bind to TBP in an ordered fashion and provide a physical and functional link between upstream activators and the RNA polymerase holoenzyme (6).

In contrast to the well characterized role of factors bound to the core promoter, the mechanism of action of a similar array of factors binding to distal enhancers remains unclear. Many

eukaryotic enhancers reside kilobases away from the genes whose expression they influence, and several models have been proposed to explain their long-range action. The scanning model suggests that proteins binding to distal elements track along the DNA until they reach the promoter where they interact with the basal machinery (7). In contrast, the looping model predicts that enhancer-bound transactivators are juxtaposed to proteins bound to the proximal promoter, with looping out or bending of the intervening DNA (8, 9). A third model suggests that binding of these transcription factors to the enhancer alters the tertiary structure of the downstream promoter (chromatin opening), allowing greater access to promoter binding transcriptional activators (9).

We have used the human β -globin cluster as a model to explore the mechanisms of promoter/enhancer interactions in the context of a multigene locus (10, 11). Tissue and developmentally specific expression of the genes of this locus (ϵ , γ , δ , and β) is dependent on sequences in the core promoters (11). However, high-level expression requires the presence of the powerful regulatory elements of the locus control region (LCR), located 6–20 kb upstream of the ϵ -globin gene (12, 13). The LCR consists of four erythroid-specific DNase I hypersensitive sites (HS1–4) (14, 15), which contain a highly conserved array of binding sites for hematopoietically restricted and ubiquitous transcription factors (10, 11). Recent evidence indicates that the HSs of the LCR function as a multiprotein holocomplex, interacting with a single gene promoter at any given time point to achieve high-level globin gene expression (16). The concept that no single site is critical for LCR activity and that considerable redundancy exists is further supported by the modest phenotypes observed with deletion of single sites (17, 18).

Using HS2, we have studied enhancer-dependent transcription of the globin genes. A tandem AP-1/NF-E2 binding motif forms the core of the HS2 enhancer and is essential for its function (19, 20). Similar sites are found in all HSs of the LCR in humans and other species, as well as the HS-40 enhancer of the α -globin cluster (21, 22). In contrast, no NF-E2 binding sites have been identified in the globin promoters. The NF-E2 motif binds a heteromeric complex consisting of a hematopoietic-specific 45-kDa subunit (p45 NF-E2) (23, 24), a member of the cap'n-collar family of transcription factors, and a ubiquitously expressed 18-kDa subunit (p18 NF-E2) (25, 26), a member of the NRL/maf family of DNA-binding proteins. Both proteins contain a basic region-leucine zipper motif. The NH₂-terminal half of p45 NF-E2 also contains a proline- and serine-rich domain, previously reported to act as a transcriptional activator (27). The DNA binding

Abbreviations: HS, hypersensitivity site; LCR, locus control region; TAF, TATA binding protein associated factor; TBP, TATA-binding protein; DB, DNA-binding domain; HA, hemagglutinin; GST, glutathione S-transferase; MEL, murine erythroleukemia.

§S.M.J. and J.M.C. contributed equally to this work.

¶To whom reprint requests should be addressed. e-mail: john.cunningham@stjude.org.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/9410051-6\$2.00/0 PNAS is available online at <http://www.pnas.org>.

specificity of the complex is conveyed by the p18 subunit, which has no transactivation potential (25).

Transcriptional activation of the β -globin gene by activator sequences of the LCR is observed in the presence of a minimal β -globin promoter possessing only an intact TATA box (28). In the context of HS2, activation is dependent on the presence of the tandem NF-E2 binding sites (19, 20). These findings suggest that an interaction between NF-E2 and the transcription initiation complex may provide a functional link between the distal LCR and the individual globin gene promoters. We report that the trans-activation domain of p45 NF-E2 is critical for its ability to support enhancer-dependent transcription of the globin genes in transient assays and in the context of the endogenous β - and α -globin loci. We demonstrate that this domain interacts with a component of the human TFIID complex, hTAF_{II}130 (29). These findings suggest a possible mechanism for long-range transcriptional activation of the globin genes.

MATERIALS AND METHODS

Plasmid Construction. Plasmids containing the wild-type or TATA box mutated human A γ -globin promoter (-356 \rightarrow +35) upstream of the firefly luciferase gene and the 1.9-kb fragment of HS2 of the human β -globin locus have been described previously (30). Yeast two-hybrid plasmids containing the GAL4 DNA-binding domain (DB) fused to polymerase complex components were made by subcloning the coding sequences of human TBP, TFIIB, and TAFs 250, 70 and 55, 30 and *Drosophila* TAF_{II}110 into PGBT9 (CLONTECH). A similar approach was used for the partial clone of hTAF_{II}130. Yeast NF-E2 expression vectors (AD-NF-E2, and NH₂ and COOH terminal mutants) were made by subcloning fragments of the human p45NF-E2 cDNA in-frame with the GAL4 activation domain in the pGAD424 vector (CLONTECH), pGEXNF-E2, and NH₂ and COOH-terminal mutants of this construct were made in a similar fashion. PCIGAL, a mammalian expression vector containing the GAL4 DNA-binding domain, was made by subcloning the blunted *HindIII*-*SalI* fragment of pGBT9 into the *SalI*-blunted *NheI* fragment of PCINeo. COOH-terminal truncation mutants of p45 NF-E2 were inserted 3' to the GAL4 DNA-binding domain, by subcloning the same fragments used to generate the yeast mutants outlined above, into the blunted *NotI*-*EcoRI* fragment of PCIGAL. An oligonucleotide encoding a consensus Kozak motif and the sequence of a hemagglutinin (HA) tag was inserted into the mammalian expression vector PCINeo, to generate PCINeoHA. NF-E2 cDNA fragments were subcloned in-frame with the HA tag. pSPTAF110 was made for *in vitro* transcription/translation studies by subcloning the complete dTAF_{II}110 cDNA into the pSP72 vector (Promega).

DNA Transfection and Transient Assays. DNA was transfected into K562 human erythroleukemia and CB3 mouse erythroleukemia cells by electroporation as described previously (31), except that CB3 cells were electroporated at a slightly lower cell density of 10⁸/ml. Five micrograms of a β galactosidase-expressing plasmid was included in all transfections in K562 cells and 10 μ g in CB3 transfections as an internal control. In transactivation experiments using the chloramphenicol acetyltransferase reporter gene under the control of tandem GAL4 binding sites, 20 μ g of pG5EC reporter plasmid was cotransfected with equimolar amounts of the PCIGALNF-E2 mutants equivalent to 1 μ g of the wild-type construct in the presence of carrier plasmid (PCINeo) to a total of 50 μ g. In CB3 cotransfection experiments, 20 μ g of the expression vector was electroporated with 20 μ g of the reporter plasmid. CB3 cells were preinduced with 1.8% dimethyl sulfoxide 72 hr before electroporation and K562 cells with 20 μ M hemin at the time of electroporation. After 48 hr, cells were harvested, lysed, and assayed as previously described (31). Results shown represent at least six different transfections with two independent plasmid preparations. Statis-

tics were calculated using Student's *t* test; values of *P* < 0.05 were considered significant.

Establishment of Stable CB3 Cell Lines Expressing p45 NF-E2. CB3 cells (32), a gift of Yacov Ben-David (Sunnybrook Health Science Center, Toronto, Canada), were stably transfected with eukaryotic vectors expressing NF-E2 and selected in 1 mg/ml G418. Clones were isolated and expanded, and expression of the transgene was analyzed by Western blotting and electrophoretic mobility shift assay using previously described procedures (27).

Yeast Two-Hybrid Assays. Yeast strains HF7c and SFY526 were obtained from CLONTECH and transformations performed using the lithium acetate method (33). All constructs initially were transformed singly into SFY526 and transformants assayed for β -galactosidase to exclude intrinsic activation. Pairs of activation domain and DNA binding domain hybrid constructs then were used to cotransform the host strains, with selection on media lacking both leucine and tryptophan. Double transformants then were assayed for β -galactosidase activity using a filter assay or, in the case of HF7c, were replated on media lacking histidine, leucine, and tryptophan with varying concentrations of 3-aminotriazole.

Expression of Glutathione S-Transferase (GST) Fusion Proteins and Affinity Chromatography. GST fusion proteins were expressed in BL21 cells. Fusion proteins were purified (34) and quantified using the Bio-Rad method after purification on glutathione-Sepharose beads (Pharmacia). SDS/PAGE gel electrophoresis with Coomassie staining was used to confirm the integrity of the full-length fusion protein. For *in vitro* protein-protein interaction assays, 1 μ g of GST or GST-fusion protein was incubated for 1 hr at 4°C with 10 μ l glutathione-Sepharose beads, which had been preblocked with 0.5% milk. After extensive washing, the beads were resuspended in 200 μ l binding buffer (10 mM Tris-HCl, pH 7.9/500 mM KCl/0.1 mM EDTA/150 μ g/ml BSA/0.1% Nonidet P-40/10% glycerol) and incubated for 1 hr at room temperature with ³⁵S methionine-labeled dTAF_{II}110 or hTAF_{II}130. After extensive washing, retained proteins were eluted by boiling in SDS loading buffer and analyzed on a 10% SDS/PAGE gel.

RNase Protection Assays. Isolation of total RNA and RNase protection analysis were performed as previously reported (27, 34).

RESULTS

HS2 Enhancer-Dependent Transcriptional Activation of the γ -Globin Promoter Requires a Functional TATA Box and Intact NF-E2 Binding Sites. To define the minimal elements necessary for enhancer-dependent transcriptional activation of the γ -globin promoter, we linked a 1.9-kb fragment of HS2, containing intact tandem AP1/NF-E2 binding sites (HS2 γ TATALUC, Fig. 1, construct 2) or an identical fragment in which the NF-E2 sites had been mutated (HS2^S γ TATALUC, Fig. 1, construct 4) to a minimal γ -promoter/luciferase gene hybrid. These constructs, and the minimal γ -promoter/luciferase hybrid lacking upstream regulatory sequences (γ TATALUC, Fig. 1, construct 1) and a construct in which the TATA box was mutated (HS2 γ TATA^SLUC, Fig. 1, construct 3) were transiently transfected into the human erythroid cell line, K562. As seen in Fig. 1, the presence of a functional TATA box was sufficient to allow enhancer-dependent transcription (Fig. 1, constructs 1 and 2). This transcriptional activation was lost with mutation of the TATA box or the NF-E2 binding sites in HS2 (Fig. 1, constructs 3 and 4). Mutation of other protein binding sites within HS2 failed to ablate enhancer activity (P.J.A., J.M.C., and S.M.J., unpublished work), suggesting that the minimal requirements for enhanced transcription are a TATA box in the promoter and intact NF-E2 binding sites in the enhancer.

p45 NF-E2 Interacts Specifically with *Drosophila* TAF_{II}110 and its Human Homologue TAF_{II}130. A direct interaction between NF-E2 and a component of the TATA box binding TFIID

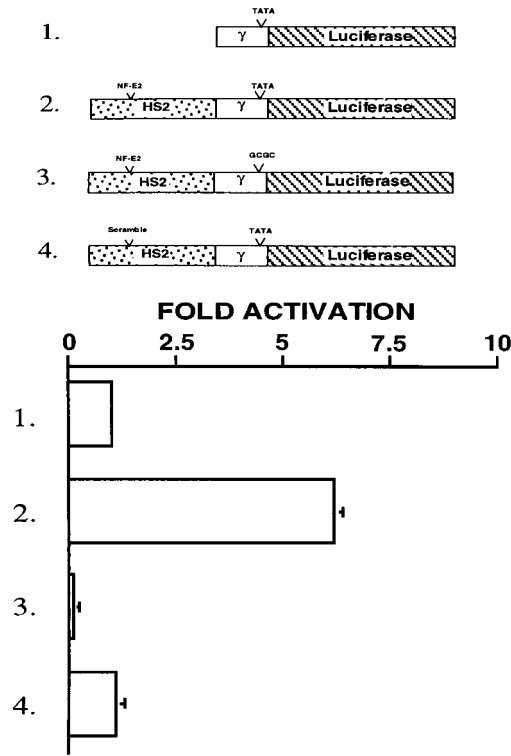


FIG. 1. A functional TATA box and the core NF-E2 sites are required for enhancer-dependent transcriptional activation. (Upper) The constructs used to examine the interaction between the HS2 enhancer and minimal γ -globin promoter. The open box represents the -356+35 minimal γ -globin promoter. The hatched box represents the luciferase reporter gene. The stippled box represents the 1,900-bp *KpnI-BglIII* fragment of HS2 of the β -globin LCR. The NF-E2 motifs are approximately 1 kb upstream of the TATA box. Constructs 3 and 4 contain mutated TATA box and NF-E2 binding motifs, respectively. (Lower) The fold activation by the enhancer relative to expression of construct 1.

complex would provide one mechanism explaining the above results (3, 35). To investigate this, we focused on the hematopoietic-specific p45 component of NF-E2, as the transcriptional activity of the complex appears to reside solely in this subunit (25, 26). Using the yeast two-hybrid assay system, we examined protein-protein interactions between p45 NF-E2 and TBP, TFIIB, and TAF_{II}s 250, 110, 70, 55, and 30. As shown in Fig. 2A, a specific interaction was observed between NF-E2 and the *Drosophila* TAF, dTAF_{II}110. No interactions were observed with the other components of TFIID, despite immunoblot analysis of yeast extracts confirming high-level expression of all transfected constructs (data not shown). Control experiments, removing either TAF_{II}110 or NF-E2 from their respective yeast plasmids, ablated reporter gene activity (Fig. 2B). We subsequently obtained a partial clone of TAF_{II}130, the human homologue of TAF_{II}110, and confirmed that this interaction also occurred between the two human proteins (data not shown).

To validate this interaction in an independent system, GST affinity chromatography was performed. GST alone or GST fused in-frame with full-length p45 NF-E2 (GST NF-E2) were coupled to glutathione-Sepharose beads and incubated under stringent conditions with ³⁵S methionine-labeled recombinant dTAF_{II}110. Specific retention of dTAF_{II}110 was observed with the GST NF-E2 beads, but not on control GST beads (Fig. 2C). In contrast, incubation of GST NF-E2 with the other components of the TFIID complex failed to show a specific interaction, in accord with our studies in the yeast two-hybrid assay (data not shown).

The Transactivation Domain of NF-E2 Colocalizes with its TAF Interacting Domain. Numerous transcription factors have been shown to interact with their cognate TAFs through their

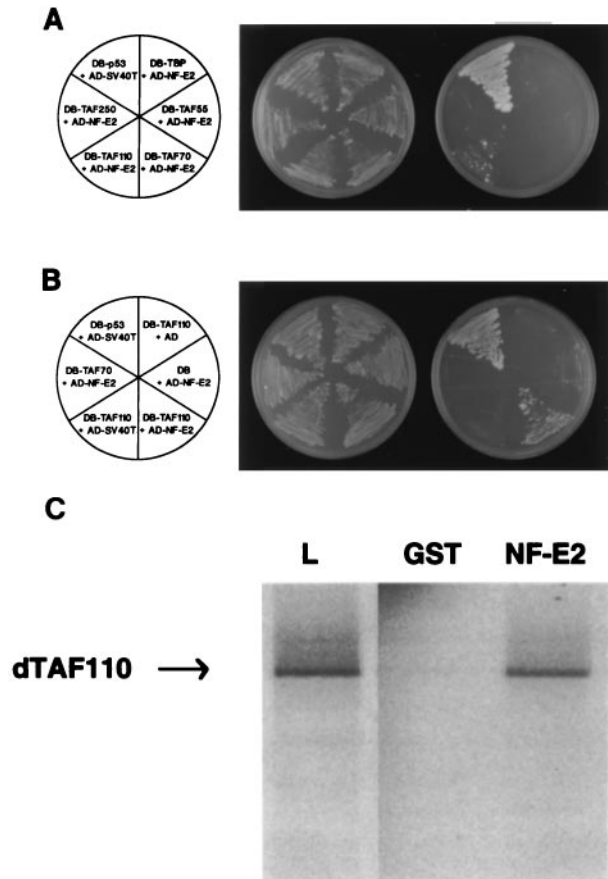


FIG. 2. Human NF-E2 interacts specifically with dTAF_{II}110. (A) The yeast two-hybrid assay demonstrates that p45NF-E2 interacts with dTAF_{II}110. The *Saccharomyces cerevisiae* reporter strain HF7C was transformed with the indicated plasmids. AD-NF-E2 contains the entire coding sequence of human NF-E2 inserted in-frame with the activation domain of GAL4 (amino acids 768-881). DB-TBP, DB-TAF_{II}55, DB-TAF_{II}70, and DB-TAF_{II}250 contain the coding sequence of the human proteins and DB-TAF_{II}110 the coding sequence of the *Drosophila* protein inserted in-frame to the GAL4 DNA binding domain (amino acids 1-147). A specific interaction between DB-p53 and AD-simian virus 40 T antigen has been reported previously. Leu- and Trp- transformants were streaked onto synthetic medium plates lacking tryptophan, leucine, and histidine to assess potential interactions (Right) and synthetic medium plates lacking tryptophan and leucine to confirm plating efficiency (Middle). The plates were incubated at 30°C for 3 days. (B) The yeast two-hybrid assay demonstrates that the NF-E2/dTAF_{II}110 interaction is specific. DB-TAF_{II}110 was plated in combination with the simian virus 40 T antigen or GAL4-AD. AD-NF-E2 was plated in combination with TAF_{II}110, TAF_{II}70, or the GAL4-DB. Transfections were grown in the absence of leucine and tryptophan to assess transformation efficiency (Middle) or in the absence of leucine, tryptophan and histidine to assess protein-protein interactions (Right). Cotransfection of AD-simian virus 40 T-antigen and DB-p53 served as the positive control. (C) GST chromatography confirms that dTAF_{II}110 binds specifically to p45NF-E2. One microgram of GST or GST-NF-E2, bound to glutathione-Sepharose was incubated with *in vitro*-translated dTAF_{II}110 as described in Materials and Methods. Bound proteins were detected by SDS/PAGE followed by autoradiography. An example of labeled dTAF_{II}110 that was loaded is shown in lane L.

transactivation domains (1). Previous studies had suggested that this domain in NF-E2 was located in the first 206 amino acids (27). This proline-rich sequence shares no structural homology with other proteins. To further localize the transactivation domain, we used mutants of NF-E2 fused to the GAL4 DNA binding domain in a eukaryotic expression vector. These constructs were cotransfected into K562 cells with a reporter plasmid containing five tandem GAL4 binding sites upstream of the chloramphenicol acetyltransferase gene. As seen in Fig. 3A, full transcriptional

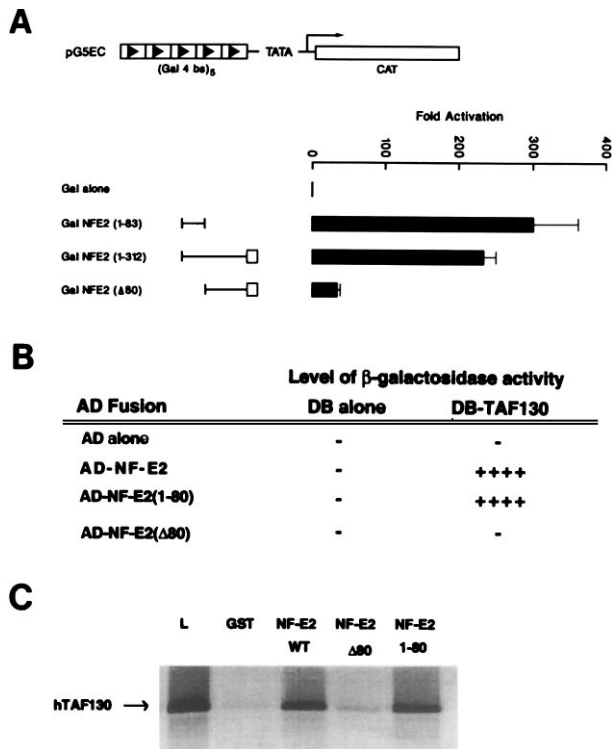


FIG. 3. Colocalization of the activation domain of NF-E2 with the TAF_{II}130 binding domain. (A) Localization of activation domain of NF-E2. K562 cells were cotransfected by electroporation with expression vectors for GAL4 fusion proteins containing the indicated regions of NF-E2 and a reporter construct containing five GAL4 DNA binding sites regulating the expression of the chloramphenicol acetyltransferase gene (pG5EC). The first 312 amino acids of NF-E2 previously have been shown to contain a proline-rich activation domain lacking structural homology to other proteins denoted by a line, and a basic region similar to other cap'n-collar transcriptional activators denoted by an open box. The expression vector used in this study was pCINeo, containing the cytomegalovirus promoter. The fold activation represents the chloramphenicol acetyltransferase activity relative to the expression of the expression vector containing the GAL4 DNA binding domain alone. (B) hTAF_{II}130 interacts with the NF-E2 activation domain in the yeast two-hybrid assay. The *S. cerevisiae* reporter strain SFY526, which contains a β -galactosidase reporter construct under the control of a GAL4 binding motif multimer, was transformed with the indicated plasmids. The AD-NF-E2 constructs contain the stated regions of the coding sequence of human NF-E2 inserted in-frame to the activation domain of GAL4 (amino acids 768–881). DB-TAF_{II}130 contains the partial coding sequence of the human protein inserted in-frame to the GAL4 DNA binding domain (amino acids 1–147). (C) The amino terminal 80 amino acids of NF-E2 interacts specifically with hTAF_{II}130 in the GST chromatography system. Fusion proteins containing GST fused to amino acids 1–374, 1–80, or 80–374 of NF-E2 or GST alone were expressed in *Escherichia coli* and bound to glutathione-Sepharose beads. The beads then were incubated with ³⁵S-labeled *in vitro*-translated hTAF_{II}130. Other experimental details are as stated in Fig. 2C.

activation was observed in the presence of the NH₂ terminal 80 amino acids of NF-E2 [GALNF-E2 (1–80)] alone. Deletion of this domain (GALNF-E2 Δ 80) reduced reporter gene expression to the levels of vector control. To assess whether this minimal activation domain was sufficient for the interaction with TAF_{II}130 we examined NF-E2(1–83) in the yeast two-hybrid assay. As seen in Fig. 3B, a GALNF-E2 (1–80) fusion protein activated the β -galactosidase reporter gene when cotransfected with TAF_{II}130. In contrast, GALNF-E2 Δ 80 was incapable of interacting. To confirm these observations, we incubated ³⁵S methionine-labeled hTAF_{II}130 with GST, GST NF-E2, and selected mutants (Fig. 3C). Beads loaded with GST NF-E2 and GST NF-E2 (1–80) specifically retained hTAF_{II}130. In contrast,

beads linked to GST or GST NF-E2 Δ 80 failed to bind significant amounts of hTAF_{II}130.

The TAF_{II}130 Interacting Domain of NF-E2 Is Required for Enhancer-Dependent Transcription of the Globin Genes. To ascertain whether the TAF_{II}130 binding domain of NF-E2 was necessary for the enhancer activity of HS2, we constructed eukaryotic expression vectors containing wild-type NF-E2, or NF-E2 lacking the TAF_{II}130 interaction domain (NF-E2 Δ 80). These were stably transfected into CB3 cells, a murine erythroleukemia (MEL) cell line in which both p45 NF-E2 alleles are nonfunctional (32). As a consequence, this line is NF-E2 null and fails to express significant levels of endogenous α - or β -globin chains (27, 32). Expression of wild-type and mutant forms of NF-E2 was verified in individual clones using electrophoretic gel mobility shift assay (Fig. 4A). The parental CB3 line, two derived

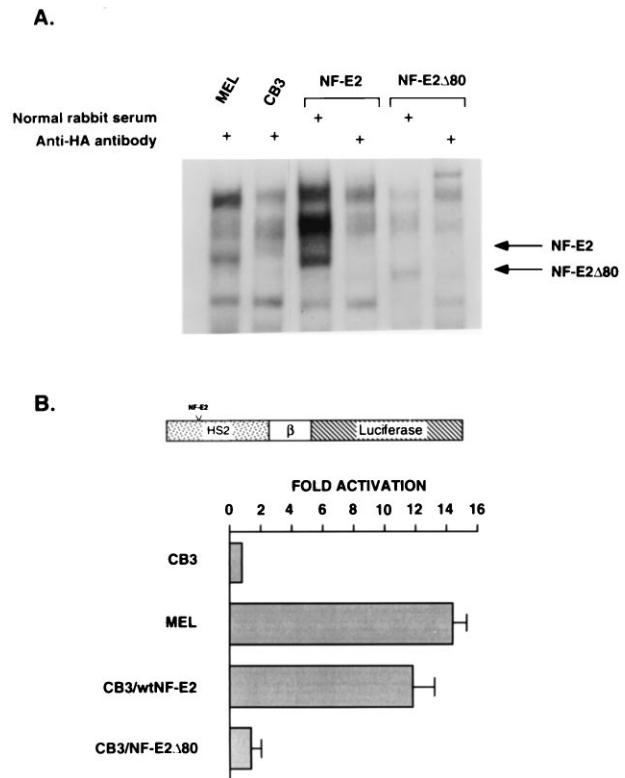


FIG. 4. The TAF_{II}130 binding domain of NF-E2 is essential for enhancer-dependent activation of the β -globin gene. (A) Wild-type NF-E2 and NF-E2 Δ 80 expression in selected CB3 clones as demonstrated in a gel mobility shift assay. CB3 cells were stably transfected with pCINeoHA vector alone or pCINeoHA vectors in which the human NF-E2 cDNA or NF-E2 Δ 80 was fused in-frame with the HA tag. Nuclear extract was prepared from individual clones and analyzed by gel mobility shift assay, dimethyl sulfoxide-induced MEL extract serving as a control. ³²P-labeled DNA fragments containing the tandem NF-E2 binding sites in HS2 were incubated with individual extracts and electrophoresed on a 4% polyacrylamide gel. Lanes 1 and 2 show MEL cell and CB3 control extracts, the latter lacking the characteristic NF-E2 complex. Lane 3 shows a representative CB3 clone transfected with pCINeoHANF-E2 demonstrating a complex that comigrates with the NF-E2 complex observed in MEL cells. Lane 4 verified the presence of HA-tagged NF-E2 with specific ablation of the NF-E2 complex by the addition of anti-HA antibody (anti-HA). Lanes 5 and 6 show similar results for a clone expressing pCINeoHANF-E2 Δ 80. (B) High-level expression of a transiently transfected β -globin promoter containing construct is dependent on wild-type NF-E2 expression in CB3 cells. CB3 clones expressing wild-type NF-E2 or NF-E2 Δ 80 or CB3 cells transfected with the pCINeo vector (control) were transiently transfected with a plasmid containing HS2 linked to a β -promoter/luciferase gene hybrid. Whole cell extracts were assayed for luciferase activity. The results shown are representative of the three transfection experiments. MEL cells were transfected as a positive control. The fold activation observed was calculated relative to the expression of the HS2 β Luciferase construct in CB3 cells.

lines for each NF-E2 expression vector, and a control MEL cell line (which constitutively expresses α - and β -globin) were then transiently transfected with a plasmid containing a β -promoter/luciferase gene hybrid linked to HS2. Enhancer-dependent activation of the β -globin promoter was observed in the MEL line and in the CB3 lines expressing wild-type NF-E2 (Fig. 4B). In contrast, the CB3 lines expressing NF-E2 Δ 80 failed to support HS2 enhancer activity, suggesting that enhancer function is dependent on the integrity of the TAF_{II}130 binding domain of NF-E2.

The TAF_{II}130 Interacting Domain of NF-E2 Is Required for Endogenous β - and α -Globin Gene Expression. Although our studies have used HS2, the presence of NF-E2 binding sites in all HS of the β -globin locus, and in HS-40 of the α -globin locus (10, 11, 21, 22) suggested that the amino terminal 80 amino acids of NF-E2 may be critical for activation of all globin genes by upstream regulatory sequences in the endogenous loci. To address this, we examined the expression of the endogenous globin genes in the CB3 lines expressing wild-type NF-E2 or NF-E2 Δ 80. As previously reported, activation of the β -globin gene to levels equivalent to those seen in MEL cells was observed in the presence of wild-type NF-E2 (Fig. 5, Table 1, and refs. 27 and 32). In comparison, NF-E2 Δ 80 failed to reactivate the endogenous β -globin gene with comparable RNA levels observed between these clones and the parental CB3 line (Fig. 5, compare CB3 and NF-E2 Δ 80; Table 1). Interestingly, activation of the endogenous α -globin locus (Fig. 5 and Table 1) also was observed only in clones expressing wild-type NF-E2. These results indicate that in CB3 cells β -globin LCR- or α -locus HS-40-enhanced expression of their respective globin genes is dependent on the activation domain of NF-E2.

DISCUSSION

We have demonstrated that enhancer-dependent transcription by HS2 is observed in the context of a minimal promoter containing only a TATA box. In addition, we have shown that the integrity

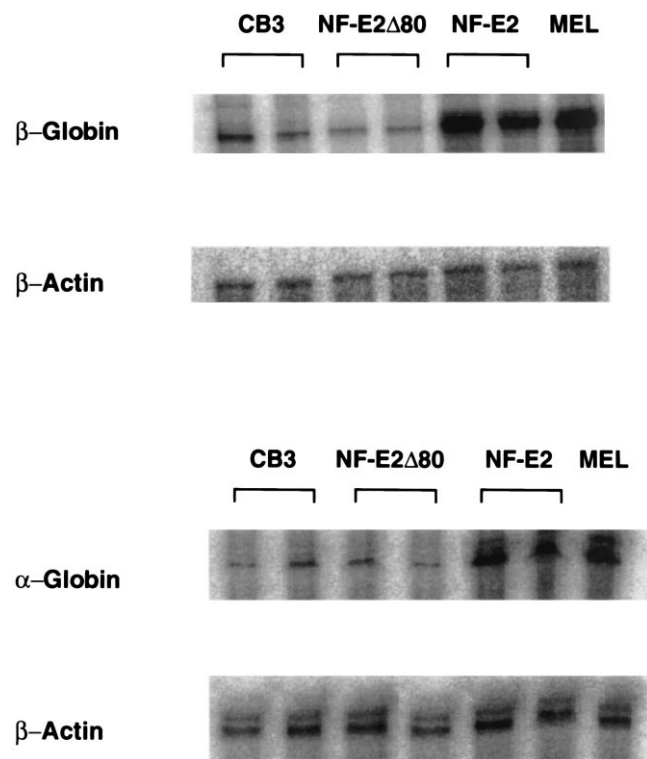


FIG. 5. High-level endogenous globin gene expression is observed only with wild-type p45NF-E2 expression in erythroid cells. RNase protection assays for α -globin, β -major mRNAs from CB3 cells, CB3 cells stably transfected with various p45 NF-E2 expressing constructs, and MEL cells. Actin levels were measured in parallel in all samples.

Table 1. Globin gene expression in transfected CB3 cell lines

	Fold activation	
	β -globin	α -globin
Empty vector	1	1
NF-E2 wild type	5.3 \pm 0.4*	4.1 \pm 0.2*
NF-E2 Δ 80	0.9 \pm 0.3	0.85 \pm 0.2

Individual CB3 clones expressing wild-type or mutant NF-E2 cDNAs were assayed for endogenous globin gene expression using specific α - and β -globin specific probes. Variability in the amount of RNA loaded per assay was corrected internally by analyzing the expression of β -actin (see Fig. 5). Comparison of activation of the globin genes were corrected with respect to untransfected CB3 cells. The results shown are the average of at least three independent experiments.

*Significantly different from CB3 cells alone ($P < 0.01$).

of the tandem NF-E2 binding sites is essential for this transcriptional activation. Both of these findings are in accord with observations made in different experimental systems (19, 20, 28) and can be linked by a model postulating a direct interaction between NF-E2 and the TFIID complex. We have demonstrated such an interaction between the activation domain of p45 NF-E2 and human TAF_{II}130. Deletion of this domain in p45 NF-E2 has dramatic functional sequelae with loss of enhancer-dependent transcription in transient transfection assays. More significantly, it also results in an inability of NF-E2 to reactivate either the endogenous α - or β -globin loci in NF-E2 null cells.

These findings support a model of DNA looping, with juxtaposition of enhancer and promoter sequences through protein-protein interactions. Previous studies have postulated the existence of a DNA loop anchored by a protein bridge between two known factors as the mechanism mediating enhancer/promoter interactions *in vivo* (7, 36). The demonstration that the simian virus 40 enhancer could stimulate transcription *in vitro* when noncovalently linked to a β -globin promoter by the protein streptavidin strengthened this model (8). This stimulation in *trans* mimics transfection, a natural phenomenon in which one chromosome affects gene expression in a paired homologue. Subsequent studies of the transcription factors Sp1 and p53 demonstrated oligomerization of these factors with coincident loop formation of the intervening DNA (37). This multimerization resulted in synergistic activation of a linked reporter gene. Recently, Cullen *et al.* (38) demonstrated that juxtaposition of the distal enhancer and proximal promoter regions of the rat prolactin gene occurred in response to estrogen. By acting through its receptor bound to the distal enhancer, estrogen stimulated the interaction between distal and proximal regulatory regions with resultant transcriptional activation. This interaction was postulated to occur through "tethering" mediated by unknown protein-protein bridges.

Studies in the β -globin locus also have suggested that DNA looping is inherent to LCR/promoter interactions. A single gene promoter interacts with the LCR directly at any given time point. During the fetal/adult switch, the LCR appears to flip-flop between the γ - and β -promoters under the influence of the trans-acting environment. Our studies identify a protein-protein interaction that may provide a functional link between the LCR and individual promoters. This is evident in the context of a transient enhancer assay (Fig. 4B) and also in the setting of the native globin loci (Fig. 5). Our findings do not account for the developmental specificity of gene expression, which appears to be mediated through binding sites for developmental proteins in the individual globin promoters (31, 39, 40). Support for this conclusion is observed in studies of transgenic mice lacking the LCR, which retain a normal developmental profile of globin gene expression, albeit at low levels (reviewed in ref. 11). Our recent studies, demonstrating that transcriptional activation of globin promoters by the factors binding to sequences adjacent to the NF-E2 core of HS2 is dependent on the binding of the developmental proteins to promoter sequences, provides a link between the transcriptional activity of the LCR and the temporal regula-

tion of the globin genes (P.J.A., J.M.C., and S.M.J., unpublished work). The importance of NF-E2 as a cornerstone of LCR/promoter interactions is suggested by numerous studies in cell lines and transgenic mice (11, 19, 20, 41). The distribution of binding motifs for this protein in all human HSs, the HSs in many other species, and the α -locus HS-40 further underscores this conclusion (21, 22, 42, 43). The topology of NF-E2 motifs with respect to other transcriptional activators binding to the HSs is also highly conserved (41), suggesting that this array of proteins may function as a multi-unit complex or enhanceosome. This arrangement is analogous to the stereospecific array of transcriptional activators found in the interferon β and IgH enhancers (44, 45). In the β -globin cluster, looping of the DNA stabilized by a direct interaction between NF-E2 and TAF_{II}130 would result in alignment of other enhancer proteins, eg., GATA-1, with the transcription initiation complex. This would facilitate an interaction between these activators and their cognate TAFs, focusing the transcriptional potential of the LCR on the polymerase complex. This hypothesis is supported by the recent demonstration of synergistic interactions between multiple enhancer-bound proteins and their respective TAFs in *Drosophila* (4).

The p45NF-E2 protein previously has been shown to be essential for high-level globin gene expression in the MEL-derived CB3 cell line (32). Similarly, expression of a dominant negative p18 mutant significantly reduced the expression of α - and β -genes in MEL cells (27). Globin expression in these cells is fully restored upon the introduction of a tethered p45-p18 heterodimer. In accord with our findings, introduction of the NH₂-terminal half of the protein also restored globin expression. These findings are in contrast to the subtle defects in hemoglobin production observed in mice with an NF-E2 null phenotype (46) and suggest that the repertoire of functional NF-E2 binding factors may be more limited in MEL cells than in the developing mouse. Although additional NF-E2-like proteins have been identified (47, 48), their ability to mediate enhancer-dependent transcription is not established. The discrepancy in the action of NF-E2 also may be influenced by the developmental stage with MEL cells arrested at a time in erythroid maturation in which NF-E2 is critical (27, 32).

Examination of the transactivation domain of NF-E2 reveals that it is proline-rich (20%) and serine-rich (10%). Previous studies have demonstrated that this class of domain is associated with factors that act as both "proximal and distal" activators of transcription (49). In contrast, glutamine-rich domains function only as proximal activators, stimulating transcription only from a position close to the TATA box. Interestingly, glutamine-rich factors have been demonstrated to transactivate via an interaction with TAF_{II}130. The binding of this TAF to the proline/serine-rich transactivation domain of NF-E2 suggests that the differential enhancer effect of transactivator domains is not due solely to a specificity of coactivator interaction but may be influenced by other factors, such as the region of interaction of the TAF protein. We currently are defining the domains of TAF_{II}130 necessary for interaction with glutamine- and proline-rich transactivators.

Drosophila and human *in vitro* transcription systems show an absolute requirement for TAFs (reviewed in ref. 1). Cellular studies in yeast suggested that TAFs may not be required for the activation of many genes (50). However, recent studies in *Drosophila* embryos demonstrate an absolute requirement for TAF_{II}s in the developmental function of the activator, bicoid (51). Thus, identification of activator/coactivator interactions are critical to our understanding of the β -globin locus. The studies described here focus on one potential mechanism of LCR/promoter interactions. However, modulation in chromatin structure mediated by the LCR, in part by p45NF-E2 (52, 53), upstream promoter-bound activators, and other components of the transcriptional machinery are necessary for optimal globin gene expression. Integration of these diverse mechanisms will be required to understand the dynamics of globin gene expression.

We thank A.W. Nienhuis, S.H. Orkin, R. Tjian, and J.H. Ihle for helpful discussions. We thank R. Tjian, R. Roeder, and P. Ney for the gift of plasmids and N. Tran and A. Becher for technical assistance. This work was supported by the National Institutes of Health, an American Society of Hematology Junior Faculty Scholars award (J.M.C.), the Cooleys Anemia Foundation (L.R.), the American Lebanese Syrian Associated Charities, the Australian National Health and Medical Research Council, and the Wellcome Trust.

1. Tjian, R. & Maniatis, T. (1994) *Cell* **77**, 5-8.
2. McKnight, S. L. (1996) *Genes Dev.* **10**, 367-381.
3. Tanese, N. & Tjian, R. (1993) *Cold Spring Harb. Symp. Quant. Biol.* **58**, 179-185.
4. Sauer, F., Hansen, S. K. & Tjian, R. (1995) *Science* **270**, 1783-1788.
5. Chiang, C. M. & Roeder, R. G. (1995) *Science* **267**, 531-536.
6. Chen, J. L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. & Tjian, R. (1994) *Cell* **79**, 93-105.
7. Brent, R. & Ptashne, M. (1984) *Nature (London)* **312**, 612-615.
8. Mueller-Sturm, H. P., Sogo, J. M. & Schaffner, W. (1989) *Cell* **58**, 767-777.
9. Felsenfeld, G. (1992) *Nature (London)* **355**, 219-224.
10. Stamatoyannopoulos, G. & Nienhuis, A. W. (1994) in *The Molecular Basis of Blood Diseases*, eds. Stamatoyannopoulos, G., Nienhuis, A. W., Majerus, P. J. & Varmus, H. (Saunders, Philadelphia), pp. 107-156.
11. Dillon, N. & Grosfeld, F. (1993) *Trends Genet.* **9**, 134-137.
12. Grosfeld, F., Blom van Assendelft, G., Greaves, D. R. & Kollias, G. (1987) *Cell* **51**, 975-985.
13. Forrester, W. C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G. & Groudine, M. (1987) *Nucleic Acids Res.* **15**, 10159-10177.
14. Tuan, D., Solomon, W., Li, Q. & London, I. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6384-6388.
15. Talbot, D., Collis, P., Antoniou, M., Vidal, M., Grosfeld, F. & Greaves, D. R. (1989) *Nature (London)* **338**, 352-355.
16. Wijgerde, M., Grosfeld, F. & Fraser, P. (1995) *Nature (London)* **377**, 209-213.
17. Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D. I., Enver, T., Ley, T. J. & Groudine, M. (1995) *Genes Dev.* **9**, 2203-2213.
18. Hug, B. A., Wesselschmidt, R. L., Fiering, S., Bender, M. A., Epner, E., Groudine, M. & Ley, T. J. (1996) *Mol. Cell Biol.* **16**, 2906-2912.
19. Ney, P. A., Sorrentino, B. P., McDonagh, K. T. & Nienhuis, A. W. (1990) *Genes Dev.* **4**, 993-1006.
20. Talbot, D. & Grosfeld, F. (1991) *EMBO J.* **10**, 1391-1398.
21. Gumucio, D. L., Shelton, D. A., Blanchard-McQuate, K., Gray, T., Tarle, S., Heilstedt-Williamson, H., Slightom, J. L., Collins, F. & Goodman, M. (1994) *J. Biol. Chem.* **269**, 15371-15380.
22. Vyas, P., Vickers, M. A., Simmons, D., Ayyub, H., Craddock, C. F. & Higgs, D. R. (1992) *Cell* **69**, 781-793.
23. Andrews, N. C., Erdjument-Bromage, H., Davidson, M. B., Tempst, P. & Orkin, S. H. (1993) *Nature (London)* **362**, 722-728.
24. Ney, P. A., Andrews, N. C., Jane, S. M., Safer, B., Purucker, M. E., Weremowicz, S., Morton, C. C., Goff, S. C., Orkin, S. H. & Nienhuis, A. W. (1993) *Mol. Cell Biol.* **13**, 5604-5612.
25. Andrews, N. C., Kotkow, K. J., Ney, P. A., Erdjument-Bromage, H., Tempst, P. & Orkin, S. H. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11488-11492.
26. Igarashi, K., Kataoka, K., Itoh, K., Hayashi, N., Nishizawa, M. & Yamamoto, M. (1994) *Nature (London)* **367**, 568-572.
27. Kotkow, K. J. & Orkin, S. H. (1995) *Mol. Cell Biol.* **15**, 4640-4647.
28. Antoniou, M. & Grosfeld, F. (1990) *Genes Dev.* **4**, 1007-1013.
29. Tanese, N. T., Saluja, D., Vassallo, M. F., Chen, J. L. & Admon, A. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 13611-13616.
30. Amrolia, P. J., Cunningham, J. M., Ney, P., Nienhuis, A. W. & Jane, S. M. (1995) *J. Biol. Chem.* **270**, 12892-12898.
31. Jane, S. M., Ney, P. A., Vanin, E. F., Gumucio, D. L. & Nienhuis, A. W. (1992) *EMBO J.* **11**, 2961-2969.
32. Lu, S., Rowan, S., Bani, M. R. & Ben-David, Y. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 8398-8402.
33. Fields, S. (1993) *Methods Companion Methods Enzymol.* **5**, 116-124.
34. Jane, S. M., Nienhuis, A. W. & Cunningham, J. M. (1995) *EMBO J.* **14**, 97-105.
35. Hernandez, N. (1993) *Genes Dev.* **7**, 1291-1308.
36. Carey, M., Lin, Y. S., Green, M. R. & Ptashne, M. (1990) *Nature (London)* **345**, 361-364.
37. Stenger, J. E., Tegtmeyer, P., Mayr, G. A., Reed, M., Wang, Y., Wang, P., Hough, P. V. & Mastrangelo, I. A. (1994) *EMBO J.* **13**, 6011-6020.
38. Cullen, K. E., Kladder, M. P. & Seyfried, M. A. (1993) *Science* **261**, 203-206.
39. Perkins, A. C., Sharpe, A. H. & Orkin, S. H. (1995) *Nature (London)* **375**, 318-322.
40. Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. & Grosfeld, F. (1995) *Nature (London)* **375**, 316-318.
41. Stamatoyannopoulos, J. A., Goodwin, A., Joyce, T. & Lowrey, C. H. (1995) *EMBO J.* **14**, 106-116.
42. Li, Q., Zhou, B., Powers, P., Enver, T. & Stamatoyannopoulos, G. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8207-8211.
43. Strauss, E. C., Andrews, N. C., Higgs, D. R. & Orkin, S. H. (1992) *Mol. Cell Biol.* **12**, 2135-2142.
44. Thanos, D. & Maniatis, T. (1995) *Cell* **83**, 1091-1100.
45. Giese, K., Kingsley, C., Kirshner, J. R. & Grosschedl, R. (1995) *Genes Dev.* **9**, 995-1008.
46. Shivdasani, R. A., Rosenblatt, M. F., Zucker-Franklin, D., Jackson, C. W., Hunt, P., Saris, C. J. & Orkin, S. H. (1995) *Cell* **81**, 695-704.
47. Caterina, J. J., Donze, D., Sun, C. W., Ciavatta, D. J. & Townes, T. M. (1994) *Nucleic Acids Res.* **22**, 2383-2391.
48. Moi, P., Chan, K., Asunis, I., Cao, A. & Kan, Y. W. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9926-9930.
49. Seipel, K., Georgiev, O. & Schaffner, W. (1992) *EMBO J.* **11**, 4961-4968.
50. Walker, S. S., Reece, J. C., Apone, L. M. & Green, M. R. (1996) *Nature (London)* **383**, 185-188.
51. Sauer, F., Wassarman, D. A., Rubin, G. M. & Tjian, R. (1996) *Cell* **87**, 1271-1284.
52. Armstrong, J. A. & Emerson, B. M. (1996) *Mol. Cell Biol.* **16**, 5634-5644.
53. Gong, Q., McDowell, J. & Dean, A. (1996) *Mol. Cell Biol.* **16**, 6055-6064.