Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1.

D I Martin and S H Orkin

References	This article cites 52 articles, 27 of which can be accessed free at: **http://www.genesdev.org#References** Article cited in: **http://www.genesdev.org/cgi/content/abstract/4/11/1886#otherarticles**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

Notes

To subscribe to *Genes and Development* go to:
http://www.genesdev.org/subscriptions/

Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1

David I.K. Martin and Stuart H. Orkin[1]

Division of Hematology–Oncology, Children's Hospital and the Dana Farber Cancer Institute, Department of Pediatrics, Harvard Medical School, Howard Hughes Medical Institute, Boston, Massachusetts 02115 USA

The murine, erythroid DNA-binding protein GF-1 (also known as NF-E1, Eryf 1), a 413-amino acid polypeptide with two novel finger domains of the C_x-C_x variety, recognizes a consensus GATA motif present in *cis* elements of the majority of erythroid-expressed genes. We have performed a structure–function analysis of this protein to evaluate its potential as a transcriptional activator and to examine the role of the finger domains in DNA binding. Using a cotransfection assay, we find that GF-1 is a potent transcriptional activator with several activation domains but that this is revealed only in heterologous cells and with reporters containing minimal promoters onto which either a single or multiple GATA-binding sites are placed. The two fingers of GF-1 are functionally distinct and cooperate to achieve specific, stable DNA binding. The amino finger is necessary only for full specificity and stability of binding, whereas the carboxyl finger is required for binding. The role of each finger is more pronounced with some GATA-binding sites than with others, suggesting a diversity of interactions between GF-1 and different target sites. The complex activation and DNA-binding properties of GF-1 are likely to contribute to the ability of this single protein to participate widely in gene expression throughout erythroid development.

Cell-type-specific gene activation is a primary determinant of cellular differentiation and reflects the complex interplay between constitutive and tissue-restricted *trans*-acting regulatory factors (Maniatis et al. 1987). Upon sequence-specific interaction with short DNA motifs in promoter and/or enhancer elements, these proteins increase or repress transcription of their target genes and thereby generate the unique distribution of RNA transcripts characteristic of particular cell types. The specific recognition of cognate DNA-binding sites by such regulators and their interactions with other proteins in a transcriptional complex are directly relevant to an understanding of the mechanisms responsible for differential gene expression accompanying development.

As a eukaryotic system in which to examine cell-type-specific gene activation and regulation, the expression of globin genes in erythroid cells has been the subject of intense scrutiny. Through ontogeny in all vertebrates, programmed shifts in both the sites of erythropoiesis and the types of globin chains expressed occur in a temporal sequence (Wood 1982). *Cis*-acting DNA elements involved in erythroid-specific and devel-

opmentally appropriate gene expression are situated within globin gene promoters, introns, and 3' enhancers, as well as within regions of the DNase I hypersensitive dominant control (or locus-activating) region (Townes et al. 1985; Choi and Engel 1986, 1988; Behringer et al. 1987, 1990; Grosveld et al. 1987; Rutherford and Nienhuis 1987; Enver et al. 1990). Although the spectrum of *trans*-acting factors interacting with these diverse elements is not fully appreciated, the presence of motifs conforming to a consensus sequence [(T/A)GATA(A/G)] (hereafter designated the GATA motif) in the majority of defined *cis* elements of globin and other erythroid-expressed genes is a consistent feature (Evans et al. 1988; Reitman and Felsenfeld 1988; Wall et al. 1988; Martin et al. 1989; Mignotte et al. 1989a,b). In several instances, mutagenesis studies have established the functional importance of these sequences in erythroid gene transcription (Reitman and Felsenfeld 1988; Martin et al. 1989; Mignotte et al. 1989; Plumb et al. 1989). Pursuit of the nuclear protein(s) interacting with this motif has served as an entry into transcriptional regulatory mechanisms operative in erythroid cells.

Nuclear protein DNA-binding experiments have revealed an erythroid factor, variously named GF-1,

[1]Corresponding author.

NF-E1, and Eryf 1, that specifically recognizes the GATA motif (Evans et al. 1988; Wall et al. 1988; Martin et al. 1989). This protein was first observed only in extracts of erythroid cells of all vertebrates examined (mouse, human, and chicken) and at all developmental stages (embryonic, fetal, and adult). Through cDNA cloning, the primary structures of the mouse, human, chicken, and frog (L. Zon and S.H. Orkin, unpubl.) proteins have been determined (Evans and Felsenfeld 1989; Tsai et al. 1989; Trainor et al. 1990; Zon et al. 1990). A striking feature of all these factors is the presence of two domains of the configuration $Cys-X_2-Cys-X_{17}-Cys-X_2-Cys$. Although this structure is reminiscent of the DNA-binding domain of C_x zinc finger proteins, such as those of the steroid receptor gene family (Evans 1988; Evans and Hollenberg 1988), it is distinct. The two finger domains are very similar in amino acid sequence, but differences between them are also highly conserved. Subsequent to the description of the erythroid proteins, it was recognized that *areA* (Kudla et al. 1990) and *nit-2* (Fu and Marzluf 1990), major regulatory factors for nitrogen metabolism in *Aspergillus* and *Neurospora*, respectively, each contain a single finger of the same general configuration, which is very similar to the vertebrate sequences, particularly in their carboxyl finger domains. This observation, taken with the cloning of cDNAs for additional proteins bearing either one or two fingers of similar structure, suggests the existence of a new family of DNA-binding proteins.

Although the erythroid proteins binding to GATA motifs have been implicated in transcriptional activation of globin and other erythroid-expressed genes, direct demonstration that they serve as transcriptional activators is lacking. As we have recently found GF-1 to be expressed in two additional hematopoietic lineages (megakaryocytes and mast cells) that are related to erythroid cells by virtue of descent from more primitive, committed progenitor and stem cells (Martin et al. 1990), we suspect that GF-1 must act in combination with other cellular factors to establish lineage-specific patterns of gene expression. Site-specific mutation in mouse embryonic stem cells followed by generation of chimeric mice has recently provided formal evidence that GF-1 is essential for development of the erythroid lineage (L. Pevny et al., in prep.).

As reported here, we have undertaken a structure—function analysis of murine GF-1 with the dual purpose of (1) assessing the potential of this protein to function as a transcriptional activator and (2) examining the properties of its apparently novel DNA-binding region. Our studies reveal a complexity of activation and DNA-binding domains that suggest specific mechanisms by which a single transcriptional factor may participate in the regulation of diverse target genes during development.

Results

Transcriptional activation by murine GF-1

Activity in transfected nonerythroid cells Transcrip-

tional activation by murine GF-1 was assessed by cotransfection of expressible GF-1 cDNA with various reporter plasmids into nonerythroid cell lines. We first examined *trans*-activation of intact genes or naturally occurring globin promoters fused to reporter sequences. Upon cotransfection of pXM–GF-1 with the mouse α1-globin gene, the human γ-globin promoter linked to human growth hormone (GH) sequences, or the entire human β-globin gene and its 3' enhancer, either with or without the dominant control region, we observed no appreciable *trans*-activation of the test sequences in monkey kidney COS, mouse NIH-3T3, or human HeLa cells (data not shown). These findings are compatible with the requirement for additional tissue-restricted factors for expression from these various constructs, the presence of negatively acting factors in heterologous cells, or the failure of GF-1 to serve as a transcriptional activator.

To avoid problems inherent in studying intact tissue-specific promoters or genes in heterologous cells, we then examined *trans*-activation of artificial promoter constructs in which various binding sites were juxtaposed to sequences derived from the human or rabbit β-globin gene promoters. Other studies have shown that addition of cognate binding sites to the proximal β-globin gene promoter permits *trans*-activation by the B-lymphocyte octamer-binding protein (Oct-2; Muller et al. 1988) or v-*myb* (Weston and Bishop 1989) in heterologous cells. A construct in which six copies of a GF-1-binding site derived from the mouse α1-globin gene promoter were placed at position −44 of the human β-globin promoter was cotransfected into NIH-3T3 cells with pXM–GF-1 or pXM vector containing a deleted form of GF-1 unable to direct wild-type GF-1 protein. As revealed by a 5'-end RNase protection assay (Fig. 1A), coexpression of wild-type GF-1 led to increased transcription of β-globin sequences. The magnitude of *trans*-activation was reduced when binding sites were positioned at −128 of the β-globin promoter (not shown). These findings indicated that GF-1 can activate transcription of a minimal promoter containing cognate binding sites.

To provide a convenient, quantitative assay of *trans*-activation, we generated a series of plasmids with secreted growth hormone as the reporter. Plasmids containing a truncated rabbit β-globin promoter without any additional binding site (TATA–GH), with a single site for the transcription factor Sp1 (Sp1–GH; Kadonaga et al. 1987), or with either one or six GF-1 sites (M1α and M6α, respectively) were cotransfected with pXM vector alone or pXM–GF-1 into COS cells. Coexpression of wild-type GF-1 led to appreciable *trans*-activation of the M1α–GH constructs (5- to 10-fold above basal expression of the reporters) but did not activate promoters with no binding sites (TATA–GH) or with one Sp1 site (Fig. 1B). The extent of *trans*-activation of either M1α–GH or M6α–GH by murine GF-1 was >100-fold in NIH-3T3 cells (Fig. 1C) and in human HepG2 and HeLa cells (not shown). Expression of human GF-1 cDNA also leads to substantial *trans*-activation in NIH-3T3 cells (Fig. 1C).
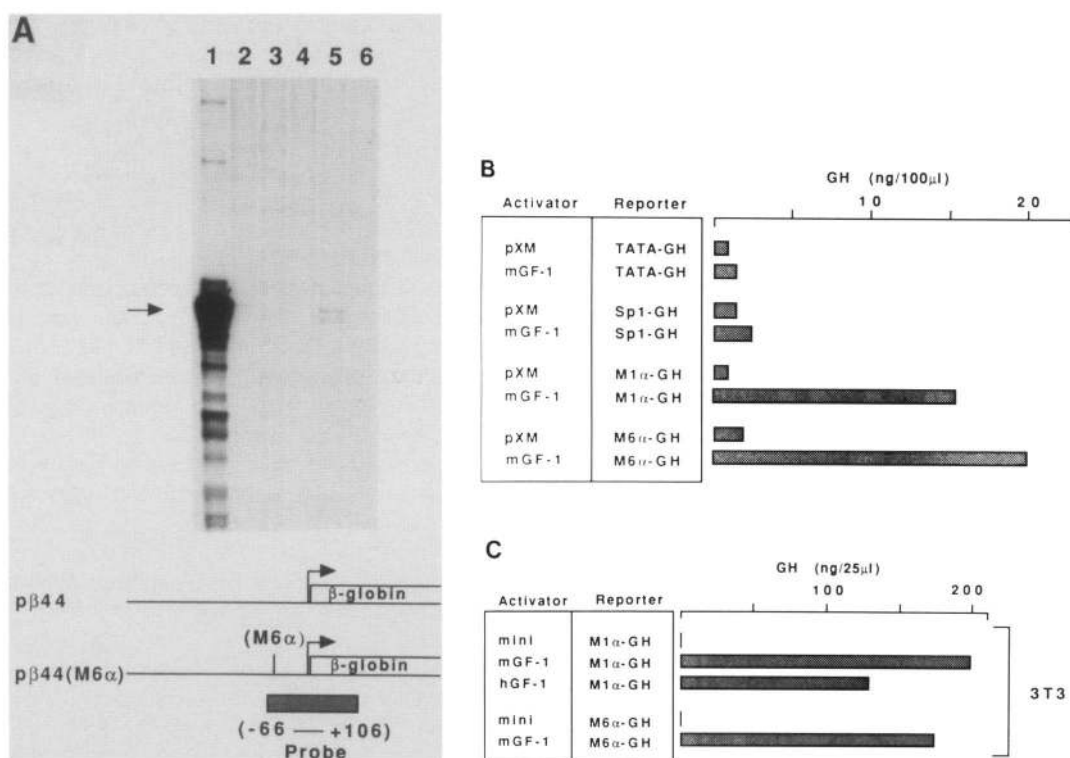
Figure 1. *Trans*-activation by GF-1 in nonerythroid cells. (*A*) *Trans*-activation of the minimal human β-globin promoter containing GF-1-binding sites. The illustrated reporter constructs were cotransfected into NIH-3T3 cells with cDNA expression plasmids containing either wild-type GF-1 or an inactive, deleted version (Δ193; see Fig. 2A). RNA was prepared 72 hr after transfection, and 20 μg of each sample was assayed by a 5′-end RNase protection assay (see Materials and methods). (Lane *1*) 0.5 μg human reticulocyte RNA; (lane *2*) tRNA; (lane *3*) β44 plus wild-type GF-1; (lane *4*) β44-globin plus Δ193; (lane *5*) M6α–β44 plus wild-type GF-1; (lane *6*) M6α–β44 plus Δ193. The protected fragment indicative of correctly initiated β-RNA is indicated by the arrow (*left*). (*B*) *Trans*-activation in monkey kidney COS cells. Cotransfections were performed as described in Materials and methods. Activators pXM and mGF-1 refer to the expression vector containing no cDNA insert and containing wild-type GF-1 cDNA, respectively. Reporters contain either no GF-1-binding sites added (TATA–GH), a single Sp1 site (Sp1–GH), one or six copies of the mouse α1-globin promoter GATA motif (M1α–GH and M6α–GH, respectively). Values are expressed as nanograms of GH/100 μl culture medium. The data shown here are from one of three independent experiments in which transfections were performed in duplicate each time. Duplicates differed by <20%. (*C*) *Trans*-activation in NIH-3T3 cells. Nomenclature is as in *B*, except mini refers to expression of a deleted form of GF-1 containing only the DNA-binding region (see Fig. 2A,B) and hGF-1 to pXM containing full-length human GF-1 cDNA. Values are expressed as nanograms of GH/25 μl culture medium. Cotransfection of TATA–GH and pXM–GF-1 into 3T3 cells leads to apparent transcriptional activation (relative to cotransfection with vector) at 10% that observed with M1α– or M6α–GH (not shown). This effect, dependent on production of GF-1 protein, is thought to result from low-affinity GATA-binding sites in the plasmid backbone.

Cotransfection of a mutant version of GF-1, designated minifactor, in which amino acid residues amino-terminal to position 194 and carboxy-terminal to 308 were removed, failed to activate transcription (Fig. 1C) and was equivalent to the use of expression vector alone (not shown).

From these findings we conclude that expressed murine or human GF-1 can serve as a potent transcriptional activator in a variety of nonerythroid cells, but this is revealed only in the context of minimal promoters containing a cognate binding site. In our studies, we have not observed any reproducible differences in the magnitude of *trans*-activation of reporters containing single versus multiple binding sites. This might reflect either maximal transcription initiation from the minimal promoter with even a single GF-1-binding site in the presence of a saturating level of expressed protein or a property of GF-1 itself.

The large difference in the degree of *trans*-activation of the minimal promoter constructs in COS and 3T3 cells suggests that the milieu into which they are introduced is important. As judged by gel-shift and Western blot analyses, we estimate that the level of GF-1 protein expressed transiently in transfected COS cells is substantially greater than that in transfected 3T3 cells, probably due to SV40 T-antigen-dependent replication of the expression vector in COS cells (not shown). Therefore, the extent of *trans*-activation is not simply related to the level of protein expressed in the nonerythroid cells. In murine erythroleukemia (MEL) cells, the steady-state level of GF-1 is intermediate between that of transfected COS and 3T3 cells (not shown). Upon introduction of the M1α–GH or M6α–GH plasmids into MEL cells by electroporation, we observe no appreciably enhanced expression relative to TATA–GH (not shown).

Localization of multiple activation domains To map regions of murine GF-1 responsible for DNA-binding and transcriptional activation, we constructed deletion mutants of the wild-type cDNA in the pXM expression vector (Fig. 2A) and cotransfected each with M1α–GH or M6α–GH reporters into COS and 3T3 cells. (Nomenclature of the mutants is described in the legend to Fig. 2A.) Cell extracts were assayed for DNA binding by gel-shift assay, using a radiolabeled fragment of the human γ-globin promoter and for *trans*-activation by measurement of secreted growth hormone. Transfection of all mutants shown in Figure 2A into COS cells leads to production of abundant protein that binds a GATA site (Fig. 2B). From the results obtained with the minifactor noted above, we conclude that amino acids amino-terminal to position 193 and carboxy-terminal to 308 are dispensable for DNA binding but not for transcription activation.

Amino-terminal deletions delimit a region within the first 63 amino acids required for *trans*-activation (Δ63) in both COS and 3T3 cells (Fig. 2A). Removal of the extreme amino-terminal segment (Δ22) only partially reduces *trans*-activation. Further amino-terminal deletions generally do not restore activity with the single exception of Δ76, which regains some residual activity. A second activation domain is revealed through study of internal deletion constructs (Δ149–188 and Δ64–193).

Deletion of other residues (Δ80–101) does not affect activity. A third domain participating in *trans*-activation is revealed by carboxy-terminal deletions (Δ308–413, Δ328–382, Δ383–413, Δ349–413). In contrast to the two regions amino-terminal to the finger domains, the carboxy-terminal activation domain appears dispensable when assayed, using the M6α–GH reporter but not the M1α–GH reporter (Fig. 2A, hatched boxes). Although the full significance of this observation is not known, it provides indirect evidence for qualitative, functional differences between the carboxy- and amino-terminal activation domains. A naturally occurring, alternatively processed version of human GF-1 that lacks exon 2 (Martin et al. 1990) is much less active in the *trans*-activation assay than wild-type mouse or human GF-1 (not shown), consistent with the localization of an activation domain to the amino terminus in the human protein.

Differences between domains necessary for activation were also established by assessing the ability of the amino- and carboxy-terminal domains to function independently when fused to a heterologous DNA-binding domain. Segments of GF-1 were joined to the DNA-binding domain of GAL4 (residues 1–147; Ma and Ptashne 1987) in the mammalian expression vector pSG424 (Sadowski and Ptashne 1989) and cotransfected into COS cells with a reporter plasmid containing a minimal promoter with five GAL4-binding sites driving
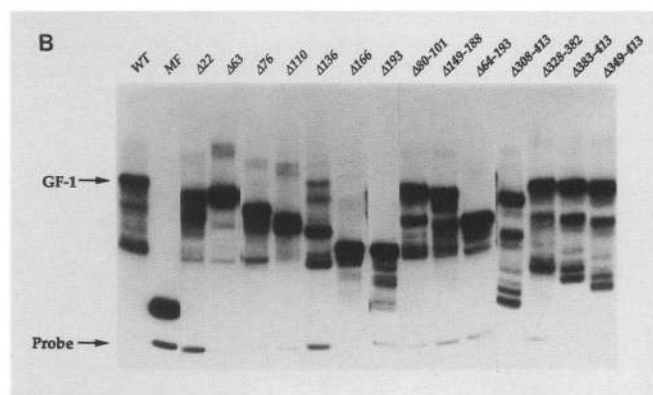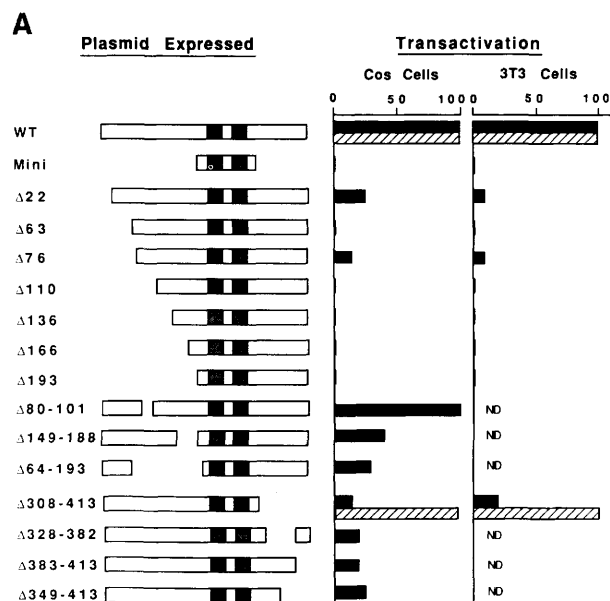

Figure 2. Transcriptional activation and DNA binding by deletion mutants of mouse GF-1. (A) Transcriptional activation. Mutants are designated by the amino acids deleted (far left), which are shown as gaps in the schematic representations. The two finger domains are indicated by shaded squares. Activation in cotransfection assays is shown as the percent of wild type, with the minifactor (mini) taken as baseline. Use of vector alone, expression of an irrelevant cDNA, and minifactor yield equivalent low-level expression from the M1α–GH or M6α–GH reporters (see Fig. 1). *Trans*-activation results with the M1α–GH reporter are represented by solid bars. Differences noted in the results obtained with the M6α–GH reporter are shown with a crosshatched bar. The results reflect the mean of at least three independent experiments for each construct with each transfection performed in duplicate on each occasion. Replicate determinations varied by <20%. Quality of transfections was assessed by examination of gel-shift activity in transfected cells (see B). (ND) Not done. (B) DNA binding of GF-1 mutant proteins. Crude extracts of COS cells transfected with the indicated constructs were tested by gel-shift assay with the human Aγ-globin promoter probe. Differences in mobility of the major DNA–protein complexes are generally correlated with the expected size of the mutant polypeptide. Faster migrating complexes reflect proteolysis in crude extracts (note their absence in the minifactor lane).

expression of chloramphenicol acetyltransferase (CAT) (Gorman et al. 1982). As shown in Figure 3A, addition of the amino-terminal 66 amino acids of GF-1 (pSG412) confers a 29-fold increase in CAT expression. Addition of the entire portion of GF-1 amino-terminal to the finger domains (pSG413) confers less activity (sixfold). Thus, the amino-terminal domain of GF-1 is necessary and sufficient for transcriptional activation. We note that this region displays a net negative charge (Fig. 3B). Addition of the carboxy-terminal domain to GAL4(1–147) (pSG445) does not lead to *trans*-activation. Thus, it is unable to activate independently and, as shown above, is only conditionally necessary for activation.

DNA binding by murine GF-1: cooperative interaction within the finger domains

As noted above, the finger domains together are sufficient for sequence-specific DNA binding (see minifactor, Fig. 2B). To investigate the roles of the individual finger domains in DNA binding, we generated a series of plasmids containing deletions or single amino acid replacements within the two finger DNA-binding domain (Fig. 4A,B). Because the two fingers exhibit a high degree of amino acid sequence similarity, we introduced equivalent mutations into each and compared the ability of expressed mutant GF-1 proteins to bind DNA and to activate transcription in nonerythroid cells. The finger domain mutations presented here perturb the DNA-binding properties of GF-1 in different ways and serve to delineate distinct functions of the two fingers.

Mutation of the amino finger reduces the stability of DNA binding To minimize effects on protein folding, the individual finger domains were deleted at positions corresponding to the intron/exon boundaries of the GF-1 gene. Deletion of the amino finger (Δ200–248) or a single substitution C207P (see Fig. 4A,B) does not prevent DNA binding as assessed by gel-shift assay (Fig. 4C). In contrast, deletion of the carboxyl finger (Δ249–290) or a single change C261P abolished binding (Fig. 4C). Mutants that bind DNA activate transcription of the Mlα–GH reporter in COS cells, whereas nonbinding mutants do not (Fig. 4A). Failure of Δ249–290 and C261P mutants to bind DNA is not due to their instability or inadequate accumulation, as shown by Western blot analysis (Fig. 4D). Although it has been difficult to establish a requirement for a metal ion in the binding of chicken Eryf 1 to DNA (Evans and Felsenfeld 1989), C261P provides strong support for a critical role of the cysteine residues in finger structure.

Although the Δ200–248 mutant protein recognizes DNA containing a GATA-binding site in a gel-shift assay, closer inspection by two different approaches reveals significant alterations in its binding properties. Previously we showed that a single GF-1 molecule simultaneously binds and contacts the pair of closely spaced GATA motifs in the human Aγ-globin promoter (Martin et al. 1989), as assessed by both DNase I footprinting and methylation interference assays. DNase I
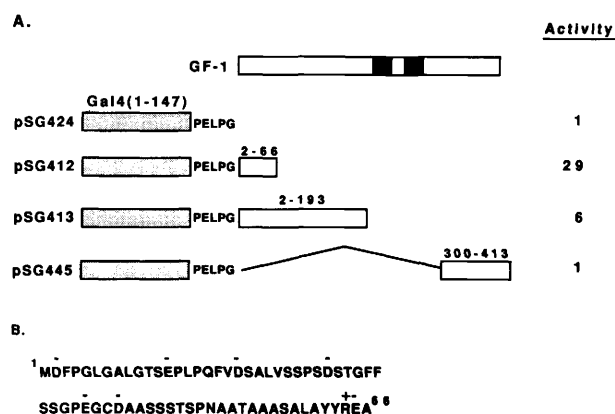

Figure 3. Transfer of the amino-terminal activation domain to a heterologous DNA-binding site. (A) Transcriptional activation by portions of GF-1 fused to the DNA-binding domain of GAL4. Numbers above the open boxes denote amino acids derived from GF-1 that are included in each fusion construct. Values at *right* reflect CAT activity of cells harvested 60 hr after cotransfection of the expression and reporter (G5BCAT) plasmids (see Methods). The results reflect the mean of three independent experiments. (B) Primary sequence of the amino-terminal 66 amino acids of murine GF-1 included in the construct pSG6412. Charged residues are indicated. The net charge is −6; homologs of GF-1 have similarly charged amino termini (see Discussion).

footprinting shows no apparent difference between wild-type GF-1 and Δ200–248 proteins on the bipartite γ-globin promoter site (Fig. 5A). As the conditions of the DNase I footprinting assay favor saturation of binding sites, which might allow two molecules of Δ200–248 to bind the γ-promoter probe, we used the methylation interference assay to examine bases contacted when only one molecule was bound. In the typical GF-1 gel-shift complex, contacted bases are evident in both GATA motifs when wild-type or minifactor GF-1 binds the bipartite γ-globin promoter site (Fig. 5B). However, Δ200–248 contacts only the distal motif (Fig. 5B). In addition, the interference of DNA binding by methylation is relaxed with Δ200–248. Normally, methylation of bases prevents recognition by the factor (i.e., contact) and results in the absence of specific bands. In Δ200–248, however, the stronger band intensity at these sites indicates that methylation is less likely to prevent protein binding. Thus, as Δ200–248 binds altered (i.e., methylated) sites more readily than wild type, it appears to have less stringent sequence specificity.

A more striking abnormality of DNA binding by Δ200–248 reflects decreased stability of binding. To evaluate the rate at which protein–DNA complexes dissociate, we performed "off-time" experiments. After initial formation of protein–DNA complexes, a 10^3-fold excess of unlabeled probe DNA was added to the binding mixture, and aliquots were removed at selected time intervals and immediately loaded and electrophoresed into a native gel. As shown in Figure 6A, wild-type GF-1 dissociates slowly from the γ-globin promoter test probe. A significant fraction of DNA is still bound after 25 min. In contrast, the Δ200–248 complex dissociates almost
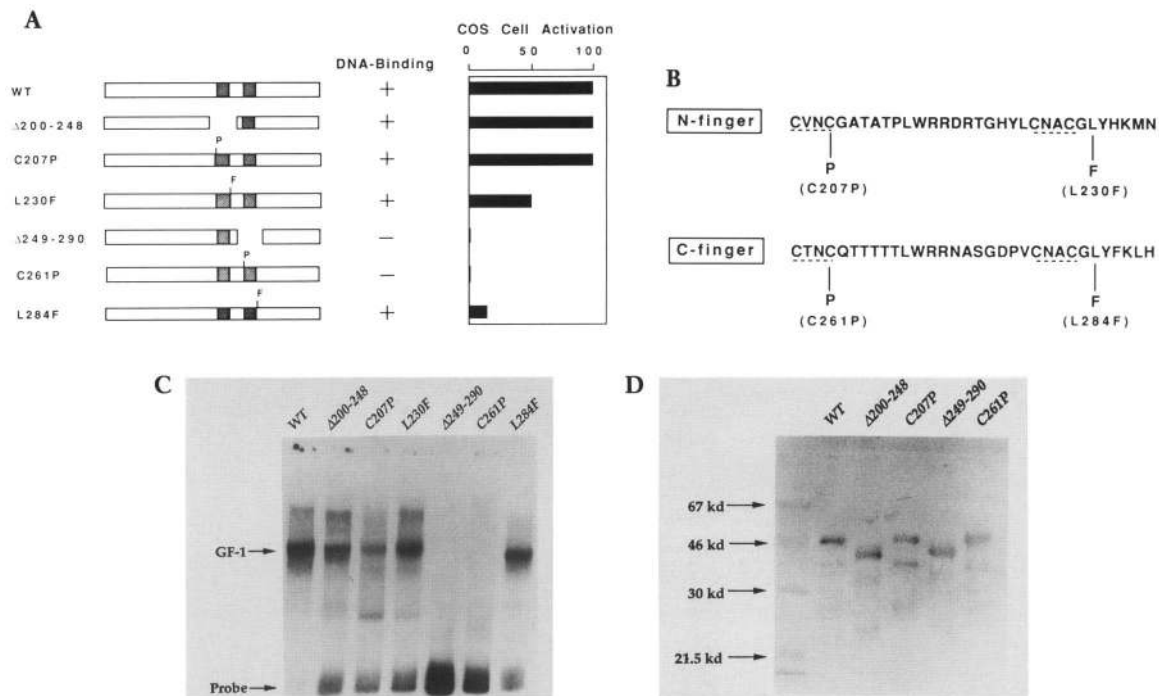
Figure 4. Mutations of the DNA-binding domain of GF-1. (*A*) DNA-binding and transcriptional activation by GF-1 finger domain mutants. Constructs designated at *left* and schematically represented in the *middle* expressed either deleted forms of GF-1 or polypeptides with single amino acid substitutions. The single-residue replacements are designated by the wild-type residue, the residue number, and the substituted amino acid (see *B*). Plasmids were cotransfected in COS cells with the M1α–GH reporter and tested for DNA binding by gel-shift assay using the Aγ-globin promoter fragment as probe and for *trans*-activation (expressed as in Fig. 2A). DNA binding is represented as + or − without reference to differences in stability or affinity of binding, as discussed in the text. (*B*) Single amino acid substitution mutations in the finger domains. The cysteine pairs of the putative fingers are underlined by dashes. (*C*) DNA binding by finger domain mutants. Extracts of COS cells transfected with the indicated expression plasmids were subjected to gel-shift assay with the Aγ-globin promoter fragment as probe. As protein concentrations of the COS extracts were not standardized for this assay, qualitative and quantitative differences in DNA binding established by further analyses are not accurately reflected here. Only the failure of Δ249–290 and C261P to bind should be noted. (*D*) Western blot assay of finger domain mutants. Protein extract (0.5 μg) of transfected COS cells was loaded in each lane and subjected to Western blot analysis with antisera to mouse GF-1. Size markers are shown at *left*. Wild-type GF-1 migrates as a ~50-kD polypeptide. Binding (Δ200–248, C207P) and nonbinding (Δ249–290, C261P) mutants are expressed at levels similar to wild type.

completely within 1 min, even though initial complex formation is robust (Fig. 6B). As the interaction of wild-type GF-1 with the two GATA motifs in the γ-globin promoter fragment is complex, we also examined the dissociation of wild-type and Δ200–248 proteins following initial binding to the mouse α1-promoter fragment, which contains only a single GATA motif. Although interaction of Δ200–248 with this site is more stable than with the γ-globin promoter fragment, dissociation is still at least fivefold faster compared with wild type (Fig. 6C,D). Thus, binding of Δ200–248 is unstable but much more so upon interaction with the bipartite γ-globin promoter site. C207P mimics Δ200–248 in the dissociation assay (not shown) and, again, points to a critical role of the cysteine residue in finger function. Our findings with the Δ200–248 and C207P mutants suggest that the amino finger is necessary for full specificity and stability of binding by GF-1 but that its role is more important at a double site than a single one.

Mutation of the carboxyl finger affects high-affinity DNA binding Deletion of the carboxyl finger in

Δ249–290 and the single substitution C261P demonstrate a requirement for this finger in sequence-specific DNA binding. Mutant L284F, bearing a substitution at a conserved residue in the carboxyl finger, exhibits a normal dissociation rate with target sequences (not shown) but binds with greatly reduced affinity (Fig. 7). Binding affinity was titrated in a gel-shift assay in which extracts of transfected COS cells containing equivalent amounts of GF-1 polypeptide (wild type or mutant) were incubated with either the Aγ- or α1-globin probes. With either target site, binding affinity was reduced by >10-fold. The corresponding mutation in the amino finger (mutant L230F) produces a different result. Although the affinity of DNA binding by L230F is within twofold of wild type, the stability of binding is reduced, though not to the extent evident with the finger deletion (Δ200–248; not shown). Thus, mutations that alter (or remove) the amino finger (Δ200–248, C207P, and L234F) decrease the stability of DNA binding, whereas mutations in the carboxyl finger either prevent (Δ249–290 or C261P) or reduce (L284F) the affinity of binding.
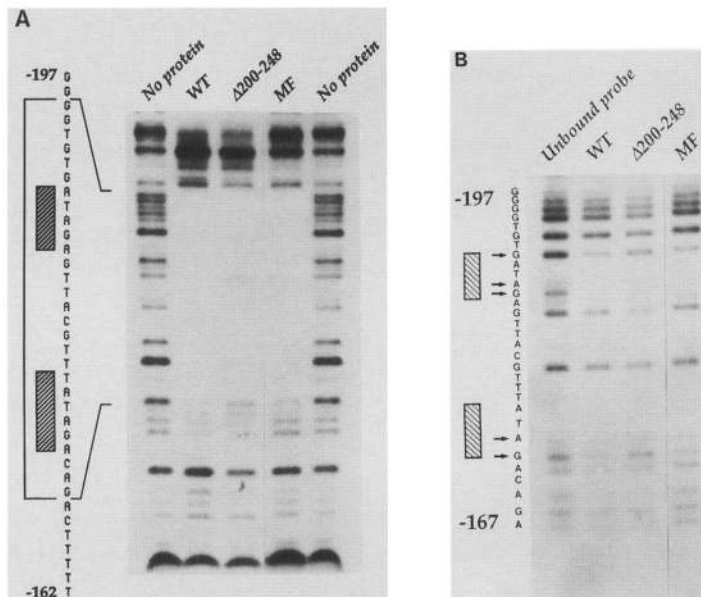
Trans-activation by the finger domain mutants in

Figure 5. DNase I footprinting and methylation interference assays of the Δ200–248 mutant. (A) DNase I protection assay of extracts of COS cells transfected with GF-1 expression plasmids. The sequence of the region of the human Aγ-globin promoter containing the bipartite GATA site is shown at *left*. The GATA motifs are indicated by hatched boxes. Wild-type (WT), minifactor (MF), and Δ200–248 yield indistinguishable protection patterns. As anticipated from its failure to bind in a gel-shift assay, Δ249–290 does not protect the Aγ-globin promoter fragment (not shown). The conditions of this assay are designed to give complete binding of probe by GF-1; under these conditions, Δ200–248, but not wild-type GF-1, tends to form a slower migrating complex that may represent two molecules binding. (B) Methylation interference assay of wild-type and mutant GF-1 bound to the Aγ-globin promoter probe. Arrows at *left* of the sequence denote bases normally contacted by GF-1. In the Δ200–248 lane, note that the intensity of bands at contact sites is intermediate between unbound and wild type in the distal GATA motif and the absence of interference at the proximal motif.

COS cells is not strictly correlated with DNA binding. Although the binding stability of Δ200–248 and C207P is reduced, *trans*-activation is normal (Fig. 4A). We suspect that this may reflect saturation of target-binding sites on the reporter in the face of the very high levels of protein expressed in COS cells. In accord with this view, we find a modest (50–75%) reduction in activation in cotransfected 3T3 cells (not shown). *Trans*-activation in COS cells by L230F is affected only minimally. *Trans*-activation by L284F, however, is reduced substantially (Fig. 4A). Thus, we infer that DNA-binding affinity, rather than the stability of the DNA-binding complex, is the primary determinant of activation as measured by cotransfection assays.

To summarize our findings from alteration of the GF-1 DNA-binding domain, (1) the two fingers are functionally distinct; (2) the carboxyl finger is required for binding; (3) the amino finger is necessary only for full specificity and stability of binding; (4) the role of each finger is more pronounced with some binding sites than with others; (5) effects on transcriptional activation are not strictly correlated with effects on DNA binding.

Discussion

The erythroid DNA-binding protein as a transcriptional activator

Several lines of evidence implicate GF-1/NF-E1/Eryf 1 as an important regulator of gene transcription in erythroid cells. These include its tissue restriction, the presence of GATA-binding sites in *cis*-acting DNA elements of virtually all characterized erythroid-expressed genes, and analyses that indicate a function for these sites in promoters and enhancers (Evans et al. 1988; Reitman and Felsenfeld 1988; Wall et al. 1988; Evans and Felsenfeld 1989; Tsai et al. 1989). Our present studies establish that murine GF-1 is a potent transcriptional activator in cotransfection experiments with artificial promoters in a heterologous cell environment. *Trans*-activation is dependent on the presence of a GATA-binding site but is not appreciably influenced by the relative number of sites in a minimal promoter. We suspect that this may reflect saturation of transcription initiation from these artificial promoter constructs.

Mutational analysis defines several regions of murine GF-1 that participate in transcriptional activation. An acidic, serine-rich amino-terminal domain of 66 residues is necessary for appreciable activation and confers activation upon fusion to a heterologous DNA-binding domain (Fig. 3A). Activation by the amino-terminal 66 residues is perhaps not surprising in view of previous characterization of acidic activation domains (Ma and Ptashne 1987; Ptashne 1988). Although we have not directly shown that the overall negative charge of this region accounts for its activation properties, conservation of net negative charge in the amino terminus of mouse, human, and chicken erythroid factors, as well as in two chicken proteins with highly related DNA-binding do-
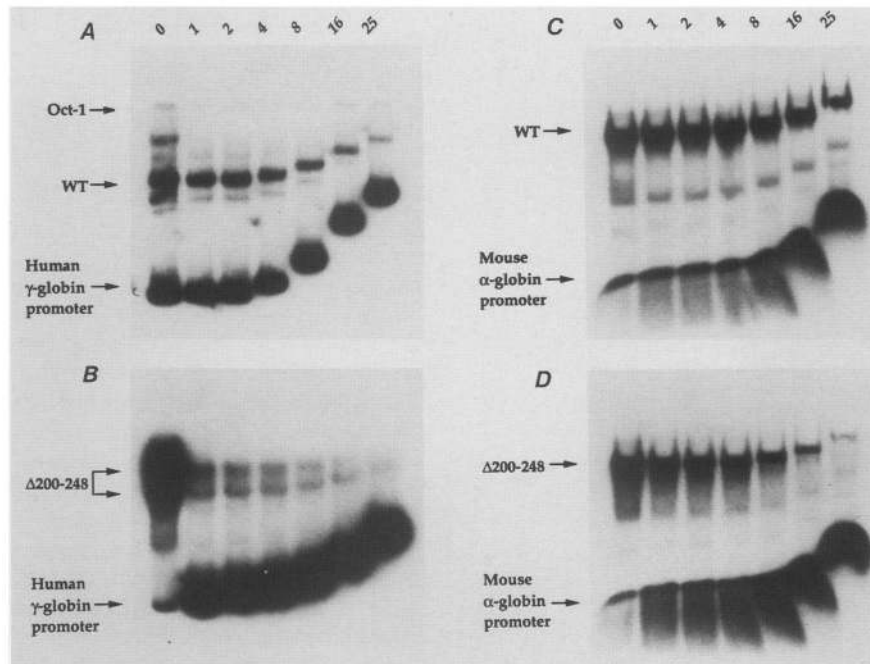
Figure 6. Unstable DNA binding by Δ200–248. Off-time assay of wild-type and mutant GF-1 binding to different GATA elements. Extracts of COS cells transfected with wild-type (WT; *A, C*) and Δ200–248 (*B, D*) expression plasmids were analyzed by gel-shift assay with the Aγ-globin promoter (*A, B*) or the mouse α1-globin promoter (*C, D*) probes. Gel-shift reactions were assembled. At time zero (lane *0*), a 10³-fold excess of unlabeled, double-stranded competitor oligonucleotide was added to each reaction mix. Aliquots were withdrawn and loaded immediately at 1, 2, 4, 16, and 25 min thereafter. The Aγ-globin fragment contains a canonical octanucleotide between the GATA motifs; note the instability of Oct-1 binding.

mains (NF-E1b and NF-E1c; Yamamoto et al. 1990), strongly suggests that this is likely to be the case.

Mutational analysis also reveals other regions required of the murine GF-1 molecule for full transcriptional activation. Taken with the findings of others (Mitchell and Tjian 1989), it appears that the existence of multiple domains within transcriptional activators is a common property of such molecules. Neither the central nor carboxyl domains of GF-1, delineated by mutants Δ149–188 and Δ308–413, are appreciably negative in charge, nor do they resemble other activation motifs (Mitchell and Tjian 1989). The different behavior of the Δ308–413 mutant upon assay with single versus multiple binding sites is particularly striking and suggests that the requirement for the carboxy-terminal activation domain can be abrogated by interactions between adjacent GF-1 molecules. The finding that the amino-terminal domain can serve as an activator when isolated from other portions of the GF-1 polypeptide, yet requires other regions for full activity in the context of the native protein, implies that a particular conformation may be functionally important to the establishment of appropriate protein–protein interactions in a transcriptional complex.

Although we have delineated domains of GF-1 that participate in transcriptional activation in heterologous cells, the function of these regions in the context of an erythroid cell environment is not addressed in our studies. That differences exist in transcriptional activation by this factor in various intracellular environments is suggested by several independent observations. First, we observe little or no *trans*-activation of minimal promoter constructs upon introduction into MEL cells, in which the endogenous GF-1 level is intermediate between that achieved transiently in 3T3 and COS cells. This finding agrees with the observation that simple grafting of GATA-binding sites onto truncated human γ- or mouse α1-globin promoters does not confer high-level, erythroid expression of linked reporters (Plumb et al. 1989; D.I.K. Martin and S.H. Orkin, unpubl.). Promoter architecture may be critical to the function of GF-1 in an erythroid cell environment but perhaps less so in heterologous cells. With the minimal promoter constructs that we have tested, interactions of nuclear proteins with GF-1 bound to DNA may be largely nonproductive within the context of MEL cells. Potential positive interaction of GF-1 (NF-E1) with factors recognizing other *cis* elements in the human β- and ζ-globin promoters, such as the CACCC motif, has been proposed by others (deBoer et al. 1988; Antoniou and Grosveld 1990; Frampton et al. 1990; Watt et al. 1990). Second, the appreciably lower level of *trans*-activation by GF-1 in COS cells, as compared with 3T3 cells, illustrates directly the influence of cellular environment. As cell-specific differences in the activity of a transcription factor have been described previously (Tora et al. 1989), the phenomenon may be quite common. Third, study of the naturally occurring T-C mutation at position −175
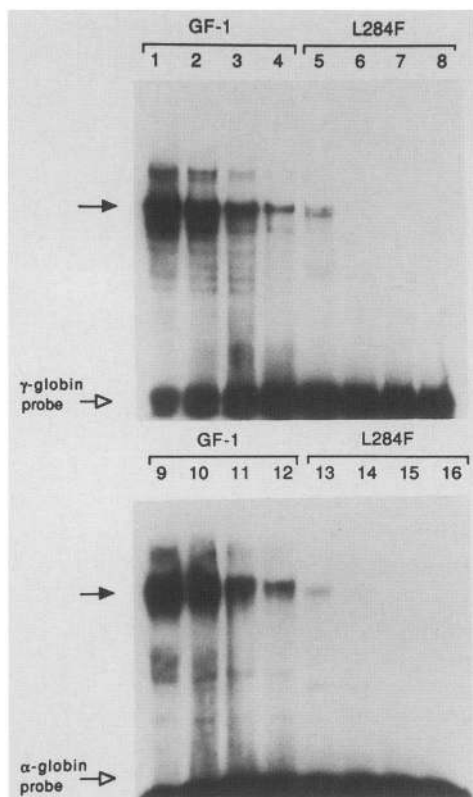
Figure 7. Reduced DNA-binding affinity of the L284F mutant. DNA-binding affinity of wild-type GF-1 and L284F assayed with different GATA elements as targets. Extracts of COS cells transfected with wild-type (lanes *1–4* and *9–12*) or L284F (lanes *5–8* and *13–16*) expression plasmids were analyzed by a quantitative gel-shift assay, using either the bipartite Aγ-globin promoter probe (lanes *1–8*) or mouse α1-globin probe (lanes *9–16*). Western blot analysis was employed first to establish that extracts of COS cells contained equivalent levels of wild-type and L284F polypeptides. COS extracts were then diluted serially in binding buffer. In each assay, the ratio of protein to poly[d(I-C)] was held constant. Lanes contain 100 ng (*1, 5, 9, 13*), 40 ng (*2, 6, 10, 14*), 20 ng (*3, 7, 11, 15*), or 10 ng (*4, 8, 12, 16*) of protein. The complexes arising from binding of GF-1 are indicated by solid arrows.

of the human γ-globin promoter that causes a fivefold increase in transcriptional activity in erythroid cells (Martin et al. 1989; Nicolis et al. 1989) has not revealed a corresponding enhancement in activation when placed in a minimal promoter and cotransfected with GF-1 into heterologous cells (not shown). Finally, the effect of promoter architecture is apparent from preliminary analysis of the erythroid-specific promoter of the erythropoietin receptor. This promoter, characterized by a single GATA-binding site near the transcriptional start sites (Youssoufian et al. 1990), functions upon transient transfection into MEL cells and is also susceptible to *trans*-activation upon transfection with GF-1–pXM in nonerythroid cells (S.H. Orkin and L. Zon, unpubl.). Ultimately, a complete resolution of the contribution of specific domains defined in heterologous cells to activation in an erythroid environment will necessitate development of assays for *trans*-activation by mutated versions of GF-1 in the context of an erythroid cell environment.

Novel aspects of the DNA-binding domain of the erythroid factor

Murine and chicken GF-1/Eryf 1 proteins constitute the founding members of a new family of finger DNA-binding proteins, which recognize a GATA motif. Although their $C-X_2-C-X_{17}-C-X_2-C$ structures superficially resemble zinc fingers of the C_x variety, the conservation of primary sequence within the putative fingers in the GATA-binding proteins is extraordinary (Fig. 8). Members of the family now include three single-finger proteins [yeast Gln3 (P.L. Minehart and B. Magasanik, in prep.)], *Aspergillus areA* (Kudla et al. 1990), and *Neurospora nit-2* (Fu and Marzluf 1990), as well as several two-finger proteins: ELT-1 [*Caenorhabditis elegans* (T. Blumenthal, pers. comm.); GF-1/Eryf 1 [murine, human, chicken, and frog (L. Zon and S.H. Orkin, unpubl.)], and NF-E1b and NF-E1c (Yamamoto et al. 1990). Although members of this family all appear to recognize the core GATA motif, amino acid differences in the fingers may serve to discriminate between core GATA motifs embedded in diverse surrounding sequences. Evidence in favor of this hypothesis is derived from two sources. First, chicken Eryf 1 does not recognize all sequences conforming to the basic consensus site with equal efficiency (Evans et al. 1988). Second, a more ubiquitously distributed protein described elsewhere binds a GATA *cis* element in the preproendothelin-1 gene promoter but does not efficiently recognize the human Aγ-globin bipartite site (Wilson et al. 1990). The contribution of individual amino acid residues to subtle sequence discrimination may be apparent upon further comparison of members of the GATA-binding family and site-specific mutagenesis.

The conservation of amino acid differences between the two fingers of the vertebrate GATA-binding proteins, and the closer resemblance of the carboxyl finger of the two finger proteins to the single fingers of yeast and fungal factors are consistent with the hypothesis that the two fingers of the higher eukaryotic proteins may have evolved to serve different functions. Our mutagenesis of murine GF-1 demonstrates this to be the case. The two finger domains are not functionally equivalent but, instead, cooperate in the intact molecule to achieve stable, sequence-specific interaction with DNA. The carboxyl finger is necessary and sufficient for binding to a GATA motif. The amino finger appears to increase sequence specificity and stability of binding; its role is greater when GF-1 binds two GATA motifs simultaneously, as in the bipartite human γ-globin site, suggesting that it may also interact directly with DNA. As the carboxyl finger in mouse GF-1 is essential for DNA binding and also resembles that present in yeast and fungal proteins, it is probable that the two-fingered DNA-binding domain arose by its duplication. Thereafter, the fingers diverged and acquired functions that act cooperatively to achieve stable, high-affinity DNA

```
GLN-3 (304)   I Q C F N C K T F K T P L W R R S P E G N T L C N A C G L F Q K L H G T M R P L S L K S D V I K K R I S
areA  (514)   T T C T N C F T Q T T P L W R R N P E G Q P L C N A C G L F L K L H G V V R P L S L K T D V I K K R N R
NIT-2         T T C T N C F T Q T T P L W R R N P D G Q P L C N A C G L F L K L H G V V R P L S L K T D V I K K R N R
ELT-1 (N)     R E C V N C G V H N T P L W R R D G S G N Y L C N A C G L Y F K W N H H A R P L V K P K K R Q Q N A Q K
mGF-1 (202)   R E C V N C G A T A T P L W R R D R T G H Y L C N A C G L Y H K M N G Q N R P L I R P K K R M I V S K R
hGF-1 (202)   R E C V N C G A T A T P L W R R D R T G H Y L C N A C G L Y H K M N G Q N R P L I R P K K R L I V S K R
NF-E1a (108)  R E C V N C G A T A T P L W R R D G T G H Y L C N A C G L Y H R L N G Q N R P L I R P K K R L L V S K R
NF-E1b (278)  R E C V N C G A T A T P L W R R D G T G H Y L C N A C G L Y H K M N G G N R P L I K P K K R L S A A R R
NF-E1c (262)  R E C V N C G A T S T P L W R R D G T G H Y L C N A C G L Y H K M N G Q N R P L I K P K K R L S A A R R


GLN-3 (304)   I Q C F N C K T F K T P L W R R S P E G N T L C N A C G L F Q K L H G T M R P L S L K S D V I K K R I S
areA  (514)   T T C T N C F T Q T T P L W R R N P E G Q P L C N A C G L F L K L H G V V R P L S L K T D V I K K R N R
NIT-2         T T C T N C F T Q T T P L W R R N P D G Q P L C N A C G L F L K L H G V V R P L S L K T D V I K K R N R
ELT-1 (C)     T E C V N C R T N T T T L W R R N G E G H P V C N A C G L Y F K L H K V E R P I T M K K D G I Q T R N R
mGF-1 (256)   T Q C T N C Q T T T T T L W R R N A S G D P V C N A C G L Y F K L H Q V N R P L T M R K D G I Q T R N R
hGF-1 (256)   T Q C T N C Q T T T T T L W R R N A S G D P V C N A C G L Y Y K L H Q V N R P L T M R K D G I Q T R N R
NF-E1a (162)  T V C S N C Q T S T T T L W R R S P M G D P V C N A C G L Y Y K L H Q V N R P L T M R K D G I Q T R N R
NF-E1b (333)  T C C A N C Q T T T T T L W R R N A N G D P V C N A C G L Y Y K L H N V N R P L T M K K E G I Q T R N R
NF-E1c (315)  T S C A N C Q T T T T T L W R R N A N G D P V C N A C G L Y Y K L H N I N R P L T M K K E G I Q T R N R
```

Figure 8. Primary sequence conservation of the finger domains of GATA-binding proteins. The finger domains of the single finger members, GLN3, *areA*, and NIT-2, are compared with the amino (*top*) and carboxyl (*bottom*) fingers of the two finger proteins. The sequence of ELT-1 was kindly provided by T. Blumenthal. Regions of conserved residues are enclosed in boxes. Numbers in parentheses indicate the amino acid residue number at the left of each sequence that is shown.

binding. Duplication of the finger domain may permit recognition of a wider variety of binding sites, as interaction with the bipartite γ-globin promoter site seems to require both.

Although the DNA-binding domain of GF-1 does not closely resemble in detail that of other described proteins, the cooperative behavior of the finger domains is reminiscent of the interaction of the two, nonhomologous zinc fingers of the steroid receptor superfamily members (Evans and Hollenberg 1988). As mutations in either of the receptor fingers generally abolish DNA binding, the functions of the individual fingers are not so readily assigned. Nonetheless, it has been proposed that the amino finger is responsible for sequence-specific interactions and the carboxyl finger is responsible for nonspecific DNA interactions and dimerization. The separation of functions of the fingers in GF-1 is analogous but more evident.

With the mutants examined so far, we have not revealed uncoupling of DNA binding and activation by mutations in the DNA-binding domain, as has been reported for the glucocorticoid receptor (Schena et al. 1989), the yeast activator HAP1 (Kim and Guarente 1989), and the muscle regulator MyoD (Davis et al. 1990), where specific mutations within the regions ascribed DNA-binding function influence transcriptional activation without demonstrably perturbing DNA binding. Although we note that DNA binding and activation by specific mutations are not precisely correlated (Fig. 4A), pronounced effects of these mutations on binding properties prevent assignment of activation functions to the DNA-binding region of GF-1. In part, our reluctance to relate mutations in the DNA-binding domain to activation changes reflects the stringent criteria that we have applied to evaluate DNA binding.

The complex interactions of the finger domains with target DNAs, as exemplified by the amino-finger mutants (Δ200–248 and C207P), suggest specific mechanisms by which GF-1 may participate in the regulation of diverse genes during cell development. Although GATA-binding sites are most commonly present as a single copy in promoters or enhancers, several instances of bipartite (or double-copy) sites are known. The functional consequences of this diversity of binding sites are speculative, as the transcriptional activation assay in heterologous cells does not accurately reflect the subtleties of DNA binding. We envision at least two ways in

which our findings relate to the function of the protein. First, activation by GF-1 in vivo may be dependent on the type of site to which it is bound (Kim and Guarente 1989; Martin et al. 1989). Thus, subtle differences in binding of diverse target sites by GF-1 may influence its activity within a transcriptional complex. Second, it may be inferred that the occupancy of different types of GATA-binding sites will vary with protein concentration. In this regard, it is of interest that GF-1 mRNA and protein levels change during mouse development (E. Whitelaw et al., in prep.) and during normal erythropoiesis (L. Zon and S. Orkin, unpubl.). Concentration-dependent, differential gene activation, similar to that established for *bicoid* in *Drosophila* embryos (Struhl et al. 1990), may contribute to GF-1 action during erythroid development.

Biological advantages of a complex transcription factor

The simple GATA consensus sequence to which the transcription factor GF-1 binds is present in the majority of *cis*-regulatory elements defined as critical for erythroid-specific gene expression. In addition, by virtue of its restricted expression in megakaryocytes and bone-marrow-derived mast cells (Martin et al. 1990), it seems probable that GF-1 also regulates numerous genes in these hematopoietic lineages. These observations pose a paradox: How can a single factor participate in complex, developmentally programmed regulation of such diverse genes? Several mechanisms may pertain. For one, promoter and enhancer architecture, as defined by the relationship of a core binding site to binding sites for other nuclear proteins, is likely to dictate combinatorial patterns that specify gene activation. Evidence is already available that GF-1 action can be influenced by such associations (Antoniou and Grosveld 1990; Frampton et al. 1990). Second, transcription factors may be altered by alternative RNA processing or by post-transcriptional modifications so as to change DNA-binding or activation properties. Although further studies are required, preliminary data have failed to reveal alternative RNA processing or substantial modification of GF-1 protein during erythropoiesis in the mouse (Whitelaw et al., in prep). Our findings concerning the activation and DNA-binding properties of murine GF-1 reported here suggest additional means by which this paradox may be resolved. First, the multiple and qualitatively different ac-

tivation domains defined in heterologous cells may provide an array of surfaces with which other nuclear factors in different cell types and at different stages of development might interact. Second, the interactive nature of the finger domains and the influence of target site structure and sequence on DNA binding are features that may contribute to complex regulation mediated through a single protein. Viewed from this perspective, the presence of multiple activation domains and a finely tuned DNA-binding unit affords a versatility to GF-1 that may endow it with substantial functional diversity. In the final resolution, the number of cell-restricted transcription factors required to achieve and establish lineage-specific gene expression may be fewer than anticipated thus far.

Materials and methods

Trans-*activation reporter constructs and cotransfection assay*

TATA–GH was generated by insertion of an oligonucleotide corresponding to the rabbit β-globin sequence (Muller et al. 1988) between the HindIII and BamHI sites of pOGH (Nichols Institute). M1α–GH and M6α–GH contained one and six copies, respectively, of the GATA-binding site of the mouse α1-globin promoter (GGGCAACTGATAAGGATTCC). Sp1–GH contained one copy of the sequence utilized previously by Muller et al. (1988). The human β-globin gene reporter truncated to position −44 (pβ44; Weston and Bishop 1989) was used to derive pβ44(M6α). For RNase protection assay of cells transfected with the pβ44 and pβ44(M6α) plasmids, total cell RNA was prepared by standard methods and treated with RNase-free DNase to remove residual plasmid DNA. The RNase protection probe was synthesized by T7 polymerase transcription of XbaI-linearized Bluescript vector into which a PCR-generated fragment spanning −66 (BamHI) to +106 (HindIII) was cloned. Hybridization and RNase treatment were as described (Melton et al. 1984).

COS cell transfections were performed by the DEAE–dextran method (Ausubel et al. 1987). Line 3T3, HepG2, and HeLa cells were transfected by the method of Chen and Okayama (1987). MEL cells were electroporated in HEPES-buffered saline at 960 μF (280 V), using a Bio-Rad Gene Pulser apparatus. For COS cell transfection, 1 μg reporter and 2.5 μg cDNA expression plasmid were added per 60 mm culture well. Cells were generally assayed for 60 hr, following removal of the DEAE–dextran and a 10% DMSO shock. For 3T3, HepG2, and HeLa transfections, precipitates containing 1.5 μg reporter and 10 μg cDNA expression plasmid DNAs were applied overnight. Assays were performed 60–72 hr after removal of the precipitates and a medium change.

To validate the cotransfection assay, we determined that activation is linearly related to activator at limiting input in 3T3 transfections and that the RNA transcribed from the M1α–GH construct in cotransfected cells is appropriately initiated (not shown). All transfections were performed in duplicate and in at least three independent experiments for each construct. Duplicate samples varied by <20%. All plasmids for transfections were banded twice in cesium gradients. For transfections with the β-globin reporter plasmids and with the GAL4 fusion constructs, TK-GH plasmid DNA was added to monitor the relative transfection efficiencies. Growth hormone determinations were made by using the Allegro radioimmunoassay kit. In general, 100 μl medium was directly assayed. Where necessary, samples were diluted to fall within a standard curve.

Site-directed mutagenesis

Mouse GF-1 was cloned into the XhoI site of BlueScript KS+ (Stratagene). Single-stranded template was generated in the Escherichia coli dut⁻, ung⁻ strain, CJ236 (Vieria and Messing 1987). Mutagenesis and selection were performed (Kunkel et al. 1987) by using 20- to 35-mer oligonucleotides prepared on an Applied Biosystems 380B DNA synthesizer. After sequencing to verify the desired mutations, inserts were excised as a KpnI–NotI fragment and recloned into a pXM plasmid altered by insertion of a linker containing KpnI and NotI sites between the resident PstI and EcoRI cloning sites. All amino-terminal deletions retain the first two amino acids (M-D) and continue with the amino acid following the number designated.

For finger deletions, the positions of the introns in the cellular gene (determined by S.-F. Tsai and kindly made available for this work) were used to guide mutagenesis.

GAL4 fusions were created by inserting PCR-generated fragments of GF-1 between the BamHI and KpnI restriction sites of the GAL4(1–147) plasmid pSG424 (Sadowski and Ptashne 1989). PCR fragments were generated by standard methods with primers encoding the amino acids at the ends of each fragment. The primers also included either BamHI or KpnI restriction sites to facilitate cloning. Amino acid sequences contained in the fusion proteins are as listed in Figure 3. The reporter for the GAL4 fusions was G5BCAT, a construct containing five GAL4-binding sites and a TATA box adjacent to the CAT gene (kindly provided by D. Last and M. Ptashne).

CAT assay

Assays for CAT activity (Gorman et al. 1982) were performed with the Promega CAT assay kit. Acetylated chloramphenicol was extracted into mixed xylenes and counted by liquid scintillation. In the GAL4 fusion experiments, pTK–GH (Nichols Institute) was cotransfected as a control to evaluate reproducibility of transfections.

DNA-binding assays

Cell extracts were prepared as described (Tsai et al. 1989). After washing once with phosphate-buffered saline (PBS), cells were scraped from plates into PBS and pelleted at 1500 rpm. The cell pellet was resuspended in 25 μl lysis buffer [10 mM HEPES (pH 7.9), 140 mM NaCl, 1.5 mM MgCl₂, and 0.5% NP-40] containing protease inhibitors (PMSF, aprotinin, leupeptin, and pepstatin) and centrifuged at 2000 rpm for 10 min at 4°C. The supernatant was removed and stored at −80°C and thereupon used for gel-shift, DNase I protection, and Western blot analyses. Western blots were performed by standard methods. After assay of protein concentration, 0.5 μg total protein was loaded in each lane of a 10% SDS-PAGE gel. Specific recognition of GF-1 was provided by a polyclonal antiserum to a peptide of GF-1 (Whitelaw et al., in prep.) kindly provided by S.-F. Tsai.

DNA probes

A 117-bp NcoI–HinfI fragment spanning −259 to −142 of the human Aγ-globin promoter (Martin et al. 1989) was prepared and 5′-end-labeled with [γ-³²P]ATP at the NcoI site by standard methods (Ausubel et al. 1987). A double-stranded oligonucleotide derived from the mouse α1-globin promoter with the sequence GATCTCCGGCAACTGATAAG-GATTCCCTG was also end-labeled and prepared by standard methods. These probes were used in DNase I footprinting, methylation interference, and gel-shift assays, as detailed in the figure legends.

Gel-shift assay

Gel-shift reactions were performed as described. In a typical 20-μl reaction, 0.5 μl COS cell extract, 1 μg of poly[d(I-C)], and 10⁴ dpm radiolabeled DNA probe were used.

DNase I footprinting DNase I footprinting of the Aγ-globin promoter NcoI–HinfI fragment was performed as described previously (Martin et al. 1989). In these reactions, 2 μl of crude extract from COS cells transfected with the appropriate expression plasmid constructs was used. Digested DNA was electrophoresed through an 8% urea–acrylamide gel.

Methylation interference The NcoI–HinfI (−259 to −141) Aγ-globin promoter fragment was end-labeled with [γ-^{32}P]ATP by standard methods and methylated as described by Maxam and Gilbert (1980). A standard gel-shift reaction was scaled up fivefold. After electrophoresis, the gel was exposed wet for 3 hr, and shifted bands were excised. Only the typical (major) GF-1 band was excised; a slower migrating band commonly produced by Δ200–248 was not included (visible in Fig. 6). DNA was eluted from excised bands, purified by Elutips (Schleicher & Schuell), and subjected to the G>A sequence reaction as described by Raymondjean et al. (1988). Approximately equal numbers of counts of each sample were loaded in each lane of an 8% urea–acrylamide gel.

Off-time assay Off-time experiments were performed as described by Li et al. (1989). A 20-μl gel-shift reaction was scaled up fourfold. At time zero, 1 μg of a double-stranded oligonucleotide derived from the sequence of the Aγ-globin promoter from −193 to −169 or the M1α sequence was added. Ten-microliter aliquots were loaded onto a 0.5× TBE/5% PAGE gel (Hoeffer Mighty Small apparatus) at varying times (as indicated) and electrophoresed at 80 V. Gels were run until the probe of the first loaded sample was near the bottom.

Gel-shift titration assay Western blot analysis was used to confirm that COS extracts to be tested contained equivalent amounts of GF-1 (and mutant GF-1) proteins. Samples were then diluted in gel-shift binding buffer. The first dilution contained 100 ng protein and 100 ng poly[d(I-C)] per microliter; subsequent dilutions were made from this mix into binding buffer and preserved the ratio of protein to poly[d(I-C)].

Acknowledgments

References

Antoniou, M. and F. Grosveld. 1990. β-Globin dominant control region interacts differently with distal and proximal promoter elements. *Genes Dev.* **4:** 1007–1013.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.A. Smith, J.G. Seidman, and K. Struhl. 1987. *Current protocols in molecular biology.* Greene Wiley, New York.

Behringer, R., R.E. Hammer, R.L. Brinster, R.D. Palmiter, and T.M. Townes. 1987. Two 3' sequences direct erythroid-specific expression of human β-globin genes in transgenic mice. *Proc. Natl. Acad. Sci.* **84:** 7056–7060.

Behringer, R.R., T.M. Ryan, R.D. Palmiter, R.L. Brinster, and T.M. Townes. 1990. Human γ- and β-globin gene switching in transgenic mice. *Genes Dev.* **4:** 380–389.

Chen, C. and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7:** 2745–2752.

Choi, O.-R. and J.D. Engel. 1986. A 3' enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. *Nature* **323:** 731–734.

———. 1988. Developmental regulation of β-globin gene switching. *Cell* **55:** 17–26.

Davis, R.L., P.-F. Cheng, A.B. Lassar, and H. Weintraub. 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60:** 733–746.

deBoer, E., M. Antoniou, V. Mignotte, L. Wall, and F. Grosveld. 1988. The human β-globin promoter: Nuclear protein factors and erythroid specific induction of transcription. *EMBO J.* **7:** 4203–4212.

Enver, T., N. Raich, A.J. Ebens, T. Papayannopoulou, F. Costantini, and G. Stamatoyannopoulos. 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* **344:** 309–313.

Evans, R.M. 1988. The steroid and thyroid hormone receptor superfamily. *Science* **240:** 889–895.

Evans, T. and G. Felsenfeld. 1989. The erythrocyte-specific transcription factor eryf1: A new finger protein. *Cell* **5:** 877–885.

Evans, R.M. and S.M. Hollenberg. 1988. Zinc fingers: Gilt by association. *Cell* **52:** 1–3.

Evans, T., M. Reitman, and G. Felsenfeld. 1988. An erythroid-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc. Natl. Acad. Sci.* **85:** 5976–5980.

Frampton, J., M. Walker, M. Plumb, and P.R. Harrison. 1990. Synergy between the NF-E1 erythroid-specific transcription factor and the CACCC factor in the erythroid-specific promoter of the human porpholbilinogen deaminase gene. *Mol. Cell. Biol.* **10:** 3838–3842.

Fu, Y.-H. and G.A. Marzluf. 1990. nit-2, the major nitrogen regulatory gene of Neurospora crassa, encodes a protein with a putative zinc finger DNA-binding domain. *Mol. Cell. Biol.* **10:** 1056–1065.

Gorman, C.M., L.F. Moffat, and B.H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2:** 1044–1051.

Grosveld, F., G.B. van Assendelft, D.R. Greaves, and B. Kollias. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* **51:** 975–985.

Kadonaga, J.T., K.R. Carner, F.R. Masiarz, and R. Tjian. 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell* **51:** 1079–1090.

Kim, K.S. and L. Guarente. 1989. Mutations that alter transcriptional activation but not DNA binding in the zinc finger of yeast activator HAP1. *Nature* **342:** 200–203.

Kudla, B., M.X. Caddick, T. Langdon, N.M. Martinez-Rossi, C.F. Bennett, S. Sibley, R.W. Davies, and H. Arst. 1990. The regulatory gene areA mediating nitrogen metabolite repression in Aspergillus nidulans. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc

finger. *EMBO J.* **9:** 1355–1364.

Kunkel, T.A., J.D. Roberts, and R.A. Zkour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154:** 367–382.

Li, R., J. Knight, G. Bream, A. Stenlund, and M. Botchan. 1989. Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. *Genes Dev.* **3:** 510–526.

Ma, J. and M. Ptashne. 1987. A new class of yeast transcriptional activators. *Cell* **51:** 113–119.

Maniatis, T., S. Goodbourn, and J.A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science* **236:** 1237–1245.

Martin, D.I.K., S.-F. Tsai, and S.H. Orkin. 1989. Increased γ-globin expression in a nondeletion HPFH mediated by an erythroid-specific DNA-binding factor. *Nature* **338:** 435–438.

Martin, D.I.K., L.I. Zon, G. Mutter, and S.H. Orkin. 1990. Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature* **344:** 444–446.

Maxam, A. and W. Gilbert. 1980. Sequencing end-labelled DNA with base specific chemical cleavages. *Methods Enzymol.* **65:** 499–560.

Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12:** 7035–7056.

Mignotte, V., F. Eleouet, N. Raich, and P.-H. Romeo. 1989a. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. *Proc. Natl. Acad. Sci.* **86:** 6548–6551.

Mignotte, V., L. Wall, E. deBoer, F. Grosveld, and P.-H. Romeo. 1989b. Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res.* **17:** 37–54.

Mitchell, P.J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245:** 371–378.

Muller, M.M., S. Rupert, W. Schaffner, and P. Mathias. 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature* **336:** 544–551.

Nicolis, S., A. Ronohi, N. Malgaretti, R. Mantovani, B. Giglioni, and S. Ottolenghi. 1989. Increased erythroid-specific expression of mutated HPFH γ-globin promoter requires the erythroid factor NFE-1. *Nucleic Acids Res.* **17:** 5509–5516.

Plumb, M., J. Frampton, H. Wainwright, M. Walker, K. Macleod, G. Goodwin, and P. Harrison. 1989. GATAAG: A cis-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. *Nucleic Acids Res.* **17:** 73–92.

Ptashne, M. 1988. How eukaryotic transcriptional activators work. *Nature* **335:** 683–689.

Raymondjean, M., S. Cereghini, and M. Yaniv. 1988. Several distinct "CCAAT" box binding proteins coexist in eukaryotic cells. *Proc. Natl. Acad. Sci.* **85:** 757–761.

Reitman, M. and G. Felsenfeld. 1988. Mutational analysis of the chicken β-globin enhancer reveals two positive-acting domains. *Proc. Natl. Acad. Sci.* **85:** 6267–6271.

Rutherford, T. and A.W. Nienhuis. 1987. Human globin gene promoter sequences are sufficient for specific expression of a hybrid gene transfected into tissue culture cells. *Mol. Cell. Biol.* **7:** 398–402.

Sadowski, I. and M. Ptashne. 1989. A vector for expressing GAL4(1–147) fusions in mammalian cells. *Nucleic Acids*

Res. **17:** 7539.

Schena, M., L.P. Freedman, and K.R. Yamamoto. 1989. Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. *Genes Dev.* **3:** 1590–1601.

Struhl, G., K. Struhl, and P.M. Macdonald. 1990. The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* **57:** 1259–1273.

Tora, L., J. White, C. Brou, D. Tasset, N. Webster, E. Scheer, and P. Chambon. 1989. The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell* **59:** 477–487.

Townes, T.M., J.B. Lingrel, H.Y. Chen, R.L. Brinster, and R.D. Palmiter. 1985. Erythroid-specific expression of human beta-globin genes in transgenic mice. *EMBO J.* **4:** 1715–1723.

Trainor, C.C., T. Evans, G. Felsenfeld, and M.S. Boguski. 1990. Structure and evolution of a human erythroid transcription factor. *Nature* **343:** 92–96.

Tsai, S.-F., D.I. Martin, L.I. Zon, A.D. D'Andrea, G.G. Wong, and S.H. Orkin. 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* **339(6224):** 446–451.

Vieria, J. and J. Messing. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* **153:** 3–11.

Wall, L., E. deBoer, and F. Grosveld. 1988. The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein. *Genes Dev.* **2:** 1089–1100.

Watt, P., P. Lamb, L. Squire, and N.J. Proudfoot. 1990. A factor binding GATAAG confers tissue specificity on the promoter of the human ζ-globin gene. *Nucleic Acids Res.* **18:** 1339–1350.

Weston, K. and J.M. Bishop. 1989. Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. *Cell* **58:** 85–93.

Wilson, D.B., D.M. Dorfman, and S.H. Orkin. 1990. A non-erythroid GATA-binding protein is required for function of the human preproendothelin-1 promoter in endothelial cells. *Mol. Cell. Biol.* **10:** 4854–4862.

Wood, W.G. 1982. Erythropoiesis and haemoglobin production during development. In *Biochemical development of the fetus and neonate.* Elsevier Biomedical Press.

Yamamoto, M., L.J. Ko, M.W. Leonard, H. Beug, S.H. Orkin, and J.D. Engel. 1990. Activity and tissue-specific expression of the transcription factor NF-E1 multigene family. *Genes Dev.* **4:** 1650–1662.

Youssoufian, H., L. Zon, S.H. Orkin, A.D. D'Andrea, and H.F. Lodish. 1990. Genomic structure and transcription of the mouse erythropoietin receptor gene. *Mol. Cell. Biol.* **10:** 3675–3682.

Zon, L.I., S.-F. Tsai, S. Burgess, P. Matsudaira, G. Bruns, and S.H. Orkin. 1990. The major human erythroid DNA-binding protein (GF-1; NF-E1; Eryf-1): Primary sequence and localization of the gene to the X chromosome. *Proc. Natl. Acad. Sci.* **87:** 668–672.