GATA Zinc Finger Interactions Modulate DNA Binding and Transactivation

Cecelia D. Trainor*, Rodolfo Ghirlando and Melanie A. Simpson

Laboratory of Molecular Biology
National Institute of Diabetes, Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD 20892

Running Title: High affinity GATA-1 binding without transactivation

* Corresponding Author
Cecelia D. Trainor
Phone: (301) 496-5889
Fax: (301) 496-0201
ceceliat@intra.niddk.nih.gov
GATA-1 and other vertebrate GATA factors contain a DNA binding domain composed of two adjacent homologous zinc fingers. Whereas only the C-terminal finger of GATA-1 is capable of independent binding to the GATA recognition sequence, double GATA sites which require both fingers for high affinity interaction are found in several genes. We propose a mechanism whereby adjacent zinc fingers interact to influence the binding and transactivation properties of GATA-1 at a subset of DNA binding sites. Using two such double GATA sites we demonstrate that the N-terminal finger and adjacent linker region can alter the binding specificity of the C-terminal finger sufficiently to prevent it from recognizing some consensus GATA sequences. Therefore, the two zinc fingers form a composite binding domain having a different DNA-binding specificity from that shown by the constituent single C-terminal finger. Furthermore, we compare two of these double sites and show that high affinity binding of GATA-1 to a reporter gene does not necessarily induce transactivation, namely, the sequence of the DNA binding site can alter the ability of GATA-1 to stimulate transcription.

INTRODUCTION

GATA factors are widely distributed, and targeted gene inactivation experiments in mice have shown that five of the six vertebrate GATA family members are essential for embryonic viability (1 - 6). While the vertebrate GATA factors contain two zinc fingers (Figure 1; 7, 8), GATA proteins with only a single finger, most similar to the vertebrate C-terminal finger, exist in fungi and worms (9 - 11). GATA-1, the original isolate, is expressed in erythroid cells, megakaryocytes and mast cells and is required for correct regulation of virtually all erythroid-specific
genes (1). These target genes have distinct temporal patterns of expression during erythropoiesis, and some genes are inhibited rather than activated by GATA-1. The basis for these dissimilar expression patterns, which is not yet fully understood, may involve differences in the DNA recognition sequences.

The GATA transcription factor family is defined by a zinc finger of the form $CX_2C-X_{17}$-CNAC and a basic linker region which is required by GATA-1 for DNA binding (7, 8, 12, 13, 14). The C-terminal finger ($CF^1$) of GATA-1 is both necessary and sufficient for DNA binding to the GATA recognition sequence (WGATAR, 15, 16). The N-terminal finger ($NF^2$) shows no independent DNA binding. It does, however, stabilize GATA-1 interactions with some naturally occurring double GATA sites that are critical for gene expression (17-21). Although the GATA motifs within these high affinity sites are found in diverse orientations and spacing, GATA-1 binds as a monomer. This suggests site-specific variability in the position and orientation of the two GATA-1 zinc fingers on DNA.

Some important binding sites that require both GATA zinc fingers for strong interaction are conserved among species. An overlapping palindromic GATA sequence (GATApal) with an adjacent inverted single site is found in at least three vertebrate hematopoietic GATA-1
promoters. Notably, this site is not found in most GATA regulated genes (21). A double site consisting of non-overlapping direct repeats is also conserved in the testis GATA-1 promoters of mice and rats (22). Both of these sequences are essential for promoter activity and their evolutionary conservation implies that they have a significant and specific function.

Here we identify two additional sites that require the N-terminal finger of GATA-1 for high affinity binding. Using derivatives of these sites we show that the N-terminal finger of GATA-1 can stabilize binding, disrupt binding, or modify the DNA binding specificity of the C-terminal finger. We compare the binding properties of one of these sites, a non-consensus GATC recognition sequence, with binding at canonical (GATA) sites and demonstrate a distinct type of binding to some GATC containing probes by several criteria, in spite of similar binding affinities. We further show that GATA-1 activity can be influenced by its mode of binding: whereas GATA-1 is unable to induce transactivation through the GATC1 binding site, transactivation does occur through several GATA sites placed in the same position on a minimal promoter. Thus the ability of GATA-1 to mediate transactivation can be modified by its DNA binding sequence. These observations provide a new mechanism through which
GATA factors could discriminate between genes and elicit distinct responses from them.
EXPERIMENTAL PROCEDURES.

Construction of reporter plasmids. Double stranded oligonucleotides (Operon Technologies) containing the GATApal, GATC1, CA1 and CA2 sites (Table 1) were phosphorylated and inserted into pαD4 at a SpeI site 38 base pairs 5’ of the TATA box (18): vector blunt ends were generated with Klenow fragment DNA polymerase. The GATApal sequence is that of the mouse GATA-1 promoter (19). pαD GP1 and -2 were prepared from pαD3 by site directed mutagenesis (QuikChange™, Stratagene).

GATA-1 expression vectors. The N-finger of human GATA-1 was prepared as a fusion with the maltose binding protein. pMAL-hNF was constructed by inserting a PCR product encompassing amino acid residues 200 to 251 of human GATA-1 between the EcoRI and SalI sites of pMAL-c2x (New England Biolabs). The PCR primers for the reaction (from 5’ to 3’) were CCGGAATTCCAGGCAAGGAGTGTGTG and ACGCGGTCGACTCATCAAC TGACAATCAGGCGCTT. The construction of plasmids producing MBP fusions with chicken and human DF peptide (pMAL-cDF, pMAL-hDF) and chicken CF peptide (pMAL-cCF) are described in the accompanying manuscript (23).

A clone producing untagged human GATA-1 CF was constructed by insertion of a PCR product digested with Ncol and BclI (coding for amino
acids 253 to 317 of hGATA-1, (24) into pET-11d (Novagen) digested with NcoI and BamHI. The PCR primers (5’ to 3’) were CCCGCCATGGCAGGTACTCAGTGACC and CCGTGATCATCAACTGGAGCCCGTTCTT. Each vector was verified by DNA sequencing.

**Expression and purification of GATA-1 proteins.**

MBP GATA-1 fusion proteins. The N-terminal MBP fusion to human GATA-1 NF was expressed in TB1 E. coli. The cells were grown overnight at 37°C in 1 liter of rich medium with glucose containing 100 µg/ml of ampicillin and 50 µM Zn(OAc)$_2$, diluted 1:4 with the same medium and induced with 0.5 mM IPTG for 4 hours. Cells were lysed in 20 mM Tris (pH = 7.4), 200 mM NaCl, 1 mM EDTA, 1mM DTT, 5 mM benzamidine by sonication. The lysates were clarified by centrifugation at 9000 x g for 30 minutes, diluted 1:5 in the same buffer and purified by affinity chromatography over an amylose resin (New England Biolabs) following the manufacturer’s recommendations. Treatment with Factor Xa and purification over a Q-Sepharose® Fast Flow column (Pharmacia) led to a purified human GATA-1 NF. The samples were adjusted to 10 % glycerol and stored at -80°C. The purification of MBP-cDF, MBP-hDF and MBP-cCF are described in the accompanying manuscript (23).
GATA-1 zinc finger peptides. Peptides were expressed in BL21(DE3) E. coli in LB containing 100 µg/mL of ampicillin and 50 µM Zn(OAc)$_2$. Expression and purification were carried out as previously described (16, 21), except that the HPLC purification step was omitted for human GATA-1 CF peptide: active fractions from the S-column were pooled, adjusted to 10% glycerol and stored at -80°C. Peptide concentrations were determined by SDS-PAGE and coomassie staining, and confirmed by evaluation of the binding activity with saturating amounts of probe.

Nuclear extracts and whole cell extracts containing GATA-1 and GATA-2 were made by standard procedures (25, 26).

**Electrophoretic mobility shift assay (EMSA).** DNA and peptides were incubated in 10 µl of a solution containing 50 mM Tris (pH = 7), 0.0125% triton, 3.2% ficoll, 0.2 mM EDTA and 4 µg/ml of poly (dI-dC) for 15 minutes and electrophoresed in an 8% polyacrylamide gel in 10 mM Hepes, 10 mM Tris and 1 mM EDTA all at room temperature. Binding assays with nuclear extracts were carried out with standard procedures as previously described (21). Oligonucleotides were purchased from Operon Technologies and purified by polyacrylamide gel electrophoresis.
Determination of Dissociation Constants. Dissociation constants, $K_d$, were determined by Scatchard analysis. For GATC1, ES and other high affinity sites, about 1.5 pmol of peptide were titrated with 0.12 to 8 pmol of DNA in a final volume of 10 µl and analyzed by EMSA. A single complex was observed and at the highest probe concentrations, all of the protein was bound. The amounts of free (F) and bound (B) DNA were determined (Molecular Dynamics Phosphorimager) and their ratio, B/F, was plotted against the concentration of the complex in moles per liter. The $K_d$ was determined from the reciprocal of the slope of the best linear fit. In the case of lower affinity interactions more protein ($\approx 4$ pmol) was included to generate sufficient complex. Certain DNA binding sites led to the formation of two complexes. In these cases, the B/F ratio of each complex was determined separately and plotted against the sum of the two bound complexes. Two examples of a Scatchard analysis are shown in Figure 2.

**DNase I Footprinting.** 15,000 cpm of the indicated DNA and 7, 21 or 63 pmol of hDF or cCF peptides were incubated in 50 µl of 10 mM Tris, 5 mM MgCl₂, 1 mM CaCl₂, 50 mM KCl, 2 mM DTT for 30 minutes and then treated with 0.0025 units of DNasel for 2 minutes, all at room temperature. Reactions were stopped by the addition of 4.5 volumes of a solution
containing 200 mM NaCl, 20 mM EDTA, 1% SDS and 100 µg/ml tRNA. DNA samples were extracted with phenol chloroform, ethanol precipitated and were electrophoresed on a 5% denaturing polyacrylamide gel which was dried and autoradiographed.

**Cell culture and transfections.** Cells were maintained and transfected as previously described (18) using 40 units of lipofectin (Gibco BRL) and 4.2 µg of DNA per sample. Cells were incubated with DNA for 5 hours and were harvested 48 hours after transfection, washed in PBS and lysed in reporter lysis buffer (Promega). Luciferase assays were performed according to manufacturer's instructions, and a liquid scintillation assay was used to determine CAT activity (27).
RESULTS

To study the contributions of the individual GATA zinc fingers to DNA binding, we have generated a series of single and double zinc finger peptides and N-terminal fusion proteins by expression in E. coli (Figure 1). Recombinant proteins (or peptides) containing both GATA zinc fingers are referred to as double finger (DF$^3$) proteins, whereas the designations NF and CF are used to indicate proteins containing only the N- or the C-terminal zinc finger respectively. The species of origin of the proteins is designated h$^4$ or c$^5$ for human or chicken, respectively (Figure 1). In this manner, the double finger of human GATA-1 (hDF, 21), the C-terminal fingers of human and chicken (16) GATA-1 (hCF and cCF) and the N-terminal finger of chicken GATA-2 (cGATA-2 NF, 28), were produced. Furthermore, all peptides but cGATA-2 NF were also produced as N-terminal fusions with the E. coli maltose binding protein (MBP$^6$, Figure 1; 23). We compare the binding properties of the single and the double finger peptides at a series of single and double DNA binding sites shown in Table 1.

A conformational change is induced when the N-finger of GATA-1 DF peptide binds to double GATA sites. Previous studies showed that complexes formed between hDF peptide and three naturally occurring
double GATA recognition sequences migrate faster in electrophoretic mobility shift assays (EMSA) than hDF peptide complexes formed with probes containing only single GATA binding sites (21). These migration differences are exemplified by the binding of hDF peptide to the double (GATC1) and single (GATApal M1) site probes in Figure 2. Fast complex migration was observed with all of the double sites tested (Table 1, data not shown), indicating that it is likely a general feature of hDF peptide complexes in which the GATA-1 N-finger participates in DNA binding. In an accompanying manuscript we show that this altered mobility is not due to differential DNA bending (23). The hDF peptide evidently adopts a different conformation as a result of the N-finger interaction with DNA (Figure 2 and data not shown).

**Binding interference between GATA zinc fingers.** Two double GATA binding sites, the GATC1 site mentioned above and the ε-globin gene silencer site (εS) were characterized to demonstrate the involvement of both GATA-1 zinc fingers in binding. A series of mutations were generated for each of these binding sequences to determine the requirements for high affinity interaction with GATA-1. These experiments led to the surprising observation that the N-finger of GATA-1 can prevent the C-
finger from binding to certain DNA sites. A description of these studies is presented below.

The ε-globin gene silencer. The human ε-globin gene silencer has three potential overlapping GATA recognition sequences (εS, Table 1,) which lie 258 bp upstream of the transcription initiation site (29, 30). To determine whether both zinc fingers of GATA-1 interact with this site, dissociation constants for the single (CF) and double finger (DF) peptides were determined by EMSA on εS and derived mutant sites (Table 1). All peptides, as well as the wild type protein, bound tightly to εS. Only single complexes were observed indicating that two protein molecules do not bind to εS simultaneously. In order to identify the DNA bases that participate in binding, a series of εS mutations, which removed the GATA sites, were tested. Removal of GATA site 1 (M1) had no effect on the binding of CF peptide, DF peptide or full-length GATA-1 (Table 1). However, mutation of either site 2 (M2) or 3 (M3) reduced the binding affinity of full-length GATA-1 (data not shown) and of hDF peptide, while only mutation of site 3 (M3) altered CF peptide binding. Therefore, εS site 3 is the preferred site for CF peptide and the binding specificity of CF and DF peptides are distinct. Thus in the 1:1 complex formed, hDF peptide
interacts with both sites 2 and 3 concurrently, suggesting that both zinc fingers interact with the DNA to generate strong binding.

To confirm the involvement of both binding sites in the interaction of hDF peptide with εS, we tested double mutations (εSM1, 2 and εSM2, 3) which lead to single site probes. CF peptides bind well to both probes (Table 1), however the binding of wild type GATA-1 was abolished (Table 1). While a lower affinity of GATA-1 for these single site probes was anticipated, the complete lack of detectable binding was surprising. Therefore, CF peptides bind with higher affinity to these single site probes than does full-length GATA-1, indicative of interference by another region of GATA-1. To localize the cause of this inhibition we also tested the binding of hDF peptide to these probes. Like the native protein, hDF peptide (which binds strongly to εS) fails to bind εSM1, 2 or εSM2, 3 (Table 1). As shown in the accompanying manuscript, MBP fusions with the double fingers of chicken and human GATA-1 also bind well to εS (23), but fail to bind to εSM1, 2 and εSM2, 3 (data not shown). We conclude that the presence of the GATA-1 N-terminal finger (finger plus adjacent arm) is sufficient to prevent the C-finger from binding with high affinity to some DNAs. The CF peptide alone binds strongly to εSM1, 2 and εSM2, 3, the DF peptide, MBP-DF fusion proteins and wild type GATA-1, not at all.
GATA-1 binding to GATC sequences. To ascertain whether a similar binding inhibition or modulation occurs with other DNA sites and GATA factors, we used a series of probes derived from (not identical to, see Table 1) those selected by GATA-2 and -3 in a site selection study (31). These probes contain a non-canonical motif having an AGATCTTA consensus. Unlike GATA-1, the N-terminal fingers of GATA-2 and -3 can bind independently to various DNA sites that include this consensus (28). In fact, the cGATA-2 NF, cGATA-3 NF and cGATA-1 CF peptides all bind with similar affinities to many such DNAs (for example GATC1, -3 and –4, Table 1; 28). The advantage of these probes over canonical binding sites is that they allow us to examine the reciprocal influence of the GATA-2 C-finger on N-finger binding (in addition to the effects of NF on CF binding).

A 26 base pair probe (GATC1, Table1) derived from a larger sequence that was selected by GATA-2 and -3 (31) was used. As shown in Figure 2 and Table 1, it contains adjacent GATA and GATC binding sites. To characterize GATA-1 binding to this DNA, mutational analyses were performed. Alteration of the GATA site 5’ of the GATC motif did not modify the binding of cGATA-1, cGATA-2, or hDF peptide (GATC2, Table 1 and 31). However, removal of five bases from the 5’ end of GATC2 did
abolish the binding of all three of these proteins (GATC3 and GATC4, Table 1). The four most 3’ bases were dispensable for binding (GATC5 and GATC6, Table 1), but the cytosines within the palindromic GATC site were essential (GATC7, Table 1). In marked contrast to these results, GATA-1 CF peptides bind to all of these probes except GATC7 (Table 1).

We note that all GATC probes used here, except for GATC7, contain a second consensus sequence, T/GAAG, previously identified as a site for GATA-1 NF (AGATCTTA, 32). While this sequence may contribute to the binding of DF peptide and GATA-1 to these probes, it is not sufficient, as no complexes are formed with GATC3 and –4 (which contain it).

In summary, while a canonical single GATA site is only 6 bp in length, our mutational analysis has shown that binding of GATA-1 and hDF peptide to GATC2 requires sequences that are at least 14 base pairs apart. Surprisingly, it does not require the canonical GATA sequence (present in GATC1).

We found it noteworthy that all of the double finger proteins tested above (i.e. GATA-1, GATA-2 and hDF peptide) failed to bind to GATC3 or GATC4, even though the GATA-1 CF and cGATA-2 NF peptides bind well. MBP-cDF also does not bind to these probes (data not shown). Thus the presence of a second zinc finger is sufficient to inhibit binding, this time
to a GATC sequence, even though the double finger proteins bind strongly to other GATC probes (GATC1, -2, -5, and -6, Table 1). Sequences flanking the consensus evidently contribute to the binding properties observed here and in (31). Analogously, the C-finger of chicken GATA-2 interferes with binding of the N-finger. While cGATA-2 NF binds to GATC3 and -4 (Table 1; 28), it is unable to do so when within the context of the wild type protein (Table 1), or cGATA-2 double finger peptide (data not shown, amino acids 277 to 396). Therefore, the native arrangement and linkage of the fingers of GATA-1 and GATA-2 alters their DNA binding specificity. Some single sites are structured such that the two-finger combination binds well (e.g. WGATAR) whereas other DNA sites (e.g. εSM1, 2 and εSM2, 3, GATC3 and -4) do not allow for a stable interaction with the double finger peptide (or proteins), GATA-1 or GATA-2.

**Covalent linkage of the N- and C-fingers is required for their interaction.** To determine whether the GATA fingers must be linked to influence one another, we performed mixing experiments with GATA-1 CF and NF. MBP-hNF, hNF, MBP-cCF and CF (h or c) peptides were tested individually and together with the GATApal, GATC4, εSM1, 2 and εSM2, 3 probes (Figure 3A and B, and data not shown). In all cases the presence of
MBP-hNF or hNF showed no influence on complex formation. We therefore conclude that the GATA-1 fingers must be covalently linked to influence one another positively (GATApal) or negatively (GATC4, εSM1, 2 and εSM2, 3).

**GATA-1 DF peptide interactions with GATC probes are not equivalent to those with canonical binding sites.**

As noted above, GATC1 is a site to which both GATA fingers bind, because it is fast migrating when complexed with hDF peptide. We initially assumed that this was due to the presence of both the GATA and the GATC consensus sequences on this probe. However, the GATC2 complex (which lacks the GATA site) is also fast migrating (data not shown), suggesting that the N-finger interacts with a non-canonical sequence within GATC2. To compare GATA protein interactions at GATC containing sites with those at canonical binding sites, C to A substitutions (CA) in GATC1 and GATC2 were tested with hDF peptide. With these mutations the fast migrating GATC1 complexes are partially converted to slower migrating forms, consistent with binding through a single finger and site (CA1 and CA2 probes, Figure 4A). Surprisingly, modifying these sequences to ones more similar to the WGATAR consensus reduced the overall binding
affinity, as well as the affinity of the N-finger of hDF peptide for the DNA. This confirms a novel type of DF peptide binding to the GATC1 probe that requires an extended non-consensus binding sequence and both zinc fingers.

**The N-finger of GATA-1 can modulate the DNA binding specificity of the GATA-1 C-finger.** Even though both the N- and C-terminal GATA-1 zinc fingers are involved in the recognition of GATC1 and other double site probes, the DF peptide has a different specificity than CF peptide. In addition to preventing DNA binding, zinc finger interactions can also lead to more subtle changes in specificity. A complex between cCF peptide and a palindromic GATA site (GATApal) is competed effectively by a single GATA sequence (β/ε), but not as efficiently by GATC1 (Figure 4B). In contrast, the DF peptide complex with the same probe displays the opposite competition pattern (Figure 4B), namely the β/ε single GATA site is a much less effective competitor than GATC1 when NF and CF are joined. Therefore the two zinc fingers act in concert and the presence of the N-finger modifies the inherent specificity of the C-finger.
In vitro footprinting reveals site-specific differences between CF and DF peptides. The interactions of DF and CF peptides with canonical binding sites or the GATC1 sequence were compared by DNase1 footprinting. GATA-1 DF and cCF peptides show strong partially overlapping footprints on the cluster of consensus GATA recognition sequences in the mouse (m) hematopoietic GATA-1 promoter (Figure 5A). The hDF peptide footprint is larger than that of cCF peptide: the protection patterns at the 3’ ends are identical, but the DF peptide footprint extends over the partial inverted GATA sequence near the 5’ end of the DNA. While DF protects the G of this partial site, cCF peptide causes enhanced cleavage at this position.

The DF peptide protects a 16 base pair region of GATC1 that includes both the GATC site and 8 base pairs 5’ of it. Surprisingly, GATA-1 CF peptide does not protect this DNA site, even at 20 fold higher protein concentration than shown here (Figure 5B), in spite of the high binding affinity ($K_d = 6$ nM). CF peptide does protect canonical GATA sites (Figures 5A, 5B), but the $K_d$ for this interaction is reported to be 0.78 nM (28). CF peptide may therefore require tighter binding affinities than DF for footprint formation, or may bind to GATC1 in a manner that still allows DNase I access to the DNA. In summary, cCF peptide binds to, but is unable
to protect the GATC1 DNA, however the addition of the N-finger (DF peptide) results in protection from DNase I.

Taken together these results confirm that GATC1 contains an extended compound site on which both GATA-1 fingers participate in complex formation. Thus GATA-1 binding to GATC1 is demonstrably different than to the consensus (GATA) recognition sequence by several criteria, suggesting that the mode of binding to these two types of sequences is distinct, and results in different conformations of the bound protein.

**High Affinity DNA binding by GATA-1 does not always induce transactivation.** GATA-1 can contact DNA containing double GATA binding sites through one or two zinc fingers and it is possible that the activity of GATA-1 might vary as a result of this differential binding. Distinctly structured binding sites could have gene-specific functions. Interactions of GATA-1 with DNA may induce conformational changes in the protein that alter the interface between GATA-1 and auxiliary factors or change the potency of its transactivation domains by repositioning them. Such conformational changes might elicit different transcriptional responses, as has been shown for a few other transcription factors
(reviewed in 33). If so, high affinity binding of GATA-1 to a target site might not always lead to transactivation.

To test this premise we used a reporter construct containing a minimal promoter with a double GATA site (pαD3, Figure 6A), previously employed to show that GATA-1 is a transactivator (18). It was shown that co-expression of GATA-1 with this reporter stimulated its transcription approximately 100 fold, whereas the same reporter gene with mutated GATA sites (pαD4) led only to a 5-fold stimulation. This activation is not absolutely dependent on the position or orientation of the GATA sites. For example, a pαD4 derivative (pαD4/5, Figure 6A) with three GATA sites placed upstream of the original sites is induced 40 fold regardless of insert orientation. Derivatives of pαD4/5 containing five or more adjacent copies of the same GATA sequence are much more active than pαD3 in erythroid cells (18). Constructs with GATA sites inserted 17, 18, 21 or 26 base pairs upstream of this position (Spe I site, Experimental Procedures) were all active in this assay (data not shown).

In this study GATC1 was compared with a canonical high affinity double GATA site, mGATApal. Both were placed in pαD4 at the distal position and tested for transactivation by cGATA-1. Reporter plasmids with one copy of each insert in either orientation and two or four adjacent
copies were compared (Figure 6B). Although these two binding sites do not
differ substantially in affinity, significant differences in transactivation
were observed. As expected, all constructs containing GATApal \(K_d = 8 \text{ nM}\)
were activated by GATA-1. In contrast, all GATC1 containing plasmids \(K_d\)
\(= 5 \text{ nM}\) yielded transactivation levels similar to the negative control,
\(\alpha D4\) (Figure 6B). This lack of transactivation is not specific to
fibroblasts as the GATC1 containing plasmids were much less active than
GATApal containing plasmids in an erythroid precursor cell line that
expresses GATA-1 (data not shown; 6C2 cells, 34). This failure to support
transactivation is not a unique property of the GATC1 sequence as cGATA-
1 expressed in QT6 cells also failed to activate \(\alpha D4\) containing a single
copy of \(\varepsilon S\) DNA (data not shown). Thus the presence of a high affinity GATA
binding site on a reporter gene does not necessarily result in
transactivation; the sequence of the DNA binding site is important.

To further examine the relationship between binding affinity and
transactivation we tested other GATA sites. The direct GATA repeats in
the \(\alpha D3\) reporter gene (figure 6A) can be converted to a sequence which
matches the consensus for GATApal by a single A to C base change (Figure
6C). Two independent identical clones containing this mutation were
generated and tested (\(\alpha D3\) GP1 \& -2). Although this nucleotide change
results in an increase in DNA binding affinity (21), constructs containing this mutation are less active than pαD3 in response to GATA-1. The activity of these new GATApal-containing plasmids (pαD3 GP1 & -2) levels off with increasing GATA-1, whereas the activation of pαD3 continues to rise (Figure 6C). Thus, the orientation of the GATA sequences (direct repeat or palindrome) relative to one another influences transactivation, and the degree of activation does not correlate with binding affinity. In summary, strong binding of GATA-1 to a reporter gene does not necessarily lead to transactivation, and the GATA-1 protein is transcriptionally inactive in response to the GATC1 sequence. These data suggest that DNA can modulate GATA-1 activity in a sequence dependent manner, by changing its conformation.

**Transactivation is restored through C to A substitution in GATC sites.** This model of GATA-1 action predicts that a GATA-1 protein lacking the N-terminal zinc finger might stimulate transcription of the GATC1 containing plasmids, even though the wild type protein does not. Because the N-finger induces the conformational changes observed here, and is required for the distinct type of interaction seen between GATA-1 and the GATC1 sequence, inactivation of the N-finger might restore the
“normal” interaction. Consequently we tested cGATA-1 mutants in which the first two cysteines of the N-terminal finger were replaced (NF\textsuperscript{-} GATA-1): these mutants have previously been shown to bind to and activate GATA dependent reporter genes (13).

In our hands NF\textsuperscript{-} GATA-1 did activate p\alpha D3 (canonical binding site, Figure 6A), but the GATC1 reporter plasmids were not stimulated (data not shown). However EMSA analysis shows that the NF\textsuperscript{-} GATA-1 mutants do not bind to GATC1 or \varepsilon S oligonucleotides, even though they do bind to other canonical GATA sequences (Figure 7A and data not shown). These binding studies confirm, once more, the involvement of both zinc fingers in GATA-1 interactions with GATC1 and \varepsilon S.

We note that as with the DF peptide, binding of the N-finger of full-length cGATA-1 to DNA also causes fast migration of the protein/DNA complex. Double site probes such as GATApal lead to faster migration than single site probes such as GATApal M1 with wild type cGATA-1, but NF\textsuperscript{-} cGATA-1 (HB2) complexes with both probes migrate slowly and at identical positions (Figure 7A). Therefore, the conformational change and altered migration induced by N-finger binding (Figure 2) is not specific to the GATA-1 DF peptide but also occurs with the native protein.
In an alternate approach, to try to restore transactivation to the GATC driven reporter genes, we used the C to A substitution binding sites, CA1 and CA2 (shown in Table 1 and Figure 4A), in the transactivation assay. We reasoned that if the lack of activation observed with the GATC1 site was due to the conformation of GATA-1 on the DNA, then restoring the “normal” conformation by converting GATC1 to a canonical binding site (GATA) might restore transactivation. This is in fact the case. Even though the C to A substitutions (CA1 and CA2) result in reduced binding affinity relative to GATC1, pαD4 plasmids containing these sequences are activated by GATA-1 (Figure 7B), unlike those containing GATC1. In summary, the activity of GATA-1 is modulated by the sequence of its DNA binding site, and the level of activation achieved does not necessarily correlate with the DNA binding affinity. DNA induced conformational changes alter GATA-1 activity.
DISCUSSION

Novel properties of GATA proteins. The N- and C-terminal zinc fingers of GATA-1 and GATA-2 influence one another in DNA binding: their activity and specificity are modified in the hDF peptide and full-length proteins such that binding to certain sites is inhibited. This change is most likely the result of a DNA dependent intramolecular interaction between the GATA fingers and (or) basic arms. Such an interaction may also be involved in the binding stabilization mediated by GATA-1 NF on other double sites. Indeed, GATA-1 has been shown to weakly dimerize on DNA through an association between the N- and C-terminal zinc fingers of separate molecules, indicating that the fingers have a specific affinity for one another (35). For similar intramolecular interactions to occur, the fingers must be capable of motion relative to one another through the 22 amino acids linking them (Figure 1). This linker region of GATA-1 has recently been shown by NMR to have little secondary structure and is consequently unlikely to independently restrict such relative movement of the fingers (36). Therefore the interference seen with GATA-1 likely reflects an association between regions of the protein contained within the two zinc fingers, as determined by their relative orientation when
bound to DNA. Similar domain interactions may also occur in the absence of DNA.

A second new attribute of GATA-1, namely the inability to activate transcription at a high affinity-binding site, is also described here. In an accompanying manuscript we show that this DNA dependence of activation does not involve site specific DNA bending (23). DNA induced modulation of transactivation has also been seen with the steroid hormone receptors which are allosterically regulated by DNA. Activation, repression, or lack of a response can occur upon DNA and ligand binding to these receptors, all dependent on the binding sequence rather than the affinity of the interaction: the DNA induces conformational changes that specify the activity of these receptors (33). Recently a similar conclusion was reached in studies with nuclear factor-κB (NF-κB): it was suggested that DNA acts as an allosteric effector of this transcription factor (37). Here we show that GATA-1 activity can also be influenced by the sequence of its cognate DNA site.

Implications for GATA factor activity. The characteristics of GATA proteins described here could influence many facets of hematopoiesis by determining the precise conformation of transcription factor binding. In erythroid cells, interaction with factors which associate with GATA-1
(e.g. EKLF, SCL, RBTN2 and PU.1; 38 - 42) and are critical for development
might be affected by the conformation of the GATA-1-DNA complexes. For
example, the N-terminal finger of GATA-1 is required for terminal
erythroid differentiation (43), and the N-terminal fingers of the GATA
proteins contain regions which interact with the essential cofactors FOG
and FOG 2 (Friend of GATA; 44 - 47). GATA-1 NF binding to double DNA
sites could influence the association and activity of these key factors. If
the GATC1 site or other sites that do not support transactivation exist in
regulatory regions of the genome, they could sequester GATA-1 in a
transcriptionally inactive state. The new DNA binding properties reported
here may presage GATA interaction with other important DNA sites that
have so far gone undetected.

Alterations in binding specificity mediated by adjacent zinc fingers
are somewhat unusual; many studies indicate that while tandem Cys$_2$His$_2$
zinc fingers can cooperate with one another to stabilize DNA binding, they
retain their individual specificities when separated by only seven amino
acids (48, 49). The N-terminal finger of GATA-1 participates in DNA
binding (13, 17, 32, 50), but previous studies did not distinguish between
contributions mediated by each finger acting independently and the
formation of a composite binding domain. We have shown that the GATA
zinc fingers within a single molecule can cooperate to form a binding domain with specificity distinct from that of the individual fingers, and that the fingers heavily influence the binding of one another. Moreover, since high affinity binding does not always lead to transactivation, the purpose of some GATA sites may be to mediate other functions. These could include transcriptional repression, alteration of chromatin structure or mediation of locus control region (LCR) activity. The GATA proteins have been implicated in each of these processes (29, 30, 32, 50 - 53). For example, hypersensitive sites (HS), HS2, HS3 and HS4 of the β-globin LCR all contain conserved double GATA sites, which may be involved in HS site formation (54, 55). This unusual versatility in DNA binding mediated through the interaction of GATA zinc fingers may have led to a corresponding functional diversity for GATA-1 and other members of the GATA family.

**Acknowledgements.** The authors are grateful to Gary Felsenfeld for excellent advice and for supporting this work. Thanks to Jim Omichinski for helpful discussions and for some of the peptides used in these studies. The authors also thank Adam Bell and Marc Reitman for suggestions on
this manuscript. We thank Donn Viviani for making some of the constructs used here.

REFERENCES

1. Orkin, S. H. (1992) Blood **80**, 575-581
2. Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S. H., and Costantini, F. (1991) Nature **349**, 257-260
3. Tsai, F. Y., Keller, G., Kuo, F. C., Weiss, M., Chen, J., Rosenblatt, M., Alt, F. W., and Orkin, S. H. (1994) Nature **371**, 221-226
4. Pandolfi, P. P., Roth, M. E., Karis, A., Leonard, M. W., Dzierzak, E., Grosveld, F. G., Engel, J. D., and Lindenbaum, M. H. (1995) Nat. Genet. **11**, 40-44
5. Molkentin, J. D., Lin, Q., Duncan, S. A., and Olson, E. N. (1997) Genes Dev. **11**, 1061-1072
6. Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R., and Grosveld, F. (1999) Development **126**, 723-732
7. Tsai, S.-F., Martin, D.I., Zon, L.I., D'Andrea, A.D., Wong, G.G. and Orkin, S.H. (1989) Nature **339**, 446-451
8. Evans, T., and Felsenfeld, G. (1989) Cell **58**, 877-885
9. Fu, Y. H., and Marzluf, G. A. (1990) Mol. Cell. Biol. **10**, 1056-65
10. Kudla, B., Caddick, M. X., Langdon, T., Martinez-Rossi, N. M., Bennett, C. F., Sibley, S., Davies, R. W., and Arst, Jr., H. N., (1990) EMBO J. **9**, 1355-1364
11. Zhu, J., Hill, R. J., Heid, P. J., Fukuyama, M., Sugimoto, A., Priess, J. R., and Rothman, J. H. (1997) Genes Dev. **11**, 2883-2896
12. Yamamoto, M., Ko, L. J., Leonard, M. W., Beug, H., Orkin, S. H., and Engel, J. D. (1990) Genes Dev. **4**, 1650-1662
13. Yang, H. Y., and Evans, T. (1992) Mol. Cell. Biol. **12**, 4562-4570
14. Omichinski, J. G., Trainor, C., Evans, T., Gronenborn, A. M., Clore, G. M., and Felsenfeld, G. (1993) Proc. Natl. Acad. Sci. U. S. A. **90**, 1676-1680
15. Evans, T., Reitman, M., and Felsenfeld, G. (1988) Proc. Natl. Acad. Sci. U. S. A. **85**, 5976-5980
16. Omichinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., and Gronenborn, A. M. (1993) Science **261**, 438-446
17. Martin, D. I., and Orkin, S. H. (1990) Genes Dev. **4**, 1886-1898
18. Evans, T., and Felsenfeld, G. (1991) Mol. Cell. Biol. 11, 843-853
19. Tsai, S. F., Strauss, E., and Orkin, S. H. (1991) Genes Dev. 5, 919-931
20. Hannon, R., Evans, T., Felsenfeld, G., and Gould, H. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 3004-300
21. Trainor, C. D., Omichinski, J. G., Vandergon, T. L., Gronenborn, A. M., Clore, G. M., and Felsenfeld, G. (1996) Mol. Cell. Biol. 16, 2238-2247
22. Onodera, K., Yomogida, K., Suwabe, N., Takahashi, S., Muraosa, Y., Hayashi, N., Ito, E., Gu, L., Rassoulzadegan, M., Engel, J. D., and Yamamoto, M. (1997) J. Biochem. (Tokyo) 121, 251-263
23. Ghirlando, R. and. Trainor, C. D. (2000) J. Biol. Chem. In press
24. Trainor, C. D., Evans, T., Felsenfeld, G., and Boguski, M. S. (1990) Nature 343, 92-96
25. Dignam, J. D., Lebowitz, R.M., and R.G. Roeder. (1983) Nucleic Acids Res. 11, 1475-1489.
26. Kumar, V., and Chambon, P. (1988) Cell 55, 145-156
27. Neuman J R, Moreney, C. A., and Russion K O. (1987) BioTechniques 5, 444-449
28. Pedone, P. V., Omichinski, J. G., Nony, P., Trainor, C., Gronenborn, A. M., Clore, G. M., and Felsenfeld, G. (1997) EMBO J. 16, 2874-2882
29. Peters, B., Merezhinskaya, N., Diffley, J. F., and Noguchi, C. T. (1993) J. Biol. Chem. 268, 3430-3437
30. Raich, N., Clegg, C. H., Grofti, J., Romeo, P. H., and Stamatoyannopoulos, G. (1995) EMBO J. 14, 801-809
31. Ko, L. J., and Engel, J. D. (1993) Mol. Cell. Biol. 13, 4011-4022
32. Whyatt, D. J., deBoer, E., and Grosveld, F. (1993) EMBO J. 12, 4993-5005
33. Lefstin, J. A., and Yamamoto, K. R. (1998) Nature 392, 885-888
34. Beug, H., Doederlein, G., Freudenstein, C. and Graf, T. (1982) J. Cell Physiol. Suppl. 1, 195-207
35. Mackay, J. P., Kowalski, K., Fox, A. H., Czolij, R., King, G. F., and Crossley, M. (1998) J. Biol. Chem. 273, 30560-30567
36. Kowalski, K., Czolij, R., King, G. F., Crossley, M., and Mackay, J. P. (1999) J. Biomol. NMR 13, 249-262
37. Menetski, J. P. (2000) J. Biol. Chem. 275, 7619-7625
38. Merika, M., and Orkin, S. H. (1995) Mol. Cell. Biol. 15, 2437-2447
39. Osada, H., Grutz, G., Axelson, H., Forster, A., and Rabbitts, T. H. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 9585-9589
40. Osada, H., Grutz, G. G., Axelson, H., Forster, A., and Rabbitts, T. H. (1997) Leukemia 11 Suppl. 3, 307-312
41. Wadman, I. A., Osada, H., Grutz, G. G., Agulnick, A. D., Westphal, H., Forster, A., and Rabbitts, T. H. (1997) EMBO J. **16**, 3145-3157
42. Rekhtman, N., Radparvar, F., Evans, T., and Skoultchi, A. I. (1999) Genes Dev. **13**, 1398-1411
43. Weiss, M. J., Yu, C., and Orkin, S. H. (1997) Mol. Cell. Biol. **17**, 1642-1651
44. Tsang, A. P., Visvader, J. E., Turner, C. A., Fujiwara, Y., Yu, C., Weiss, M. J., Crossley, M., and Orkin, S. H. (1997) Cell **90**, 109-119
45. Tsang, A. P., Fujiwara, Y., Hom, D. B., and Orkin, S. H. (1998) Genes Dev. **12**, 1176-1188
46. Lu, J. R., McKinsey, T. A., Xu, H., Wang, D. Z., Richardson, J. A., and Olson, E. N. (1999) Mol. Cell. Biol. **19**, 4495-4502
47. Holmes, M., Turner, J., Fox, A., Chisholm, O., Crossley, M., and Chong, B. (1999) J. Biol. Chem. **274**, 23491-23498
48. Klug, A., and Schwabe, J. W. (1995) FASEB J. **9**, 597-604
49. Klug, A. (1995) Ann. N. Y. Acad. Sci. **758**, 143-160
50. Merika, M., and Orkin, S. H. (1993) Mol. Cell. Biol. **13**, 3999-4010
51. Imagawa, S., Yamamoto, M., and Miura, Y. (1997) Blood **89**, 1430-1439
52. Grosveld, F., Dillon, N., and Higgs, D. (1993) Baillieres Clin. Haematol. **6**, 31-55
53. Boyes, J., Omichinski, J., Clark, D., Pikaart, M., and Felsenfeld, G. (1998) J. Mol. Biol. **279**, 529-544
54. Pomerantz, O., Goodwin, A. J., Joyce, T., and Lowrey, C. H. (1998) Nucleic Acids Res. **26**, 5684-5691
55. Stamatoyannopoulos, J. A., Goodwin, A., Joyce, T., and Lowrey, C. H. (1995) EMBO J. **14**, 106-116
**Abbreviations:**

1. CF, C-terminal zinc finger and C-terminal basic arm
2. NF, N-terminal zinc finger and adjacent arm
3. DF, double finger peptide
4. h, human
5. c, chicken
6. MBP, E. coli maltose binding protein
7. EMSA, electrophoretic mobility shift assay
8. m, mouse
9. LCR, locus control region
10. HS, hypersensitive site
FIGURE LEGENDS

Figure 1. GATA zinc finger peptides. The amino acid sequence of the human GATA-1 double finger (hDF) peptide is presented. Structural features of hDF are derived from the NMR solution structures of cGATA-1 CF peptide on DNA (16), and of hGATA-1 NF (36). The arrow indicates the start of hCF peptide, which is represented by the solid line just below hDF peptide. The solid line immediately below that represents cCF peptide which is highly homologous to hCF. MBP-hCF contains the same GATA-1 derived amino acids as hCF, while MBP-cCF contains amino acids 156 to 223 (23). The broken line indicates the GATA-1 derived amino acids present in MBP-hNF and free hNF. Free hNF peptide contains four extra amino acids at the N-terminus (I, S, E, and F) which are derived from the expression vector. The amino acid sequence of cGATA-2 NF is from (28).

Figure 2. The binding of the N-finger of GATA-1 DF peptide induces a conformational change. 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75 or 2 pmol of the indicated $^{32}$P labeled DNA were incubated with 1.5 pmol of hDF peptide in a final volume of 10 µl and analyzed by EMSA. The amounts of bound and free probe were determined with a Molecular Dynamics
Phosphorimager. The corresponding Scatchard analyses are shown. The absolute value of the reciprocal of the slope obtained from the best linear fit, corresponds to $K_d$. The arrow indicates a canonical GATA site while the GATC site is underlined.

Figure 3. GATA-1 zinc fingers must be covalently linked to influence one another during DNA binding. (A) 0.125, 0.5 and 1 pmol of the indicated probes were mixed with MBP-cCF, MBP-hNF or both. Samples were incubated for 10 minutes at room temperature and analyzed by EMSA. Proteins were quantified by SDS-PAGE and coomassie staining, and approximately equivalent amounts of each protein were used. (B) hNF and MBP-hNF were tested with MBP-cCF, and cCF peptides as in A. Again, equivalent amounts of the N- and C-finger peptides were used.

Figure 4. GATA-1 CF and DF peptides show distinct binding specificities on GATC containing probes.

(A) 0.25, 0.5, 0.75, 1.5, or 2 pmol of the indicated $^{32}$P labeled DNA was incubated with 10.5 pmol (3X) of hDF peptide for 15 minutes at room temperature and analyzed by EMSA. A titration of GATC1 with 3.5 pmol (1X) of hDF peptide is also shown. P marks the free probe. Dissociation
constants, $K_d$, were determined by Scatchard analysis. For probes that give two complexes, $B/F$ was calculated for the individual complex and this ratio was plotted against the sum of the two bound species. Because similar values were obtained for the upper and lower complexes with CA1 and CA2, the average value for the two complexes is reported.

(B) 3.5 pmol of cCF or hDF peptides were incubated with $^32$P labeled mGATApal DNA and competed with 3.3, 6.6, 10 and 13 pmol of unlabeled $\beta/\epsilon$ DNA (a single site from the chicken $\beta/\epsilon$ enhancer element, Table 1) or GATC1 DNA. P indicates the free probe.

**Figure 5. In vitro DNase I footprinting shows that both GATA-1 fingers participate in the protection of GATC1.**

330 bp BssHII to PvuII restriction fragments from plasmids containing (A) the GATA sites from the hematopoietic mGATA-1 promoter, (B) the GATC1 sequence ($p\alpha D4$/GATC1) or a double GATA site consisting of direct repeats of the consensus binding site ($p\alpha D3$), were endlabeled with Klenow fragment DNA polymerase, and subjected to DNase1 footprinting. Lanes labeled - have no GATA peptide; the triangles indicate increasing added amounts of the peptides indicated. The brackets mark the protected regions of the DNA, the lines indicate the corresponding protected
sequence, and the arrows indicate the GATA sites. * indicates enhanced cutting. The antisense strand is labeled for the mGATA-1 promoter and the αD3 probes.

**Figure 6. The amount of transactivation induced by GATA-1 does not correlate with its affinity for the target DNA.**

(A) A diagram of two reporter plasmids (containing the chloramphenicol acetyl transferase gene, CAT) used in previous studies, and results of those studies are indicated. To generate pαD4/5, the GATA sites in pαD3 were mutated to yield pαD4, and three copies of the oligonucleotide shown above pαD4/5 were ligated to a Spe I restriction site upstream of this modification. Transactivation of pαD3 was always set to 100% and results with other constructs were normalized to this value. Arrows indicate GATA sites. The black box indicates the position of the αD-globin gene double GATA sites. The gray boxes indicate the GATA sites placed in the distal position.

(B) Transactivation with the test plasmids and with the negative control plasmid, pαD4, is shown. Oligonucleotides containing either mGATApal (left graph) or GATC1 (right graph) were placed in the distal position of pαD4 and tested with GATA-1 in QT6 cells. 3 µg of each reporter plasmid
and 0.8 µg of an expression vector producing chicken GATA-1 were
cotransfected into QT6 cells along with a plasmid expressing luciferase
as a control for transfection efficiency. CAT activity was normalized to
luciferase activity. Results displayed (as transactivation) are the ratio of
CAT activity from each reporter in the presence versus the absence of
GATA-1, and are the average of at least 3 independent experiments. Bars
represent the standard error. The white box with – indicates mutated
GATA sites in pαD4. Gray boxes indicate the position and number of
oligonucleotide inserts, and arrows indicate their orientation. The dotted
lines across the graphs mark the activity of the negative control, pαD4.

(C) An A to C base substitution in the double GATA site of pαD3 was made
by site directed mutagenesis, and this change resulted in increased GATA-
1 binding affinity as indicated. Two independent identical clones were
isolated (GP1 and GP2) and transfected into QT6 cells with incremental
two fold increases in GATA-1 expression plasmid (62 to 1000 ng). CAT
activity in counts per minute is shown in comparison to that of pαD3.
Arrows indicate GATA sites.

**Figure 7.** A single C to A base change in the GATA binding site
of pαD4/ GATC1 results in increased transactivation.
(A) An NF- mutant of GATA-1 does not bind to GATC1 or εS DNA. A vector expressing a mutant of cGATA-1 in which the first two cysteines of NF were changed to serines (HB2; 13) was transfected into QT6 cells. Whole cell extracts were made 48 hours later and compared by EMSA with extracts from cells transfected with an expression plasmid producing wild type cGATA-1. Probes are indicated above the gel and the sequences can be found in Table 1. Arrows indicate full-length GATA-1 while the arrowhead indicates a degradation product. The wild type GATA-1 complex with GATApal is fast migrating relative to the single site complex (GATApalM1) and to the NF- GATA-1 (HB2) complexes.

(B) The CA1 and CA2 oligonucleotides (Table 1 and Figure 4A) were inserted into the Spe I site of pαD4 in 1, 2 or 4 copies and tested in the transactivation assay as described in Figure 6. The number following the construct name indicates the copy number of the oligonucleotide insert in each construct.
Table 1. The sequence of the DNA probes and competitors used in this study is presented. All DNA sequences in this manuscript are listed from 5’ to 3’. Canonical GATA recognition sequences are marked by arrows, whereas GATC sequences are underlined. Arrows numbered 1 to 3 mark three potential GATA sites in εS DNA. The dissociation constants for the interactions were determined by Scatchard analysis as described in the Experimental Procedures section and are given in nM. Representative Scatchard plots are shown in Figure 2. Values are the mean of at least 3 experiments and standard errors were < 50% of the mean. For interactions in which no complex was visible we estimated the $K_d$ by determining the limits of detection of our assay. The column labeled GATA-1 represents EMSA analysis of wild type GATA-1 from several sources: GATA-1 (h or c) from erythroid cell extract and recombinant human and chicken GATA-1 from QT6 cell extract were tested and gave comparable results. The column marked GATA-2 contains EMSA results using recombinant cGATA-2 expressed in QT6 cells. + indicates an intense signal while – indicates no detectable binding. ++ represents strong binding but the $K_d$ was not determined. Lower case letters demarcate base changes.
| Name       | DNA Sequence            | hGATA-1 DF peptide | hGATA-1 CF peptide | cGATA-1 CF peptide | GATA-1 | GATA-2 | GATA-2 NF peptide |
|------------|-------------------------|--------------------|--------------------|--------------------|--------|--------|------------------|
| εS         | GAATGGAGAGATGGATATCATTTTGAAG | 8                  | 5                  | 5                  | +      |        |                  |
| εSM1       | GAATGGAGAGATGGATATCATTTTGAAG | 5                  | 5                  | +                  |        |        |                  |
| εSM2       | GAATGGGAGAatTGGATATCATTTTGAAG | 300                | 5                  | +                  |        |        |                  |
| εSM3       | GAATGGGAGAGATGGATAtTTTTTGAAG | 50                 | 100                |                    |        |        |                  |
| εSM1,2     | GAATGGGAGAGatGATATCATTTTGAAG | >20,000            | 8                  | 15                 | -      |        |                  |
| εSM2,3     | GAATGGGAGAGatTGGATATTTTTGAAG | >20,000            | 19                 | 14                 | -      |        |                  |
| GATC1      | AGCTTCGGGATAAGATCTTTAATTC | 5                  | 6                  | +                  | +      | 18     |                  |
| GATC2      | AGCTTCGaaATAAGATCTTTAATTC | 5                  | 20                 | +                  | -      |        |                  |
| GATC3      | AggttgcaGATCTTTATTTt      | >20,000            | 21                 | -                  | -      | 20     |                  |
| GATC4      | CGaaATAAAGATCTTTAATTC     | >20,000            | 30                 | -                  | +      | 31     |                  |
| GATC5      | AGCTTCGGGATAAGATCTTTAA    | 20                 | 34                 | +                  |        |        |                  |
| GATC6      | AGCggGGGGATAAGATCTTTAA    | 19                 | ++                 | +                  |        |        |                  |
| GATC7      | AGCTTCGGGATAAcATTTAA      | >20,000            | >20,000            | -                  |        |        |                  |
| mGATA pal  | AGTCCATCTGATAAGACTCTAGTTGCTGCC | 8                  |                    | +                  | +      |        |                  |
| mGATA palM1 | AGTCCcagTGATAAGACTTCAGTGCTGCC |                    |                    |                    |        |        |                  |
| cGATA palM1 | TCGCTcagAGATAAGGGGTCTCAGAGTG |                    |                    |                    |        |        |                  |
| β/ε        | AGCTTCGGTTTGCAGATAAACATTTAATCA |                    |                    |                    |        |        |                  |
GATA Peptides

Human GATA-1 Double Finger (hDF)

Chicken
GATA-2 N-finger (cGATA-2 NF)

This amino acid is a C in the native protein.
* This cysteine is an S in the peptide.
hGATA-1 DF Peptide

Free Probe

Fast Complex

Bound (10^{-8} M)

B/F

50
40
30
20
10
0

2.5 5 7.5 10 12.5 15

Bound (10^{-8} M)

K_d nM r

GATC1 2 0.93

cGATApalM1 14 0.8

AGCTTCGGGATAAAGATCTTAAATTC  TCGCTcagAGATAAGGCCTTCAGAGTGC

Downloaded from http://www.jbc.org/ by guest on July 24, 2018
A. GATA-1 Double Finger peptide, relative concentration

\[
\begin{align*}
GATC1 &\quad AGCTTCGGGATAAAGATCTTAAATTC & 5 \text{ nM} \\
CA1 &\quad AGCTTCGGGATAAAGAT\textcolor{red}{a}TTAATTC & 13 \text{ nM} \\
CA2 &\quad AGCTTCGA\textcolor{red}{a}AAATAAAGAT\textcolor{red}{a}TTAATTC & 22 \text{ nM}
\end{align*}
\]

B.

\[
\begin{align*}
\text{CF peptide} & & \text{DF peptide} \\
\text{Competitor} &\quad \beta/\varepsilon \text{ enhancer} &\quad GATC1 \\
\text{Competitor} &\quad \beta/\varepsilon \text{ enhancer} &\quad GATC1
\end{align*}
\]
A.  

5' AGTCCATCTGATAAGACTTATCTGCTGCCC 3'
A.

\[ \text{p}\alpha\text{D3} \]

\[ \text{GATCCAAGGATAAG} \]

\[ \text{GATCCCGCAGATAAAGCATG} \]

\[ \text{p}\alpha\text{D4/5} \]

B.

\text{GATApal}

\[ \text{p}\alpha\text{D4} \]

\[ \text{GATC1} \]
C.

Consensus:

GATApal ATCA/TGATAAG 8 nM

Reporter Plasmid GATA Sites

pαD3 GGATAAGATAAG 22 nM

pαD3 GP1 & 2 GGATCAGATAAG 8 nM
A.

Wild Type cGATA-1
NF cGATA-1 (HB2)

B.

Transactivation

- pD3
- pD4
- pDCA1-1
- pDCA1-2
- pDCA1-4
- pDCA2-1
- pDCA2-2
- pDCA2-4

Transactivation graph showing levels of expression for different conditions.
