Published in final edited form as:

Biochem Biophys Res Commun. 2008 May 16; 369(4): 1052–1056. doi:10.1016/j.bbrc.2008.02.136.

Substitution of DNA-Contacting Amino Acids with Functional Variants in the Gata-1 Zinc Finger: A Structurally and Phylogenetically Guided Mutagenesis

Tyson R. Vonderfecht¹, Daniel L. Schroyer¹, Brandy L. Schenck¹, Virginia M. McDonough², and Michael J. Pikaart^{1,3}

¹Department of Chemistry, Hope College, Holland, Michigan 49423

²Department of Biology, Hope College, Holland, Michigan 49423

Abstract

DNA binding functionality among transcription factor proteins is afforded by a number of structural motifs, such as the helix-turn-helix, helix-loop-helix, and zinc finger domains. The common thread among these diverse structures is their sequence-specific binding to essential promoter or other genetic regulatory sequences with high selectivity and affinity. One such motif, present in a wide range of organisms from bacteria to vertebrates, is the Gata-type zinc finger. This family of DNA-binding proteins is characterized by the presence of one or two (Cys)₄ metal binding sites which recognize the protein's eponymous binding site, GATA. Unlike other conserved DNA binding domains, Gata proteins appear to be restricted to binding consensus GATA sequences, or near variations, in DNA. Since the architecture of the Gata finger seems built around recognizing this particular sequence, we set out to define the allowable range of amino acid substitutions along the DNA-binding surface of a Gata finger that could continue to support sequence specific DNA binding activity. Accordingly, we set up a one-hybrid screen in yeast based on the chicken Gata-1 C-terminal zinc finger. Mutant libraries were generated at five amino acids identified in the Gata-DNA structure as likely to mediate sequence-specific contacts between the Gata finger and DNA. These libraries were designed to give as exhaustive amino acid coverage as possible such that almost all alternative amino acids were screened at each of the five probed positions. Screening and characterization of these libraries revealed several functional amino acid substitutions at two leucines which contact the DNA at the 3' and 5' flanks of the GATA binding site, but no functional substituents for amino acids near the core of the binding site. This pattern is consistent with amino acid sequences of known DNA-binding Gata fingers.

INTRODUCTION

The recruitment and binding of a limited set of specific transcription factors is an early step in the commitment of bone marrow stem cells to differentiate into functional blood cells during hematopoiesis. Among the specific transcription factors active during blood cell development are those of the Gata family, so named because they bind to 5'-WGATAR-3' (W = T or A; R = G or A) sequences in DNA. These sequences are found in abundance in the regulatory sequences of the genes encoding the globin proteins, which must eventually be produced in

³Corresponding author; Telephone: 616-395-7382, Fax: 616-395-7118, Email: pikaart@hope.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

great amounts in red blood cells [1]. Gata proteins contain two adjacent zinc finger domains of conserved amino acid sequence with the C-terminal Gata zinc finger (C-finger) sufficient to direct specific binding to the GATA binding motif in DNA. However, the N-terminal zinc finger (N-finger) has also been shown to possess its own DNA binding ability, with a subtly altered specificity in that it prefers a base other than A in the fourth position of GATA [2]. In its normal full-length protein context, the N-finger has been shown to participate in driving high affinity interaction at palindromic GATA sites and in transcriptional activation [3,4]. In addition to directing DNA binding, the Gata zinc fingers have been shown to be involved in protein-protein recognition as well, such as with the Gata cofactor FOG [5].

Several structural determinations of Gata Zn finger /DNA complexes by NMR have been published, two on the chicken Gata-1 C-finger [6,7] and one of the *Aspergillus nidulans* AreA zinc finger, in a comparison of the wild-type and a mutant peptide ([8,9]. These three structures are generally similar, revealing multiple contacts from amino acid side chains in close proximity to the GATA DNA major and minor grooves (Figure 1). In particular, five amino acids impose side chain atoms very close ($\leq 2 \text{ \AA}$) to those of the nucleotide base pairs. Referring to them according to the numbering of the cGata-1 C finger, these amino acids are Leu 17 and Arg 19 from the top of the fingertip, and Asn 29, Leu 33, and Leu 37 lying within the several turns of α -helix which extends C-terminal from the zinc finger. In order to explore the range of allowed amino acid substitutions that support both proper protein folding and DNA binding of the Gata finger, we have performed a site-directed mutagenesis survey of the amino acids noted above to complement our knowledge from existing evolutionary sequence variations within functional Gata-1 finger sequences.

MATERIALS AND METHODS

Construction of reporter plasmid and cGata-1 expression vector

A lacZ reporter plasmid was constructed by inserting an oligonucleotide containing a repeat of four GATA sites into the promoter of a pLacZi reporter plasmid (Clontech). In addition to this GATA molecule, variations which contained AATA, GGTA, and AGTA were also prepared. The plasmids were transfected into *S. cerevisiae* strain W303 by integration into the Ura3 locus, and correct integration was confirmed by Southern blotting of isolated DNA (data not shown). The chicken cGata-1 C-finger and its basic tail [6] was amplified by PCR and inserted in-frame into pGAD424 (Clontech) to create pGADgata, an expression vector that creates a GAL4 activation domain/C-finger fusion protein. The C-finger region in the resultant plasmid was verified by DNA sequencing. Expression of the wild-type Gata-1/Gal activation domain fusion gave robust β -gal expression from the GATA-containing promoter but not from the scrambled GATA controls.

Construction of mutant plasmid libraries

The library oligonucleotides were designed using overlapping oligonucleotides with degeneracies at the individual codon to be screened, which were ligated into pGADgata. Each plasmid library was amplified by transfection into competent *E. coli* cells, yielding approximately 10^5 colonies per library. Following library amplification and plasmid purification, the presence of degeneracy at the desired nucleotide positions was confirmed by sequencing through the mutant regions.

Screening of mutant plasmid libraries

Yeast containing the integrated Gata-1 activated lacZ reporter plasmid were electroporated with the mutant plasmid libraries and grown on minimal medium plates. After overnight growth at 30° C, colonies were patched onto fresh plates and assayed for β -galactosidase activity with 5-Bromo-4-chloro-3-indolyl β -D-galactopyranoside (X-gal) according to standard methods

[10]. We estimated that screening 100 colonies in the Leu and Arg sublibraries and 120 colonies in the Asn library will assure with >95% probability that all possible amino acids will be represented in each library. Plasmids from X-gal positive colonies were sequenced to determine the mutation in C-finger and rescreened for β -galactosidase activity with *o*-nitrophenyl β -D-galactopyranoside (ONPG) according to standard methods [10] to confirm activity.

Measurement of Gata finger/DNA dissociation constants

Electrophoresis mobility shift assays to measure Gata DNA wild type and mutant interactions were performed as described earlier [11].

Transactivation studies

Activation of a globin promoter by full-length cGata-1 and mutants was measured in a similar manner to that described earlier [12]. The method was updated in that, first, transfection was performed using Lipofectin, and second, the GATA sequence-containing α D6 promoter was subcloned out of the chloramphenicol acetyltransferase reporter plasmid used previously into a *Renilla* luciferase promoter (pGL2-Basic, Promega). Following electroporation, the two types of luciferase activities were assayed using Dual Luciferase Reporter chemistry (Promega) and measured in a luminometer (Turner).

RESULTS

One hundred colonies from each of the Leu 17, Arg 19, Leu 33, and Leu 37 sublibraries and 120 colonies in the Asn 29 library were screened and assayed for β -galactosidase activity with X-gal. Positive colonies were picked followed by reporter plasmid isolation and sequencing to determine the mutation in the C-finger (Table 1). Functional C-finger mutants had Tyr, Ile, or Met at the Leu 17 position, and Tyr or Thr at the Leu 37 position. The Arg 19, Asn 29, and Leu 33 positions yielded no functional alternate C-finger sequences. The assay for the Leu 37 sub-libraries was less clear-cut than the other mutation positions in that these libraries gave rise to many light blue colonies. As a result, colonies that gave blue colors significantly lighter than the blue colors of wildtype colonies were not counted. Sequencing of these light blue colonies showed that their mutant C-fingers contained Met, Val, Gly, Asn, Glu, Ser, His, or Lys at the Leu 37 position.

In order to compare for DNA binding ability directly among the various mutants, coding sequence for the C-fingers from the yeast one-hybrid plasmids were isolated from the yeast plasmids and recloned into glutathione S-transferase (GST) fusion expression vectors (pGEX6, Promega). GST fusion plasmids for each mutant were introduced into the appropriate bacterial strain for GST-Gata peptide production. Each mutant-containing bacterial strain was grown, induced to express its Gata peptide, and harvested for protein. GST-Gata peptides were purified over glutathione-agarose columns. Purified peptides were measured for concentration. Binding affinities were measured using quantitative EMSA to a fluorescently labeled oligonucleotide containing a consensus GATA binding site (Figure 2). All six Gata mutants identified as ONPG positives gave binding comparable to that of wildtype Gata.

Finally, as a biological test of functionality, the Leu 17 and 37 mutants were tested for ability to direct transcriptional activation. Yeast one-hybrid activity by growth in liquid culture and β -galactosidase assay by ONPG, which were consistent with the agar/X-gal assays, was correlated with transactivation assays in cell when introduced into a full-length chicken Gata-1 cDNA in an RSV promoter-driven expression vector. These full-length Gata expression vectors were cotransfected into QT6 cells along with a firefly luciferase reporter gene under the control of a globin promoter with multiple GATA binding sites. As a transfection control, a renilla luciferase gene driven by the tk promoter was also included. Two days following transfection,

cells were harvested and tested for relative firefly and *Renilla* luciferase activities as indicators of full-length Gata-1 transactivation potential. As shown (Figure 3), activation was quite consistent among the wildtype Gata-1 and the mutants with the exceptions of, first, a Leu17 to Ser mutation included as a negative control, and the Leu37 to Tyr. Thus, five of the six mutations identified as functional from the yeast one-hybrid screen exhibited completely normal transactivation potential in the context of full-length protein.

DISCUSSION

The ability of transcription factors such as Gata-1 to stimulate gene expression requires that they bind tightly and exclusively to their cognate binding sites at the regulatory loci of the genes to be regulated. A base-pair specific interaction between DNA and protein must by definition include a mechanism whereby the protein is able to discriminate a series of DNA base pairs of a given composition from all other possible base pair combinations. Several families of structural platforms have evolved to provide DNA binding domain functionality, including the basic (zipper and helix-loop-helix) domain, the helix-turn-helix domain, and the zinc-binding domain [13]. The zinc binding group includes the canonical zinc finger proteins, whose members are characterized by possession of one or more tetrahedral zinc coordination sites formed by four cysteine and/or histidine amino acid side chains. Gata proteins represent a subfamily characterized by a CXXC(X)₁₇₋₁₉CXXC metal binding domain and subsequent basic region (Figure 5). There are currently several hundred known Gata family members: a review on the evolutionary relationships amongst the Gatas included ninety-three [14]; the SBASE database [15] includes 328, while the EMBL-maintained SMART database [16] counts 462. Many of these sequences are of unknown function and unconfirmed (even unlikely) ability to bind DNA. Aside from the four cysteines, the only absolutely conserved amino acid is the tryptophan at position 18, presumably needed for the proper folding of the zinc fingertip. Other amino acids that are widely conserved, or substituted conservatively, are those involved in non-specific DNA interactions, such as the aromatic residue 34. Of particular relevance to the mechanism of sequence specificity are those amino acids most likely responsible for sequence-specific DNA contacts as discussed above for cGata-1 and AreA: Leu 17, Arg 19, Asn 29, and Leu 33 and 37. These are generally conserved in most of the classical Gatas, such as the vertebrate Gata proteins in hematopoietic and endodermal tissues and the fungal proteins such as AreA [17–19]. An interesting exception is a recently discovered gain-of-function mutation in the human Gata-2 C-finger of the amino acid referred to here as Leu17, to valine in a screen of patients with chronic myeloid leukemia [20]. Farther afield, several more exotic Gata family members have been found in which most or all of these five amino acids are different, such as the MTA1 (metastasis-associated) from rat and human [21,22], and the *C. elegans* Egl-27 protein [23,24]. MTA1 and Egl-27 have been shown to be components of chromatin remodeling complexes (NuRD) and as such may not necessarily be direct DNA binding proteins [25,26], but the nonconservative nature of the amino acid substitutions at the positions of Leu 17, Arg 19, Asn 29, Leu 33 and Leu 37 suggests that these amino acids side chains are not critical for folding of the finger *per se*.

The sequences of several fungal proteins which are known to bind specifically to DNA are instructive in demonstrating the possibility for at least some sequence variation at the key contacts in the context of specific binding. These include the *Neurospora* whitecollar-1 and -2 (WC-1 and WC-2) proteins, transcription factors which regulate gene expression in response to blue light, and the cutinase promoter Palindrome Binding Protein (PBP) in the plant pathogenic fungus *Fusarium solani* [27–29]. The sequences of WC1 and 2 and PBP (Figure 4) indicate the ability of chemically very divergent functionalities at two of cGata-1's DNA-contacting leucines to support DNA binding, including an acidic residue at position 17 and a basic one at position 37.

In addition to these Gata homologs, a few studies have characterized spontaneous or induced mutations of the *Aspergillus* AreA Gata protein, using growth on various nitrogen sources as a selection/screening phenotype [30]. That work found a widespread flexibility at the Leu 17 position, with Val, Thr, Ile, Asn, Ala, and Met, and Leu 37, with Phe, Val, and Ser all providing functional variants. At the Leu 33 position, substitution with Cys was reported as functional while substitution by Phe was nonfunctional. For Arg 19 and Asn 29, no functional substitutions were found while Leu or Glu in place of Arg 19 and Ser for Asn 29 were nonfunctional. Finally, further mutagenesis results come from a study on the *Ustilago maydis* Urbs1 protein in which an Arg19 to Leu mutation resulted in loss of function [31].

In sum, the phylogenetic surveys are in good agreement with the mutagenesis results presented here in that the protein/DNA interaction in the Gata-1 protein tolerates a wide range of chemistry at the flanks of the binding site; that is, the leucine at position 17 according the chicken Gata-1 C-finger structure, which lies outside the base G in the GATA consensus binding site, and the leucine at position 37, lying downward from the terminal A base in GATA (Figure 1). The interaction between the Gata zinc finger and DNA has been described as hydrophobic in nature due to the presence of the leucine amino acid side chains from the protein and thymine methyl groups in the DNA [32]. While isoleucine provided a functional replacement for leucines at both positions 17 and 37 and methionine at 17, each also tolerated at least one polar amino acid as well. Considered along with our earlier work showing that the thymine methyl groups are much less important than polar moieties on guanine and adenine bases in directing binding stability [11], it would seem more accurate to characterize the Gata-DNA interaction as including a wide range of van der Waals interactions in addition to uniquely hydrophobic ones.

ACKNOWLEDGMENTS

The authors acknowledge support for this work from the Howard Hughes Medical Institute (Faculty Development Grant for Interdisciplinary Research; VMM and MJP) and from the National Institutes of Health (R15 HL066009-02; MJP). MJP also thanks Dr. Mary Ann Osley and colleagues at the University of New Mexico for their hospitality over a sabbatical stay during which a portion of this work was completed.

REFERENCES

1. Orkin SH. GATA-binding transcription factors in hematopoietic cells. *Blood* 1992;80:575–581. [PubMed: 1638017]
2. Pedone PV, Omichinski JG, Nony P, et al. The N-terminal fingers of chicken GATA-2 and GATA-3 are independent sequence-specific DNA binding domains. *Embo J* 1997;16:2874–2882. [PubMed: 9184231]
3. Trainor CD, Omichinski JG, Vandergon TL, et al. A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA-binding domain for high-affinity interaction. *Mol Cell Biol* 1996;16:2238–2247. [PubMed: 8628290]
4. Trainor CD, Ghirlando R, Simpson MA. GATA zinc finger interactions modulate DNA binding and transactivation. *J Biol Chem* 2000;275:28157–28166. [PubMed: 10862757]
5. Tsang AP, Visvader JE, Turner CA, et al. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 1997;90:109–119. [PubMed: 9230307]
6. Omichinski JG, Trainor C, Evans T, et al. A small single-finger peptide from the erythroid transcription factor GATA-1 binds specifically to DNA as a zinc or iron complex. *Proc Natl Acad Sci U S A* 1993;90:1676–1680. [PubMed: 8446581]
7. Tjandra N, Omichinski JG, Gronenborn AM, et al. Use of dipolar ^1H - ^{15}N and ^1H - ^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat Struct Biol* 1997;4:732–738. [PubMed: 9303001]

8. Starich MR, Wikstrom M, Schumacher S, et al. The solution structure of the Leu22-->Val mutant AREA DNA binding domain complexed with a TGATAG core element defines a role for hydrophobic packing in the determination of specificity. *J Mol Biol* 1998;277:621–634. [PubMed: 9533884]
9. Starich MR, Wikstrom M, Arst HN Jr, et al. The solution structure of a fungal AREA protein-DNA complex: an alternative binding mode for the basic carboxyl tail of GATA factors. *J Mol Biol* 1998;277:605–620. [PubMed: 9533883]
10. Rose, MD.; Winston, F.; Hieter, P. *Methods in Yeast Genetics: A Laboratory Course Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1990.
11. Mott BH, Bassman J, Pikaart MJ. A molecular dissection of the interaction between the transcription factor Gata-1 zinc finger and DNA. *Biochem Biophys Res Commun* 2004;316:910–917. [PubMed: 15033488]
12. Evans T, Felsenfeld G. trans-Activation of a globin promoter in nonerythroid cells. *Mol Cell Biol* 1991;11:843–853. [PubMed: 1990287]
13. Matys V, Fricke E, Geffers R, et al. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 2003;31:374–378. [PubMed: 12520026]
14. Lowry JA, Atchley WR. Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain. *J Mol Evol* 2000;50:103–115. [PubMed: 10684344]
15. Vlahovicek K, Murvai J, Barta E, et al. The SBASE protein domain library, release 9.0: an online resource for protein domain identification. *Nucl. Acids Res* 2002;30:273–275. [PubMed: 11752313]
16. Letunic I, Copley RR, Pils B, et al. SMART 5: domains in the context of genomes and networks. *Nucl. Acids Res* 2006;34:D257–D260. [PubMed: 16381859]
17. Charron F, Nemer M. GATA transcription factors and cardiac development. *Semin Cell Dev Biol* 1999;10:85–91. [PubMed: 10355032]
18. Scazzocchio C. The fungal GATA factors. *Curr Opin Microbiol* 2000;3:126–131. [PubMed: 10745000]
19. Weiss MJ, Orkin SH. GATA transcription factors: key regulators of hematopoiesis. *Exp Hematol* 1995;23:99–107. [PubMed: 7828675]
20. Zhang SJ, Ma LY, Huang QH, et al. Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci U S A*. 2008
21. Toh Y, Kuninaka S, Endo K, et al. Molecular analysis of a candidate metastasis-associated gene, MTA1: possible interaction with histone deacetylase 1. *J Exp Clin Cancer Res* 2000;19:105–111. [PubMed: 10840944]
22. Toh Y, Pencil SD, Nicolson GL. A novel candidate metastasis-associated gene, mta1, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. *J Biol Chem* 1994;269:22958–22963. [PubMed: 8083195]
23. Ch'ng Q, Kenyon C. egl-27 generates anteroposterior patterns of cell fusion in *C. elegans* by regulating Hox gene expression and Hox protein function. *Development* 1999;126:3303–3312. [PubMed: 10393110]
24. Herman MA, Ch'ng Q, Hettenbach SM, et al. EGL-27 is similar to a metastasis-associated factor and controls cell polarity and cell migration in *C. elegans*. *Development* 1999;126:1055–1064. [PubMed: 9927605]
25. Solari F, Bateman A, Ahringer J. The *Caenorhabditis elegans* genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. *Development* 1999;126:2483–2494. [PubMed: 10226007]
26. Xue Y, Wong J, Moreno GT, et al. NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol Cell* 1998;2:851–861. [PubMed: 9885572]
27. Ballario P, Vittorioso P, Magrelli A, et al. White collar-1, a central regulator of blue light responses in *Neurospora*, is a zinc finger protein. *Embo J* 1996;15:1650–1657. [PubMed: 8612589]
28. Linden H, Macino G. White collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in *Neurospora crassa*. *Embo J* 1997;16:98–109. [PubMed: 9009271]
29. Carattoli A, Cogoni C, Morelli G, et al. Molecular characterization of upstream regulatory sequences controlling the photoinduced expression of the albino-3 gene of *Neurospora crassa*. *Mol Microbiol* 1994;13:787–795. [PubMed: 7815938]

30. Wilson RA, Arst HN Jr. Mutational analysis of AREA, a transcriptional activator mediating nitrogen metabolite repression in *Aspergillus nidulans* and a member of the "streetwise" GATA family of transcription factors. *Microbiol Mol Biol Rev* 1998;62:586–596. [PubMed: 9729601]
31. An Z, Zhao Q, McEvoy J, et al. The second finger of Urbs1 is required for iron-mediated repression of *sid1* in *Ustilago maydis*. *Proc Natl Acad Sci U S A* 1997;94:5882–5887. [PubMed: 9159169]
32. Omichinski JG, Clore GM, Schaad O, et al. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* 1993;261:438–446. [PubMed: 8332909]

Figure 1. Structure of the C-terminal chicken Gata-1 zinc finger bound to DNA (from PDB ID 2GAT, [7]), highlight in dark grey the five amino acid side chains in closest contact with the DNA major groove along the helical axis (at right) and perpendicular to it (left). The DNA sequence GATA runs from top to bottom on the frontmost strand in the figure at right.

Figure 2.

In vitro DNA binding of purified mutant Gata finger peptides. The Gata finger cDNA sequence from positive LacZ clones was removed from the Gal activation plasmid and subcloned into the bacterial expression plasmid pGEX6, resulting in fusion in frame to a glutathione S-transferase (GST) domain to facilitate purification. GST/Gata fusion proteins were expressed in bacterial culture and purified over glutathione chromatography columns. Electrophoretic mobility shift assays were performed by titration with fluorescently-labelled oligonucleotide containing a GATA binding site (lefthand column). Ratios of bound and unbound oligonucleotide were plotted (center column) and fit to single-site binding isotherms to determine dissociation constants (K_D) for each Gata mutant, shown at right (+/- standard error).

Binding affinity is decreased, and complex size increased, for all mutants and wildtype Gata compared to previously reported values due to the presence of the GST domain.

Figure 3.

Transactivation potential of Gata C-finger mutants. **A.** Positive colonies from the one-hybrid mutant screen as judged by production of blue color on X-gal were re-screened using ONPG. The unmodified Gata finger/Gal activation plasmid was included as a positive control, denoted wt. The N29R (shaded bar) is a clone of an X-gal negative (white) colony and is included as a negative control. **B.** Positive finger mutants were introduced into the full length chicken Gata-1 cDNA in an expression plasmid driven by the RSV promoter. These plasmids were cotransfected in avian QT6 fibroblasts along with a reporter plasmid carrying the luciferase gene linked to the chicken α -globin promoter, a Gata-responsive promoter. Cell lysates were assayed for luciferase activity and are shown relative the transcriptional activation driven by the wildtype full length Gata. The L17S (shaded bar) mutant was another inactive clone from the one-hybrid screen and is shown as a negative control, as well luciferase transfected without the Gata expression plasmid ("no gata"). All mutants identified as positive in the one-hybrid screen showed transactivation potential similar to that of the wildtype protein, with the exception of the L37Y mutant.

5 10 15 20 25 30 35
 Chicken Gata1 K R A G V T C S N C E T S T T T L W R --- R S P M G A - P V C N A C G L Y Y K L H Q V N R
 Human Gata1 K R A G 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
 Aspergillus AreA Q N G P 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
 Neurospora WC1 G N M V R D C A N C H T R N T P E W R --- R G P S G N R D L C N S C G L R W A K Q T G R V S
 Neurospora WC2 V A E E Y V C T D C G T L C S P E W R --- K G P S G P K T L C N A C G L R W A K K E K K N
 Fusarium PBP T S E E Y V C T D C G T L D S P E W R --- K G P S G P K T L C N A C G L R W A K K E K K R N
 Human MTA1 P G A G R A C E S C Y T T Q S Y Q W Y S W G P P N M Q C - R L C A S C W T Y W K K Y G G L K
 C elegans Egl27 G P S G R A C H H C Y G A E S K D W H --- H A N G L L - - L C T D C R L H Y K K Y G Q L R
 60% consensus p..sppC*ss*C.TspT*sb*WR...+sssG.pHLCNACGL*aa*+p*ph*..

Figure 4.

Zinc finger sequences from a select number of Gata family homologs. The human and chicken C-terminal fingers in the upper two rows illustrate the degree of sequence conservation present among vertebrate Gata's. The sequences shown in the four middle rows display the wider range of sequences among fungal DNA binding single-finger Gata proteins. The lower two sequences, the vertebrate Metastasis Associated protein 1 (part of the NuRD remodeling complex) and its invertebrate homolog Egl27, are presented to include proteins which do not appear to function in a sequence-specific DNA binding role but nonetheless contain the characteristic Gata amino acid sequence.

Table 1

Appearance of active DNA binding mutants in the Gata one-hybrid screen. Numerous blue colonies were obtained from the Leu 17 and 37 libraries. Upon recovery of the Gata plasmids and sequencing, these clones were found to contain mutations encoding the indicated amino acid(s). No positive colonies were obtained for the Arg19 and Leu 33 libraries. Two blue positives were recovered from the Asn 29 library, both of which contained asparagine codons; this was possible at this position because this library was constructed in a way that did not exclude asparagine.

Library	Colonies screened	Number of X-gal positives	Mutation
Leu 17 RNK	100	10	Ile, Met
Leu 17 NVK	100	6	Tyr
Arg 19 KNK	100	0	-
Arg 19 NHK	100	0	-
Asn 29 NNK	120	2	(Asn)
Leu 33 RNK	100	0	-
Leu 33 NVK	100	0	-
Leu 37 RNK	100	27	Ile
Leu 37 NVK	100	12	Tyr, Thr