THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 21, Issue of May 22, pp. 13273–13279, 1998
Printed in U.S.A.

A Conserved C-terminal Domain in PBX Increases DNA Binding by the PBX Homeodomain and Is Not a Primary Site of Contact for the YPWM Motif of HOXA1*

(Received for publication, October 31, 1997, and in revised form, March 3, 1998)

Nancy C. Green‡, Isabel Rambaldi‡, Joseph Teakles‡§, and Mark S. Featherstone‡§¶‖

From the ‡McGill Cancer Centre, and Departments of §Medicine (Division of Experimental Medicine) and ¶Oncology, McGill University, Montréal, Québec H3G 1Y6, Canada

HOX proteins are dependent upon cofactors of the PBX family for specificity of DNA binding. Two regions that have been implicated in HOX/PBX cooperative interactions are the YPWM motif, found N-terminal to the HOX homeodomain, and the GKFQ domain (also known as the Hox cooperativity motif) immediately C-terminal to the PBX homeodomain. Using derivatives of the E2A-PBX oncoprotein, we find that the GKFQ domain is not essential for cooperative interaction with HOXA1 but contributes to the stability of the complex. By contrast, the YPWM motif is strictly required for cooperative interactions *in vitro* and *in vivo*, even with mutants of E2A-PBX lacking the GKFQ domain. Using truncated PBX proteins, we show that the YPWM motif contacts the PBX homeodomain. The presence of the GKFQ domain increases monomer binding by the PBX homeodomain 5-fold, and the stability of the HOXA1·E2A-PBX complex 2-fold. These data suggest that the GKFQ domain acts mainly to increase DNA binding by PBX, rather than providing a primary contact site for the YPWM motif of HOXA1. We have identified 2 residues, Glu-301 and Tyr-305, required for GKFQ function and suggest that this is dependent on α-helical character.

Hox genes are widely conserved regulators of anteroposterior patterning during animal development (1). The 39 mammalian *Hox* genes are expressed in overlapping domains and confer positional identity along the body axes (1, 2). The *Hox* genes encode homeodomain-containing transcription factors that bind a degenerate site with a TAAT core (3). Whereas modest binding-site preferences exist, these appear insufficient to account for specificity of action *in vivo*. The PBC (PBX and CEH-20) subfamily of homeodomain proteins (4), including PBX (pre-B-cell transformation-related gene) (5) and EXD (extradenticle) (6), are cofactors that greatly increase the specificity of HOX proteins. *PBX* is the mammalian homolog of *Drosophila exd*. The latter was first identified in a genetic screen for patterning defects in the fly (6, 7). The homeodomain proteins encoded by *PBC* genes cooperatively interact with HOX proteins to bind an extended site on DNA and regulate transcription *in vivo* (8–17). PBC proteins improve HOX specificity due to the increased size of the cooperative binding site and the strength of DNA binding and by modulating recognition of cooperative binding sites by different groups of HOX proteins (15, 16, 18–24).

The human *PBX* family comprises three genes, *PBX1*, *PBX2*, and *PBX3* (25). *PBX1* was initially discovered through the study of the t(1:19) chromosomal translocation causing 25% of all acute pre-B cell leukemia (26, 27). This translocation fuses the powerful transcriptional activation domains of E2A to the homeodomain and C terminus of PBX1, creating the novel transcription factor, E2A-PBX. The oncogenicity of E2A-PBX has been demonstrated in focus-forming assays, myeloid differentiation assays, and in mice (5, 28, 29). Studies done to uncover the regions of E2A-PBX required for transformation have identified the transcriptional activation domains as essential (30, 31). However, the homeodomain was found to be dispensable for oncogenic transformation in focus-forming assays and transgenic mice (30, 31), although it is required to block myeloid differentiation (31). This raised the question of how E2A-PBX could alter target gene expression without a DNA-binding domain.

HOX proteins could interact with homeodomain-containing E2A-PBX and localize it to regulatory sites on DNA. This supposes that E2A-PBX would cause transformation by the misregulation of *Hox* gene targets. There are a number of arguments supporting this model. The leukemia induced by E2A-PBX is characterized by a block in differentiation and *Hox* genes are known to regulate hematopoietic differentiation (32). Misexpression of *Hox* genes is known to cause leukemia (33–36). There is also evidence that the combined misexpression of *Hox* genes and the PBX-related gene *Meis1* causes leukemia in BXH-2 mice (37).

It is therefore of interest to examine closely the interactions between HOX proteins and E2A-PBX to better understand how PBX contributes to HOX specificity and to assess the above model of E2A-PBX transformation. Two regions outside the homeodomains have been implicated in HOX/PBX interactions. The YPWM motif (and variants thereof) is found in many HOX proteins N-terminal to the homeodomain (38). This motif is required for any significant interaction with PBX or EXD (11, 14, 17, 20, 24, 39–42). In PBX, a conserved region immediately following the C terminus of the homeodomain has also been shown to influence cooperative interactions (40, 43–45) and has been designated the Hox cooperativity motif (HCM) (44). The HCM, referred to as the GKFQ domain in this paper, has been proposed as a possible contact site for the YPWM motif (21, 40, 44).

Here we report that the GKFQ domain of PBX is not essential to form a cooperative complex with HOXA1 but does contribute to complex stability. We have mapped important residues of the GKFQ domain and shown that its function is likely

* This work was supported in part by funds from the Cancer Research Society Inc. and a Chercheur-Boursier award from the Fonds de la Recherche en Santé du Québec (to M. S. F.) and an internal postdoctoral award from a Center Grant from the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (to N. C. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed: McGill University, 3655 Drummond St., Montréal, Québec H3G 1Y6, Canada. Tel.: 514-398-8937; Fax: 514-398-6769; E-mail: featherstone@medcor.mcgill.ca.

to depend on an α-helical structure. We further demonstrate that the primary YPWM contact site is located in the PBX homeodomain and not in the GKFQ domain. The GKFQ domain does, however, increase monomer PBX binding. We therefore conclude that the GKFQ domain contributes to many HOX/PBX interactions by stabilizing contacts with DNA. A secondary role in interprotein interactions is not excluded and may be more important for interaction with a subset of HOX proteins including HOXB7.

EXPERIMENTAL PROCEDURES

Plasmid Construction—The E2A-PBX mutants were constructed by two different approaches. EP ΔGKFQ was generated by polymerase chain reaction mutagenesis so as to create two new restriction sites: an *Eco*RI site at base pair 863 (GAATCC to GAATTC) of the PBX1A open reading frame, resulting in a single silent change, and a *Hin*dIII site at base pair 926 (CAGCTG to AAGCTT). The latter introduced two codon changes: T309K and V311F. The E2A-PBX mutants A302P, F298P, and N303A/I304A/Y305A were generated by introducing a 68-mer double-stranded oligonucleotide into these new restriction sites. EP pswt,[1] which reintroduced wild-type E2A-PBX sequence while retaining the new restriction sites, was also produced in this manner. All other mutants were subsequently generated by site-directed mutagenesis (Sculptor Mutagenesis System, Amersham Pharmacia Biotech) in M13. All E2A-PBX derivatives were cloned in the pSG5 expression vector (46) (Stratagene) allowing both *in vitro* transcription/translation and expression in mammalian cells. PBX 233–319 and 233–294 were constructed by polymerase chain reaction amplification using primers that incorporated start and stop codons and *Nde*I and *Bam*HI sites to clone into pET16b (Novagen). All constructs were verified prior to use. The primers used were: 5′GGGAATTCCATATGGCGCGGCGGAAGAGAC-3′, 5′CCCAAGCTTGGATCCTCAATGGGCTGACACATTGG3′, and 5′CGGGATCCTCAGTTCTTCTTGTACCG3′.

Protein Expression and Purification—The truncated HOXA1 proteins used in these experiments (amino acids 185 to the C terminus) were expressed as His-tagged fusion proteins and batch-purified from *Escherichia coli* MC1061 using an imidazole elution protocol as described previously (20). They were stored at −80 °C in 200 mM Na_2PO_4 (pH 7.4), 500 mM NaCl, 10 mM β-mercaptoethanol, 0.01% Triton X-100, 300 mM imidazole, and 20% glycerol.

His-tagged PBX constructs were expressed in *E. coli* BL21(DE3)pLysS. Cultures were grown at 37 °C to an A_{600} of 0.9, induced with isopropyl-1-thio-β-D-galactopyranoside to 1 mM and grown at 30 °C for 3 h. Cells were pelleted by centrifugation at 5,000 × *g* for 20 min, resuspended in binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9, supplemented with urea to 6 M), and sonicated. The extract was clarified by centrifugation at 12,000 × *g* for 20 min followed by filtration through a 0.45-μm filter. Protein was batch-purified by incubating the clarified extract with nickel-charged chelating Sepharose beads (Amersham Pharmacia Biotech) for 1 h at 4 °C followed by washing twice with 10 ml of binding buffer and three times with 10 ml of 1× washing buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9, supplemented with urea to 6 M). Protein was refolded by stepwise 1 M decreases in the urea concentration in the washing buffer with a 20-min equilibration period between each buffer change. Protein was eluted with 1× elution buffer (1 M imidazole, 0.5 M NaCl, 20 mM Tris-HCl, pH 7.9) and stored at −80 °C in 20% glycerol. Protein concentrations were determined by the Bradford assay (47).

The E2A-PBX proteins were produced using a TnT *in vitro* coupled transcription/translation kit (Promega). The amount of translated protein was quantified by [^{35}S]Met labeling.

Transient Transfection—Transient transfections were done in human embryonic kidney 293 cells as described previously for P19 cells (20). Human embryonic kidney 293 cells were cultured in α-minimal essential medium supplemented with 10% fetal calf serum (Life Technologies, Inc.) and antibiotics (penicillin and streptomycin) (Sigma).

Electrophoretic Mobility Shift Assay (EMSA), HOX Titrations, and Dissociation Rate Experiments—EMSA was done as described previously (11) with the following modifications: 8 ng of His-tagged HOXA1 and 2.0 μl of *in vitro* translated E2A-PBX derivatives or 16 ng of His-tagged PBX protein were used per 10 μl of reaction. Final buffer conditions were 10 mM Tris (pH 7.5), 75 mM NaCl, 30 mM imidazole, 1 mM dithiothreitol, 1 mM EDTA, 5.4 μg of bovine serum albumin, 12.7% glycerol, and 80 ng of poly(dI-dC)/10 μl. The following ^{32}P-labeled DNA probe was used (one strand shown): 5′ TCACCATGATTGATGGGC-GACTGCTCGG 3′. This contains the "G6" cooperative binding site 5′-TGATTGATGG-3′. Quantification of steady-state bands was determined using a Fuji Bas2000 Imager.

A range of HOX protein concentrations (0–200 nM) was incubated as described above either alone or with 30 ng of PBX 233–294. Quantification of bands was determined using a Fuji Bas2000 Imager. Titration data was plotted using the GraFit program (48) and fitted to the equation describing cooperative binding: $y = ([L]^n * \text{capacity})/(K_d + [L])^n$ where $[L]$ is the concentration of ligand and n is the cooperativity factor.

Dissociation rate experiments were performed by adding 450 ng/μl of unlabeled double-stranded competitor DNA (identical to the labeled DNA binding site) at time 0 following a 50 min preincubation period at room temperature of proteins with labeled DNA. Samples (7 μl) were loaded at 3, 10, 15, 20, 25, and 40 min. The fraction of labeled DNA bound at each time point was determined using a Fuji Bas2000 Imager. The log of this value was plotted against time to get a line of best fit where the slope is equivalent to the negative K_d, the dissociation rate constant. Half-lives were determined using the equation $t_{1/2} = -\log(0.5)/K_d$.

RESULTS

The GKFQ Domain Is Not Essential for Cooperative Interactions between HOXA1 and E2A-PBX Nor Is It the Primary Contact Site for the YPWM Motif—Thirteen residues immediately following the PBX homeodomain are highly conserved across species (Fig. 1*B*). We call this the GKFQ domain, after the first 4 residues. The GKFQ domain has been reported to be important for cooperative interactions between HOX and PBX (40, 43, 44, 49) and, based on comparison to MATA1/MATα2, has been predicted to adopt an α-helical structure (50–52). A deletion of the 13 conserved amino acids in this region was made to study its importance (Fig. 1*A*). In addition, we have conducted an alanine/proline scan through the GKFQ domain (Fig. 1*C*). Whereas most residues were mutated to alanine, those that are already alanine or that are alanine in one or more of the known homologs of PBX1, were instead mutated to proline. Proline was chosen since it should disrupt the putative α-helical structure (53). Two other mutations outside the GKFQ region were introduced in wild-type E2A-PBX to facilitate the creation of these mutants by the introduction of new restriction sites (see "Experimental Procedures"). This variant of wild-type, called EP pswt, contains the changes T309K and V311F. The change T309K is found naturally between PBX1 and EXD, whereas residue 311 lies outside of the most conserved region (Fig. 1*B*). There was no difference between the activity of EP pswt and wild-type E2A-PBX as measured by the dissociation rates of cooperative complexes (data not shown). In transfection experiments, transcriptional activation by EP pswt was approximately twice that of wild-type E2A-PBX (Fig. 3).

Initially, we wished to assess the overall importance of the GKFQ domain to HOX/E2A-PBX steady-state cooperative DNA binding and to assess the possibility that the GKFQ domain serves as the primary contact site for the YPWM motif (Fig. 2). The mutant EP ΔGKFQ (Fig. 1*A*) lacks the GKFQ domain due to an internal deletion of residues 296 through 308. Nonetheless, EP ΔGKFQ is still able to form a heterodimeric complex with HOXA1 (Fig. 2, compare *lanes 5* and *7*). The GKFQ domain is therefore not essential for HOXA1/E2A-PBX interactions. By contrast, mutation of the key tryptophan and methionine residues of the YPWM motif of HOXA1 (HOXA1 WM→AA) abolishes interactions with EP pswt and EP ΔGKFQ (Fig. 2, compare *lanes 6* and *8*). Together, these results strongly suggest that the GKFQ domain of PBX is not the primary site of contact of the HOX YPWM motif. Mutation of the PBX GKFQ domain or the HOX YPWM motif does not have equivalent consequences, as would be the case if these two domains were mutual primary contact sites.

[1] The abbreviations used are: EP, E2A-PBX1A; EP pswt, EP pseudo wild type; EMSA, electrophoretic mobility shift assay.

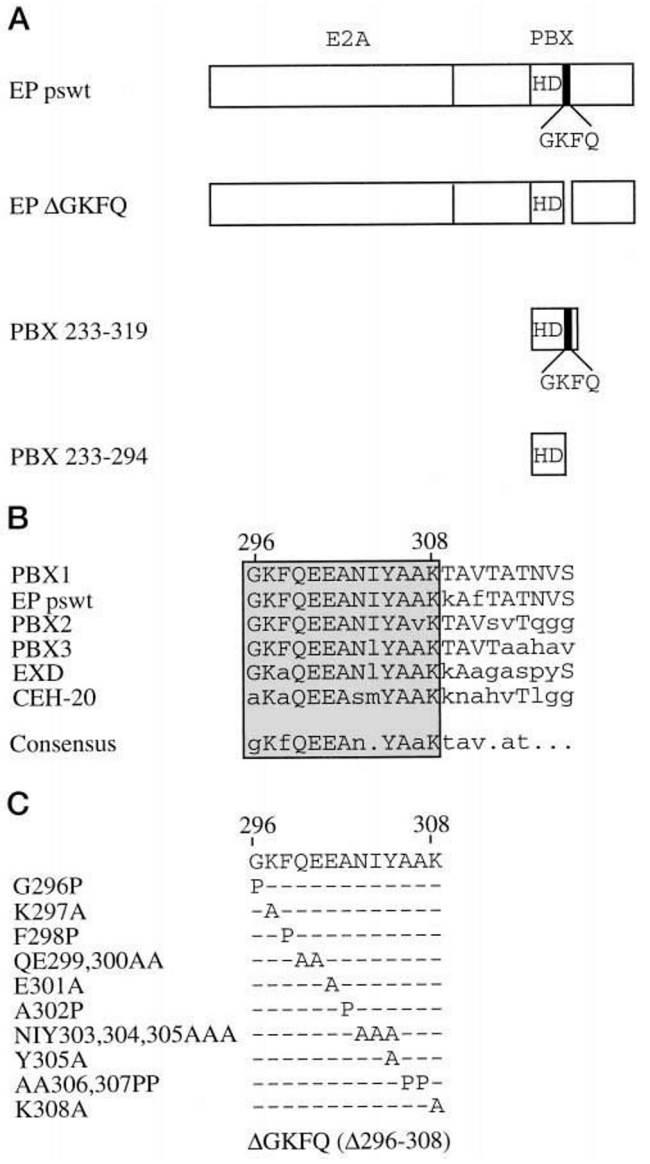

FIG. 1. **A description of proteins used in this study and the degree of GKFQ domain conservation among homologs.** *A,* derivatives of E2A-PBX1A and PBX1. E2A-PBX1A pseudo wild type (EP pswt, see "Experimental Procedures") was used as the parent for the alanine scan of the GKFQ domain (*black box*). The E2A and PBX1A halves of the oncoprotein are marked. HD designates the PBX homeodomain. EP ΔGKFQ is a deletion mutant of EP pswt lacking amino acids 296 through 308 (PBX1A numbering system is used throughout) (25). Two other truncated PBX proteins were constructed containing the isolated PBX homeodomain (PBX 233–294) or the PBX homeodomain with a short C-terminal extension including the GKFQ domain (PBX 233–319). *B,* the GKFQ domain of human PBX1 and its homologs including human PBX2 and PBX3 (25), *Drosophila* EXD (7), and the *Caenorhabditis elegans* CEH-20 (57). The residues in each homolog and EP pswt that differ from PBX1 are indicated in *lowercase characters*. A consensus is provided below where residues in *uppercase* are 80% identical and those in *lowercase* are at least 50% identical to PBX1. The *shaded area* contains those residues that were subjected to an alanine scan. *C,* residues of the GKFQ domain were generally mutated to alanine. However, those that were already alanine or were alanine in one of the homologs of PBX1, were mutated to proline. A deletion of the entire 13 amino acids from 296 through 308 (EP ΔGKFQ) was also constructed. All proteins shown in *C* are derived from EP pswt. PBX 233–319 and 233–294 are derived from wild-type PBX1.

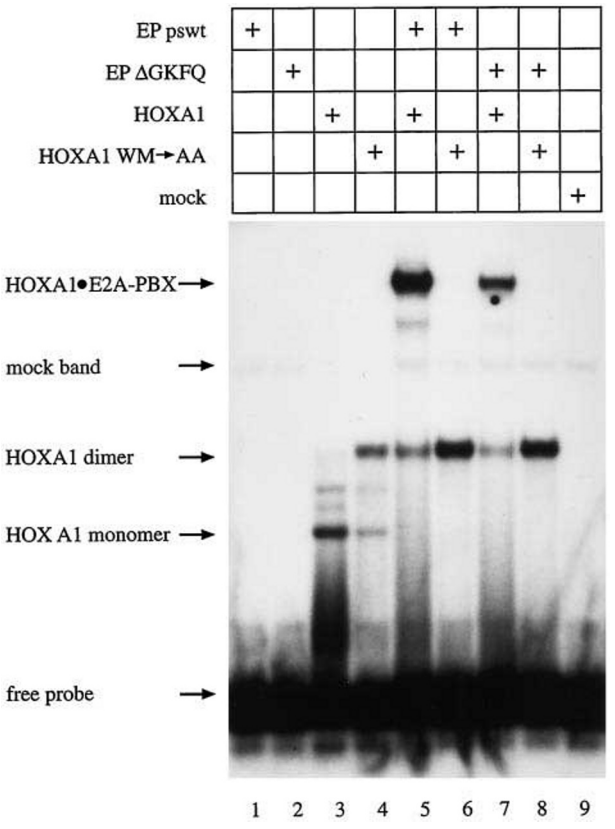

FIG. 2. **The GKFQ domain contributes, but is not essential to HOX/PBX interactions *in vitro*.** Mutations of the GKFQ domain (EP ΔGKFQ) and the YPWM motif of HOXA1 (HOXA1 WM→AA) were tested for cooperative interaction by EMSA. The HOXA1 proteins were bacterially expressed and purified, whereas the E2A-PBX proteins were *in vitro* translated. A nonspecific band derived from the reticulocyte lysate is indicated. This and all subsequent experiments were performed with a double-stranded radiolabeled probe having a G6 cooperative binding site (5′-TGATTGATGG-3′) (20).

Transient transfection experiments confirmed these results *in vivo*. Expression vectors for either EP pswt or the mutant EP ΔGKFQ were cotransfected with a luciferase reporter bearing a minimal promoter driven by cooperative binding sites for HOX·PBX complexes (pML(5XHOX/PBX)) (11). Comparable to our results in EMSA, deletion of the GKFQ domain from EP pswt reduced but did not abolish transcriptional activation (Fig. 3, compare *bars 3* and *4*). The EP ΔGKFQ mutant retains 18% of the activity of EP pswt and 40% of true E2A-PBX wild type, which was less active in this assay.

The Binding Site for the YPWM Motif Is in the PBX Homeodomain—Since the YPWM motif does not primarily interact with the GKFQ domain, the possibility that its major binding site resides in the PBX homeodomain was addressed using truncated recombinant proteins. PBX 233–319 contains the homeodomain plus 24 residues spanning the C-terminal GKFQ domain, whereas PBX 233–294 is the isolated homeodomain (Fig. 1*A*). The GKFQ domain was not essential to form a DNA-bound heterodimer with HOXA1 (Fig. 4*A*, compare *lanes 5* and *8*). Rather, the PBX homeodomain alone was able to interact with HOXA1 to form a heterodimeric complex bound to DNA. Moreover, this interaction was highly YPWM dependent (Fig. 4*A*, compare *lanes 8* and *9*), demonstrating that the primary binding site for the YPWM motif of HOXA1 is in the PBX homeodomain rather than the GKFQ domain.

Based on the percentage of probe bound, the HOXA1·PBX 233–294 heterodimeric complex is 90-fold more abundant than would be predicted by chance association of the monomeric constituents on the same fragment of DNA. This is strongly suggestive of cooperative DNA binding. To prove that the interaction between HOXA1 and PBX 233–294 was indeed cooperative, we first established a linear range of HOXA1 monomer

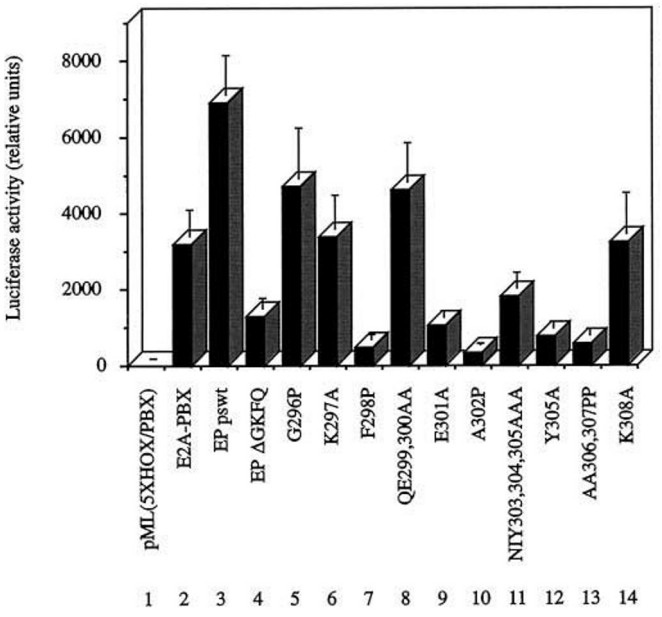

FIG. 3. **The GKFQ motif contributes, but is not essential to HOX/PBX interactions *in vivo*.** The histogram summarizes the results of transient transfection assays comparing the effect of deletion of the GKFQ domain (EP ΔGKFQ) of E2A-PBX and the alanine-proline scan on transcriptional activation. 1 μg of expression plasmid for the indicated proteins was transfected into HEK 293 cells along with carrier DNA, a luciferase reporter plasmid carrying five copies of a HOX·PBX cooperative binding site (pML(5XHOX/PBX)) and an internal *lacZ* control. Standard error of the mean is shown, $n = 3$.

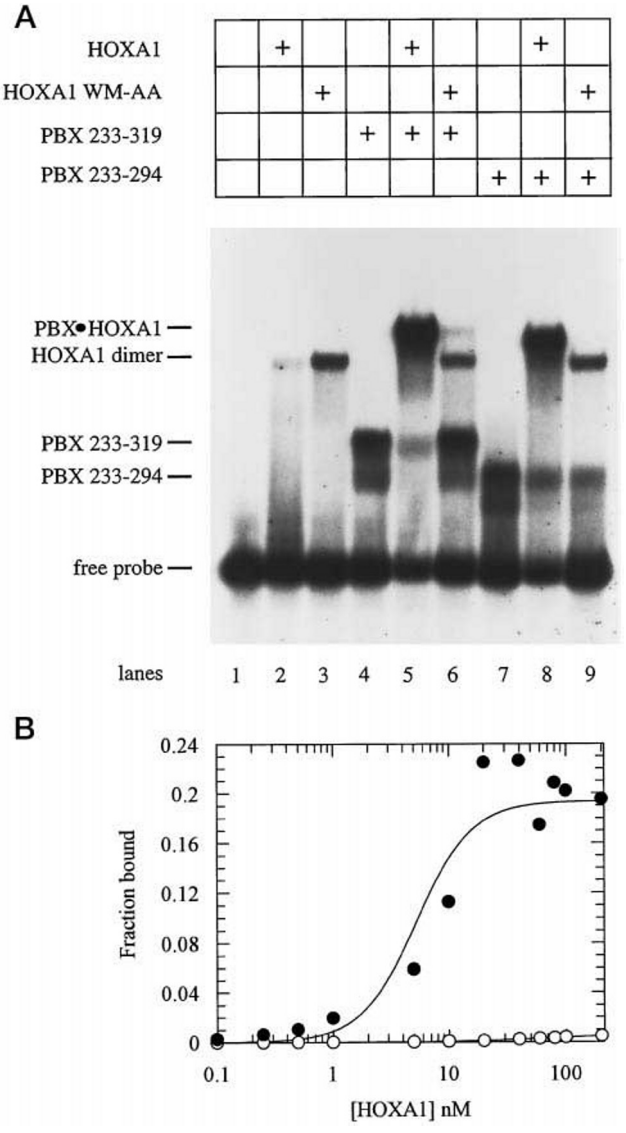

FIG. 4. *A*, the binding site for the YPWM motif is in the PBX homeodomain. The EMSA compares the relative abilities of the isolated PBX homeodomain, PBX233–294, and the PBX homeodomain plus the GKFQ domain, PBX233–319, to bind as heterodimers with HOXA1 and HOXA1(WM→AA). *B*, HOXA1 and PBX233–294 form cooperative complexes. Symbols used: ○, HOXA1; ● HOXA1/PBX233–294. The *x* scale shows the log concentration of the HOX partner used in the titration.

binding spanning three orders of magnitude. The HOXA1 prepared for this experiment displayed a lower amount of monomeric DNA binding activity than the preparation used in Fig. 4*A*. In general, HOXA1 displays poor DNA binding activity as a monomer (54), but variation in the extent of binding does occur between preparations. We then measured the amount of probe bound by the heterodimeric complex in the presence of a fixed amount of PBX 233–294. A plot of the fraction of probe bound *versus* the concentration of HOXA1 (log scale) revealed a strongly sigmoidal curve (Fig. 4*B*), the signature of cooperative interactions (55). The use of HOXA1 mutated in the YPWM motif failed to yield detectable levels of heterodimeric complex formation with PBX 233–294 (data not shown.) This experiment firmly establishes the cooperative nature of the interaction between HOXA1 and the PBX homeodomain and makes clear that this is dependent on the YPWM motif in the HOX partner. We conclude that the YPWM directly contacts the PBX homeodomain to promote cooperative DNA binding.

The GKFQ Domain Is Dependent on α-Helical Character and Requires Conserved Residues—We wished to determine the importance of specific residues of the GKFQ domain to HOX/PBX interactions in the context of full-length E2A-PBX. The effects of GKFQ domain mutations (Fig. 1*C*) on the stability of the HOXA1/E2A-PBX complexes were assessed in dissociation rate experiments (Fig. 5). Deletion of the entire GKFQ domain (EP ΔGKFQ) decreased complex stability by 49% (24.1 min from 46.6 min, Fig. 5, compare *lanes 1* and *2*). This was consistent with the transient transfection data wherein the deletion of the GKFQ domain caused 82% loss of transcriptional activation (Fig. 3, compare *lanes 3* and *4*). Two types of mutation, those predicted to disrupt α-helical character (alanine to proline) and those predicted to reveal functional residues (mutations to alanine) were then tested.

Three of four mutations to proline (F298P, A302P, and A306P/A307P) had similar deleterious effects. In transient transfections, between 91 and 95% of transactivation activity was lost (Fig. 3, compare *bar 3* with *bars 7*, *10*, and *13*). In dissociation rate assays, 39–44% of complex stability was lost compared with EP pswt (Fig. 5, compare *bar 1* with *bars 5*, *8*, and *11*). These results imply that α-helical character is important in this region. The G296P mutation had no effect, perhaps because α-helical structure at the start of the region is not present or not required. Of residues mutated to alanine, 2 were shown to decrease both transactivational activity and the stability of E2A-PBX·HOXA1 complexes on DNA. The E301A mutant decreased E2A-PBX·HOXA1 transactivation by 85%, similar to deletion of the whole GKFQ region (Fig. 3, compare *bars 4* and *9*). Dissociation rate assays were consistent with this observation, with complex stability on DNA decreased to 57% of EP pswt values, similar to the destabilization observed when the whole region was deleted (Fig. 5, compare *bars 2* and *7*). Mutation of Tyr-305 either alone, Y305A, or in the triple mutant, N303A/I304A/Y305A, reduced transactivation relative to EP pswt values (Fig. 3, compare *bar 3* with *bars 11* and *12*) by 89 and 74%, respectively, and complex stability to 68% of EP

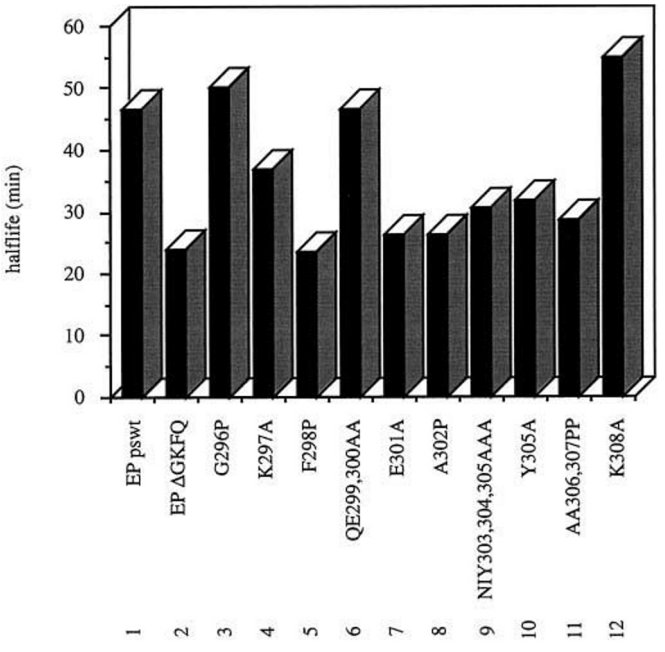

FIG. 5. **Mapping GKFQ domain residues which are critical to HOXA1/E2A-PBX complex stability.** The histogram is a summary of dissociation rate experiments in which a vast excess of cold competitor (see "Experimental Procedures") was added to preformed HOXA1/E2A-PBX complexes bound to radiolabeled probe and loaded at 6 different time points over the course of 40 min. A half-life was calculated based on the rate of dissociation of the radiolabeled probe over time. Average values, $n = 2$ or 3 are shown.

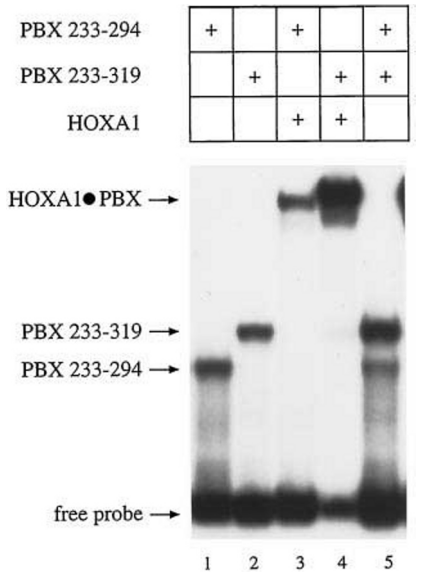

FIG. 6. **The GKFQ domain increases PBX monomer binding.** The EMSA compares the monomer binding of PBX 233–319 to that of the isolated PBX homeodomain, PBX 233–294. Their binding was compared as monomers in heterodimeric complexes with HOXA1 and in competition for probe in the same reaction. Twice as much labeled probe was added in *lane 5* to permit better visualization of both bands. There was an approximately 60-fold molar excess of protein over probe.

pswt values (Fig. 5, compare *bar 1* with *bars 9* and *10*). Since Tyr-305 is the only strictly conserved residue in the NIY triplet and both mutations have a similar effect, Tyr-305 is considered to be the functional residue.

The GKFQ Domain Contributes to PBX Monomer Binding—Since the GKFQ domain is not critically involved in protein-protein interactions through the YPWM motif, it was considered that it may improve the stability of HOX·E2A-PBX complexes by increasing the affinity of DNA binding by PBX. The importance of the GKFQ domain for PBX monomer binding was assessed with truncated proteins (Fig. 1*A*). Equal amounts of PBX homeodomain (PBX 233–294) and PBX homeodomain with the GKFQ domain (PBX 233–319) bind DNA with comparable efficiency (Fig. 6, *lanes 1* and *2*). However, when these two proteins must compete with one another for DNA binding sites, the presence of the GKFQ domain confers 5-fold better binding (Fig. 6, *lane 5*). This corresponds well with the 6-fold difference seen between the amount of cooperative complex formed with HOXA1 and each of these recombinant PBX proteins (Fig. 6, *lanes 3* and *4*).

The loss of the GKFQ domain also caused a similar 6-fold decrease in the stability of cooperative complexes formed with HOXA1. The complex of HOXA1 with PBX 233–319 had a half-life of 36 s (average of three experiments). Removal of the GKFQ (HOXA1·PBX 233–294) decreased the half-life markedly. It is difficult to measure the amount of complex remaining at time points earlier than 20 s after addition of competitor. In one experiment, there was no remaining HOXA1·PBX 233–294 complex at 20 s. In a second experiment, a small amount of complex remaining at 20 s allowed us to calculate a half-life of 6 s, six times shorter than that obtained with the GKFQ domain.

DISCUSSION

Two Models Describe the Roles of the YPWM Motif and the GKFQ Domain—Although domains important for cooperative interactions between HOX and PBX proteins have been identified, the manner in which they act remains unclear. We have examined two reported domains of importance for HOX/PBX interactions: the YPWM motif found N-terminal to the HOX homeodomain, and the conserved region immediately C-terminal to the PBX homeodomain, which we have named the GKFQ domain. Previously, two models have been proposed to describe the relative activities of these elements. In the first, the GKFQ domain is the primary contact site for the YPWM motif and is required for cooperative interaction with HOX proteins (44). In the second model, the YPWM motif contacts the PBX homeodomain, possibly in the loop separating helices 1 and 2 (45, 56) and/or in a hydrophobic pocket formed between helices 1 and 2 (45).

In support of a direct interaction between the HOX YPWM motif and the PBX GKFQ domain, two groups have reported that deletion of the GKFQ domain eliminates nearly all interactions with HOXB7 (40, 45). On the basis of these results and molecular modeling, it was proposed that these two regions contact each other in the DNA-bound heterodimer (21, 44). The GKFQ domain was thus designated the Hox cooperativity motif or HCM (44).

The second model is supported by data showing that HOXA5 can weakly interact with the isolated homeodomain of PBX and that addition of a free peptide containing the HOXA5 YPWM motif increases the monomer binding of the PBX homeodomain (43). Fusion of the YPWM motif to the PBX homeodomain also caused a similar increase in monomer binding, whereas the addition of the GKFQ domain to the PBX homeodomain increased its monomer binding 5-fold. Mutations in the GKFQ domain generally had a similar effect on monomer binding and cooperative interactions with HOXA5 (43). It was therefore proposed that the YPWM motif contacts the PBX homeodomain directly, and that the role of the GKFQ domain may be restricted to a direct DNA binding function.

The GKFQ Domain Is Not Essential for HOXA1/PBX Interaction—Our results are consistent with a primary interaction of the YPWM motif of HOXA1 with the homeodomain of PBX. In contrast to the first model, we find that the GKFQ domain is

not absolutely required for HOX/PBX interactions. Nonetheless, in the context of full-length E2A-PBX1A, we find that the GKFQ domain improves HOX·E2A-PBX·DNA complex stability *in vitro* (Fig. 5) and increases transactivational activity *in vivo* (Fig. 3). Our deletion of positions 296–308 removed the most highly conserved residues C-terminal to the homeodomain. Although it is possible that compensatory residues lie in the moderately conserved region that follows the GKFQ domain, we note that deletion of the entire region C-terminal to residue 309 of PBX had no effect on cooperativity with either HOXB7 or ENGRAILED-2 (45). We suggest that there is a differential requirement for the GKFQ domain, depending on the HOX partner involved. Thus, the HOXB7 protein (44, 45), unlike HOXA1 (this study), HOXA5 (43), and ENGRAILED-2 (45), is unable to interact substantially with PBX in the absence of the GKFQ. The linker region separating the YPWM motif of HOXB7 from the homeodomain is shorter than for any of the HOX proteins that show some independence from the GKFQ domain. It is possible that the HOXB7 linker length or composition imposes physical constraints, leading to stronger dependence on GKFQ domain function. Regardless of the explanation, the more modest contribution of the GKFQ domain to cooperativity with HOX proteins from paralog groups 1 and 5 suggests that the term "Hox cooperativity motif" is not a generally applicable designation for this region of PBC family proteins.

The YPWM Motif Contacts the PBX Homeodomain—To resolve whether the YPWM motif contacted the GKFQ domain or the PBX homeodomain, we compared the relative importance of the YPWM motif and the GKFQ domain for cooperativity. If the YPWM motif primarily contacted the GKFQ, the removal of the GKFQ domain should be as deleterious to cooperative interactions with HOX proteins as the loss of YPWM function. Instead, we find that mutation of the YPWM motif is far more disruptive than the deletion of the GKFQ domain (Fig. 2), demonstrating that the latter is not the primary contact site for the YPWM motif. Furthermore, we have shown that the isolated PBX homeodomain is able to cooperatively interact with HOX proteins in a YPWM-dependent manner (Fig. 4), establishing that a binding site for the YPWM motif is found within the PBX homeodomain. It does remain possible that the GKFQ domain serves a secondary role in contacting the YPWM motif, perhaps helping to form or stabilize the primary YPWM contact site in the PBX homeodomain. Our results are supported by the findings of other groups (43, 45), which implicate the 3 amino acid insertion in the extended loop of the PBX homeodomain and hydrophobic residues in helix 1 as HOX YPWM motif contact sites.

The GKFQ Domain Contributes to PBX Monomer Binding—The presence of the GKFQ domain increased by 5-fold the ability of the PBX homeodomain to bind DNA as a monomer when competing against the PBX homeodomain alone (Fig. 6, *lane 5*). The enhanced monomer binding of PBX233–319 was mirrored in the 5-fold abundance of PBX233–319·HOXA1 over PBX233–294·HOXA1 complex observed in steady-state binding conditions (Fig. 6, *lanes 3* and *4*) and in the 6-fold increase in the stability of the cooperative complex in dissociation rate experiments. Thus, the increase in monomer PBX binding is sufficient to explain the increased stability seen in E2A-PBX·HOXA1 cooperative complexes.

The GKFQ Domain Contributes to E2A-PBX·HOXA1 Complex Stability—It was thought, based on comparison to MATa1/MATα2 proteins, that the GKFQ domain of PBX might form an α-helix (50–52). Our mutational analysis supports this hypothesis since three mutations to proline (F298P, A302P, and A306P/A307P) destabilize E2A-PBX·HOXA1 complexes by 39–49% (Fig. 5) and result in 91–95% decreases in transactivation of a reporter gene (Fig. 3). The disruption of α-helical structure in this region would result in the mispositioning of functional GKFQ domain residues. It is possible that some of the residues mutated to proline are actually functional rather than structural residues. This is unlikely to be the case for the A302P and A306P/A307P mutants since the alanine residues could only provide a few hydrophobic interactions, and the loss of these contacts would be unlikely to cause the large effects observed. Phe-298 is the only residue mutated to proline that may also be functional, even though it is not strictly conserved in the PBC family (Fig. 1*B*). Lu and Kamps (43) have shown that an F298A mutation, a natural substitution occurring in EXD and CEH-20 (Fig. 1*B*), does decrease PBX233–326·HOXA5 steady-state binding and PBX233–326 monomer binding, indicating that Phe-298 may also play a role in GKFQ domain function.

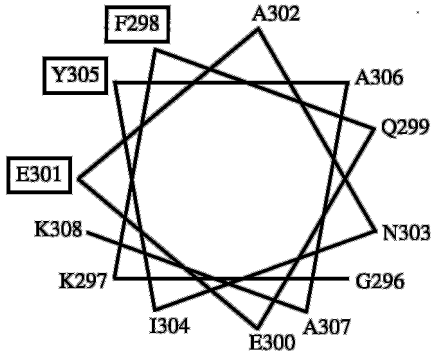

FIG. 7. **An α-helical wheel representation of the GKFQ domain.** The *boxed* amino acids are those that decreased stability and transactivation when mutated to alanine in this study (Glu-301 (*E301*), Tyr-305 (*Y305*)) or decreased monomer PBX 233–326 binding in the study by Lu and Kamps (43) (Phe-298 (*F298*)).

In this study, Glu-301 and Tyr-305 were shown to be the 2 main functional residues in the GKFQ domain. The E301A mutation decreased complex stability to 57% (26.4 *versus* 46.6 min) and transactivation to 15% of the EP pswt value. Glu-301 is an acidic residue that could participate in hydrogen bonds or in an ion pair with either lysine or arginine. Glu-301 has been mutated by Lu and Kamps (43) to arginine. They observed a large decrease in steady-state monomer binding of a PBX233–326 and a corresponding decrease in cooperative complex formation with HOXA5, consistent with the effects we observe. The highly polar nature of Glu-301 argues against it participating in hydrophobic contacts to the YPWM motif of HOXA1 or HOXA5.

The Y305A mutation decreased cooperative complex stability by 32% and transactivation by 89%. Tyr-305 could participate in a hydrogen bond or in hydrophobic interactions but could also stack with other aromatic residues. Mutation of Tyr-305 by Lu and Kamps (43) to phenylalanine showed no effect in their assays, suggesting that Tyr-305 does not form a hydrogen bond. Coupled with the results presented here, we suggest that Tyr-305 participates in important hydrophobic interactions or in a stacking interaction with another aromatic amino acid in the PBX homeodomain, thus stabilizing conformation and promoting contacts to DNA.

An α-helical wheel projection (Fig. 7) predicts that Glu-301, Tyr-305, and Phe-298 would lie on the same face of an α-helix, suggesting that they interact with PBX homeodomain residues that are close together in space. Further structural studies will be required to address the specific contacts made by Glu-301 and Tyr-305 to the PBX homeodomain.

Acknowledgments—We thank Dr. Michael Phelan for recombinant HOXA1 protein and Peter Pawelek for help with the GraFit computer program.

REFERENCES

1. Krumlauf, R. (1994) *Cell* **78,** 191–201
2. Zeltser, L., Desplan, C. and Heintz, N. (1996) *Development* **122,** 2475–2484
3. Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994) *Cell* **78,** 211–23
4. Bürglin, T. R. and Ruvkin, G. (1992) *Nature Genet.* **1,** 319–320
5. Kamps, M. P., Look, A. T. and Baltimore, D. (1991) *Genes Dev.* **5,** 358–68
6. Peifer, M. and Wieschaus, E. (1990) *Genes Dev.* **4,** 1209–23
7. Rauskolb, C., Peifer, M. and Wieschaus, E. (1993) *Cell* **74,** 1101–1112
8. Chan, S.-K., Jaffe, L., Capovilla, M., Botas, J. and Mann, R. S. (1994) *Cell* **78,** 603–15
9. van Dijk, M. A. and Murre, C. (1994) *Cell* **78,** 617–24
10. Pöpperl, H., Bienz, M., Studer, M., Chan, S.-K., Aparacio, S., Brenner, S., Mann, R. and Krumlauf, R. (1995) *Cell* **81,** 1031–1042
11. Phelan, M. L., Rambaldi, I. and Featherstone, M. S. (1995) *Mol. Cell. Biol.* **15,** 3989–3997
12. Lu, Q., Knoepfler, P. S., Scheele, J., Wright, D. D. and Kamps, M. P. (1995) *Mol. Cell. Biol.* **15,** 3786–3795
13. van Dijk, M. A., Peltenburg, L. T. and Murre, C. (1995) *Mech. Dev.* **52,** 99–108
14. Chan, S.-K., Popperl, H., Krumlauf, R. and Mann, R. S. (1996) *EMBO J.* **15,** 2476–2487
15. Shen, W.-F., Rozenfeld, S., Lawrence, H. J. and Largman, C. (1997) *J. Biol. Chem.* **272,** 8198–8206
16. Chan, S.-K., Ryoo, H.-D., Gould, A., Krumlauf, R. and Mann, R. S. (1997) *Development* **124,** 2007–2014
17. Zhao, J. J., Lazzarini, R. A. and Pick, L. (1996) *EMBO J.* **15,** 1313–1322
18. Knoepfler, P. S., Lu, Q. and Kamps, M. P. (1996) *Nucleic Acids Res.* **24,** 2288–2294
19. Chan, S.-K. and Mann, R. S. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5223–5228
20. Phelan, M. L. and Featherstone, M. S. (1997) *J. Biol. Chem.* **272,** 8635–8643
21. Chang, C.-P., Brocchieri, L., Shen, W.-F., Largman, C. and Cleary, M. L. (1996) *Mol. Cell. Biol.* **16,** 1734–1745
22. Shen, W. F., Chang, C. P., Rozenfeld, S., Sauvageau, G., Humphries, R. K., Lu, M., Lawrence, H. J., Cleary, M. L. and Largman, C. (1996) *Nucleic Acids Res.* **24,** 898–906
23. Lu, Q. and Kamps, M. P. (1997) *Oncogene* **14,** 75–83
24. Shanmugam, K., Featherstone, M. S. and Saragovi, H. U. (1997) *J. Biol. Chem.* **272,** 19081–19087
25. Monica, K., Galili, N., Nourse, J., Saltman, D. and Cleary, M. L. (1991) *Mol. Cell. Biol.* **11,** 6149–6157
26. Kamps, M. P., Murre, C., Sun, X.-H. and Baltimore, D. (1990) *Cell* **60,** 547–555
27. Nourse, J., Mellentin, J. D., Galili, N., Wilkinson, J., Starbridge, E., Smith, S. D. and Cleary, M. L. (1990) *Cell* **60,** 535–545
28. Kamps, M. P. and Wright, D. D. (1994) *Oncogene* **9,** 3159–66
29. Dedera, D. A., Waller, E. K., LeBrun, D. P., Sen, M. A., Stevens, M. E., Barsh, G. S. and Cleary, M. L. (1993) *Cell* **74,** 833–43
30. Monica, K., LeBrun, D. P., Dedera, D. A., Brown, R. and Cleary, M. L. (1994) *Mol. Cell. Biol.* **14,** 8304–8314
31. Kamps, M. P., Wright, D. D. and Lu, Q. (1996) *Oncogene* **12,** 19–30
32. Lawrence, H. J., Sauvageau, G., Humphries, R. K. and Largman, C. (1996) *Stem Cells* **14,** 281–91
33. Blatt, C., Aberdam, D., Schwartz, R. and Sachs, L. (1988) *EMBO J.* **7,** 4283–4290
34. Hatano, M., Roberts, C. W., Minden, M., Crist, W. M. and Korsmeyer, S. J. (1991) *Science* **253,** 79–81
35. Nakamura, T., Largaespada, D. A., Lee, M. P., Johnson, L. A., Ohyashiki, K., Toyama, K., Chen, S. J., Willman, C. L., Chen, I.-M., Feinberg, A. P., Jenkins, N. A., Copeland, N. G. and Shaughnessy Jr., J. D. (1996) *Nature Genet.* **12,** 154–158
36. Borrow, J., Shearman, A. M., Stanton Jr., V. P., Becher, R., Collins, T., Williams, A. J., Dubé, I., Katz, F., Kwong, Y. L., Morris, C., Ohyashiki, K., Toyama, K., Rowley, J. and Housman, D. E. (1996) *Nature Genet.* **12,** 159–167
37. Nakamura, T., Largaespada, D. A., Shaughnessy Jr., J. D., Jenkins, N. A. and Copeland, N. G. (1996) *Nature Genet.* **12,** 149–153
38. Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle, C. and Boncinelli, E. (1986) *Nature* **324,** 664–668
39. Knoepfler, P. S. and Kamps, M. P. (1995) *Mol. Cell. Biol.* **15,** 5811–5819
40. Chang, C.-P., Shen, W.-F., Rozenfeld, S., Lawrence, H. J., Largman, C. and Cleary, M. (1995) *Genes Dev.* **9,** 663–674
41. Johnson, F. B., Parker, E. and Krasnow, M. A. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 739–743
42. Neuteboom, S. T. C., Peltenburg, L. T. C., van Dijk, M. A. and Murre, C. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 9166–9170
43. Lu, Q. and Kamps, M. P. (1996) *Mol. Cell. Biol.* **16,** 1632–1640
44. Chang, C.-P., De Vivo, I. and Cleary, M. L. (1997) *Mol. Cell. Biol.* **17,** 81–88
45. Peltenburg, L. T. C. and Murre, C. (1997) *Development* **124,** 1089–1098
46. Green, S., Issemann, I. and Sheer, E. (1988) *Nucleic Acids Res.* **16,** 369
47. Bradford, M. M. (1976) *Analytical Biochemistry* **72,** 248–254
48. Leatherbarrow, R. J. (1992) GraFit version 3.0. Erithacus Software Ltd. Staines, UK
49. Peltenburg, L. T. C. and Murre, C. (1996) *EMBO J.* **15,** 3385–3393
50. Phillips, C. L., Stark, M. R., Johnson, A. D. and Dahlquist, F. W. (1994) *Biochemistry* **33,** 9294–9302
51. Stark, M. R. and Johnson, A. D. (1994) *Nature* **371,** 429–32
52. Li, T., Stark, M. R., Johnson, A. D. and Wolberger, C. (1995) *Science* **270,** 262–269
53. Richardson, J. S. and Richardson, D. C. (1989) Prediction of protein structure and the principles of protein conformation. Plenum Press, New York
54. Phelan, M. L., Sadoul, R. and Featherstone, M. S. (1994) *Mol. Cell. Biol.* **14,** 5066–5075
55. Ptashne, M. (1992) A Genetic Switch. Cell Press, Cambridge
56. Lu, Q. and Kamps, M. P. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 470–474
57. Waterston, R., Martin, C., Craxton, M., Huynh, C., Coulson, A., Hillier, L., Durbin, R., Green, P., Shownkeen, R., Halloran, N., Metzstein, M., Hawkins, T., Wilson, R., Berks, M., Du, Z., Thomas, K., Thierry-Mieg, J. and Sulston, J. (1992) *Nat. Genet.* **1,** 114–123