

Zinc Fingers 1–7 of EVI1 Fail to Bind to the GATA Motif by Itself but Require the Core Site GACAAGATA for Binding*

(Received for publication, September 1, 1995, and in revised form, October 23, 1995)

Archibald S. Perkins[‡] and Jeong H. Kim

From the Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06520-8023

EVI1 is a zinc finger oncoprotein that binds via fingers 1–7 to the sequence GACAAGATAA. The target genes on which EVI1 acts are unknown. This binding motif overlaps with that for the GATA transcription factors, (T/A)GATA(A/G), and GATA-1 can bind to and activate transcription via a GACAAGATAA motif. The possibility has been raised that, when overexpressed in leukemogenesis, EVI1 may function by interfering with the differentiation-promoting action of GATA factors. To explore this, we have assessed the affinity of EVI1 for the GATA binding sites derived from erythroid-specific GATA-1 target genes, and found only low affinity interactions. We examined the contacts between EVI1 and DNA by methylation interference studies, which revealed extensive contacts between EVI1 and its binding site. The importance of the contacts for high affinity binding was shown by *in vitro* quantitative gel shift studies and *in vivo* cotransfection studies. To examine what types of sequences from mouse genomic DNA bind to EVI1, we isolated and sequenced five EVI1-binding fragments, and each showed the GACAAGATA site. The data presented contribute to our knowledge of the binding specificity of EVI1, and yield a clearer picture of what sequences can, and cannot, act as targets for EVI1 action.

The current molecular understanding of myeloid leukemogenesis has come largely from the identification and characterization of genes that contribute to this multistep process. Proviral tagging in retrovirally induced tumors has been a powerful way of identifying myeloid-specific oncogenes. One such oncogene, *Evi1*, was identified as a common site of retroviral insertion in murine myeloid leukemias (1, 2). Rearrangements in *Evi1* have also been documented in human myelodysplasias and leukemias, indicating its involvement in human disease (3–9). Both the retroviral insertions and the chromosomal alterations at *Evi1* result in transcriptional activation of the gene, suggesting that the gene acts as a dominant oncogene in leukemia. Interestingly, most acute myeloid leukemias bearing *EVI1*-activating alterations are of the M0, M1, or M2 class, and are usually CD34-positive, suggesting an immature phenotype. Cell lines with such alterations are dependent on hematopoietic growth factors, indicating that *Evi1* does not abrogate growth factor requirements (4, 10).

cDNA cloning and analysis showed that the murine *Evi1* encodes a 1042-residue, 145-kDa protein with 10 zinc finger

motifs that are separated into two domains by 481 amino acids (10). There is also an acidic domain at the C-terminal end, which may function in transcriptional activation, although, as presented in this paper, *Evi1* does not appear to function as a transcriptional activator. A shorter isoform of *Evi1*, which migrates as 88 kDa, is produced via alternative splicing in both human (11) and mouse (12). This form lacks zinc fingers 6 and 7, as well as 269 of the adjacent C-terminal amino acids. Since these fingers are important for DNA binding (13), it is likely that this isoform has different DNA binding characteristics than the 145-kDa isoform, although this has not been carefully examined.

The finding of 10 zinc fingers in the protein argues that *Evi1* encodes a sequence-specific DNA-binding protein that plays a role in nucleic acid regulation, most likely RNA transcription. The consensus binding sites for both the first and second sets of zinc fingers have been identified; fingers 1–7 bind to TGACAA-GATAA (14) or GACAAGATAAGATAA (13), and the second set of fingers, numbers 8–10, bind to GAAGATGAG (15). The first *EVI1* binding site, TGACAA-GATAA, shows overlap with the binding site for the GATA family of transcription factors, which bind to the consensus motif (A/T)GATA(A/G) (16, 17). This site was first identified as a common motif present in *cis*-acting elements of erythroid-specific genes, and through mutagenesis studies was found to be functionally important for erythroid gene transcription (18–21). The GATA family of transcription factors now includes GATA-1 (22, 23), GATA-2 (24–26), GATA-3 (27–29), GATA-4 (30, 31), GATA-5 (32), and GATA-6 (32), each of which has a distinct pattern of expression. While GATA-2 expression can be detected in early myeloid precursors (33), it is expressed in other cells as well. To date, no myeloid-specific GATA family member has been reported.

The overlap in binding sites for *EVI1* and the GATA factors suggested the possibility that *EVI1* may bind to GATA sites located in *cis* to GATA target genes and influence their transcription. Support for this theory comes from the observation that ectopic expression of *Evi1* in primary bone marrow cells inhibits differentiation in response to erythropoietin (34). Since differentiation along the erythroid lineage is dependent on GATA1 activity (35), and since there is overlap between the GATA-1 and *EVI1* binding sites, it is possible that loss of erythropoietin responsiveness results from *EVI1*-induced repression of GATA-1-responsive gene(s) harboring an *EVI1* binding site. Also in support of this, as presented below, *Evi1* can repress GATA1-mediated activation of responsive reporter constructs in cotransfection studies in NIH 3T3 cells. *Bona fide* target genes for *EVI1* have not been reported, and the importance of *Evi1*-induced repression of GATA factor function within cells has not been carefully addressed. In this paper, we report the further characterization of the interaction between zinc fingers 1–7 of *EVI1* and its binding site, both *in vitro* and *in vivo*. These studies indicate a high degree of specificity of *EVI1*'s interaction with DNA and show that *EVI1* has negli-

* This work was supported by Grant 92-52 from the James S. McDonnell Foundation. 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: Dept. of Pathology, Yale University School of Medicine, P. O. Box 208023, New Haven, CT 06520-8023. Tel.: 203-785-6843; Fax: 203-785-7467; E-mail: perkins@biomed.med.yale.edu.

ble affinity for known GATA sites. Our data suggest that, theoretically, only a small subset of GATA sites, those conforming to the full EVI1 binding site (TGACAAGATAA or GACAA-GATAAGATAA), can be bound *in vivo* by both proteins. However, none of the known GATA target sites conforms to this motif.

EXPERIMENTAL PROCEDURES

Cell Lines and DNA Transfections—NIH3T3 cells were cultured at 37 °C/5% CO₂ in Dulbecco's modified Eagle's medium supplemented with 10% calf serum and penicillin/streptomycin. Assays for chloramphenicol acetyltransferase (CAT)¹ and β -galactosidase were as described (36). NIH3T3 cells were transfected with a total of 15 μ g of DNA/10-cm plate by the calcium phosphate precipitate method (37) with 2 min of "shock" with 15% glycerol at 4–8 h after transfection. Cells were assayed 36–60 h later. All assays were done in triplicate. Acetylation products were analyzed by ethyl acetate extraction and thin layer chromatography, followed by quantitation using a PhosphorImager (Molecular Dynamics). Statistical analysis was done using the SYSTAT 5 software program.

Plasmid Construction and Production of Recombinant Protein—A bacterial expression plasmid for a truncated version of murine EVI1 comprising the first domain of zinc fingers was constructed for use in overexpressing and isolating EVI1 protein. This plasmid, pMBP-EVI1(1–254), was constructed by inserting a mutated *Evi1* gene in frame and downstream from the gene for the maltose-binding protein (MBP) in plasmid pMalC (New England Biolabs; Ref. 38). So that only the first set of zinc fingers was produced, a stop codon was engineered at amino acid 254 in EVI1 in plasmid pKS-FLEVI1 (a pBluescript derivative containing the full-length *Evi1* cDNA; Ref. 10) by site-directed mutagenesis (Amersham Corp.; Ref. 39) on a single-stranded phagemid substrate using oligonucleotide (5'-GGCCAAGGCTGACTTCCCTG-GAACC-3'). The mutagenized *Evi1* coding region was transferred to pMalC by amplifying it by polymerase chain reaction (PCR) with primers (5'-CGGAATTCATGGCGCTGACATCCACGAAGAA-3' and 5'-CGTCTAGATTGGTCCACTCTCGTCAACCTTGACAATGTC-3'), digesting with *EcoRI* and *XbaI*, and ligating into similarly digested pMalC vector. pMBP-EVI1(1–254) was transferred into DH5 α strain of *Escherichia coli*, which were then used for production and purification of EVI1(1–254) essentially as described (14). The chimeric *Evi1*-VP16 activator construct was made by amplifying sequences encoding amino acids 1–250 (zinc fingers 1–7) of *Evi1* from an *Evi1* cDNA (plasmid p58.2–1 (10)) by PCR using oligonucleotides (5'-GAATTCATGGCGC-CTGACATCCACGAAGAA-3' and 5'-GAATTCTCCAGGAAGTGAAT-GCCTTGGCC-3'), cutting this product with *EcoRI*, and ligating it into *EcoRI*-cut pNLVP16 (40). This plasmid contains the SV40 early region promoter upstream of the start of transcription and is thus active in mammalian cells.

Methylation Interference and Missing Base Contact Probing—Methylation interference was performed as described (41). A 244-bp double-stranded DNA fragment bearing a single wild type 11-bp binding site for EVI1 (TGACAAGATAA) was generated by PCR using pAP34 (a pCAT promoter derivative containing one EVI1 binding site at the *BamHI* site) as a template, and using an M13 forward primer and an oligonucleotide primer having sequence (5'-GAAGTGGCGGAGTTAG-3') that symmetrically flanks the binding site, one of which was ³²P-labeled; analyses were separately performed using top strand-labeled and bottom strand-labeled substrates. These radiolabeled PCR products were purified, methylated with dimethylsulfate and bound to MBP-EVI1(1–254) (see "Plasmids and Protein"). The DNA-protein reaction mixture was fractionated on a nondenaturing gel (4% polyacrylamide, 0.5 \times Tris borate-EDTA (0.045 M Tris borate, 0.001 M EDTA)). Bound and free fractions were isolated, treated with piperidine, fractionated on a 6% sequencing gel, and autoradiographed as described (41).

The missing base contact probing was performed as described (42) and was essentially the same as that described above, except that, instead of modification with dimethylsulfate, the DNA was treated with 0.1 M formic acid for depurination, or hydrazine for depyrimidation. These modified DNAs were then bound to MBP-EVI1(1–254), fractionated, cleaved with piperidine, analyzed on a 6% sequencing gel, and autoradiographed as described (42).

¹ The abbreviations used are: CAT, chloramphenicol acetyltransferase; MBP, maltose-binding protein; PCR, polymerase chain reaction; bp, base pair(s).

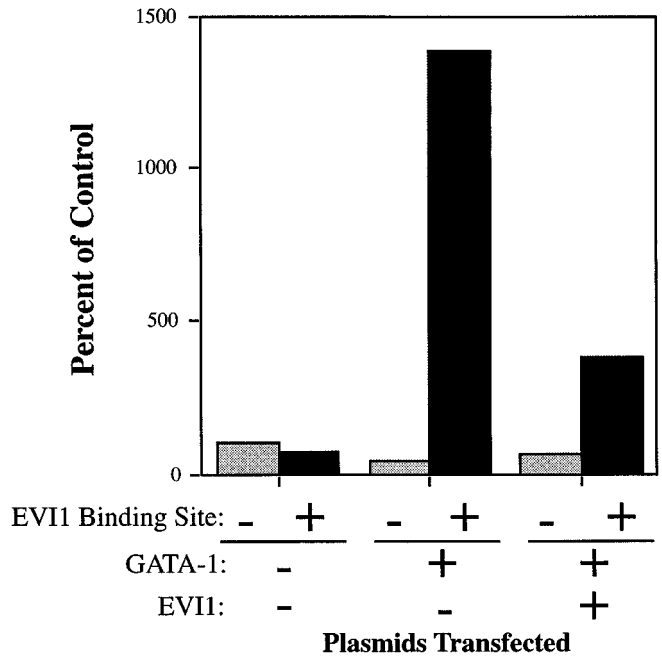


FIG. 1. Cotransfection studies in NIH 3T3 cells showing that GATA-1 activates reporters containing the GACAAGATAA motif and that *Evi1* can repress this activation. 10-cm plates (in duplicate) were transfected with a tk-CAT reporter (5 μ g/plate), with or without the EVI1 binding site as indicated, along with GATA-1 (2.5 μ g/plate) and *Evi1* (0.5 μ g/plate) expression plasmids as indicated. CAT activity in cells is expressed as a percentage of the activity in cells transfected with tk-CAT (no binding site) alone.

Electromobility Shift Assay—Protein-DNA equilibrium binding assays were performed by mixing 0.5 pmol (10,000 cpm) of ³²P-labeled oligonucleotide and 0.35 pmol of protein in a 15- μ l reaction containing 25 mM HEPES, pH 7.5, 50 mM KCl, 5 mM MgCl₂, 50 μ M ZnCl₂, 1 mM dithiothreitol, 10% glycerol, 100 μ g/ml bovine serum albumin, and 250 μ g/ml poly(dI-dC) (Pharmacia Biotech Inc.). Binding was conducted at 30 °C for 20 min, after which the reaction was fractionated on a 4% polyacrylamide, 0.1% bisacrylamide gel, buffered with 0.5 \times Tris borate/EDTA buffer. After electrophoresis and drying, the gel was exposed to a PhosphorImager screen, and quantitation was performed on a Molecular Dynamics PhosphorImager. To assure that assays were performed in the presence of excess DNA, titrations of both DNA and protein were done prior to doing the competition with mutant oligonucleotides. The affinity of EVI1 for a cold competitor was calculated as the slope of a plot of ((1 - Y)/Y) versus amount of competitor, where Y is the ratio of the fraction of radiolabeled wild type oligonucleotide bound in the presence of competitor to the fraction bound with no competitor (43). This affinity was then expressed as a percentage of that for the wild type competitor (TGACAAGATAA). All binding assays were done in triplicate.

RESULTS

GATA-1 Can Transactivate via an EVI1 Binding Site, and *Evi1* Represses This Transactivation—The binding site for the first set of zinc fingers of EVI1, GACAAGATAA, contains a GATA consensus sequence, (A/T)GATA(A/G), making it likely that the GATA family of transcription factors can bind to the EVI1 motif. To examine this, we cotransfected a GATA-1 expression vector with a CAT reporter bearing an EVI1 binding site upstream of the herpes simplex virus thymidine kinase gene promoter region into NIH 3T3 cells. As shown in Fig. 1, GATA-1 specifically activates the CAT reporter with the EVI1 binding site, but not the parental vector. When an *Evi1* expression plasmid was cotransfected as well, there was a significant reduction in CAT activity, indicating that *Evi1* can repress GATA-1-induced transactivation (Fig. 1). *Evi1* did not repress the basal activity of the parental reporter (Fig. 1), indicating that EVI1 does not interfere with the basal transcriptional

TABLE I
Relative affinity of EVI1 binding to known and mutant GATA sites

Asterixes indicate identities with the wild type sequence. % Sim. indicates the percent similarity with the wild type sequence; % Affin. represents the percent affinity relative to the wild type. The final column gives the literature reference for the sequence.

Sequence	Identity	% Sim.	% Affin.	Ref.
TGACAAGATAAG	Wild type	100	100	(14)
G**T*****	Chicken α^D globin	83	7	(63)
*****	Chicken α^D globin G \rightarrow T, T \rightarrow C	100	100	
*C*****A	PCNA	83	3	(64)
*****A	PCNA, C \rightarrow G	92	100	
G*CA*****	Mouse β^{maj1}	67	28	(20)
GC*ACT*****	"GATA" consensus	58	6	(16, 17)
GC*ACT*****	Mouse α_1 globin	58	2	(20)
GAACCT*****	Mouse carbonic anhydrase 1	50	8	(20)
AT*AT*****TA	Transcobalamin	50	4	(65)
G*GAT*****GA	Mouse β^{maj3}	50	6	(20)
G*GG*G*****TC	Human gelatinase	50	8	
GCTTCT***G**	Mutant GATA	42	8	

factors, but most likely represses GATA-1-mediated activation by site occupancy of the EVI1 binding site. These results are in agreement with those of others (34), and suggest a possible mechanism of *Evi1* action in leukemogenesis: that it may interfere with the normal activation of genomic GATA family target genes essential for differentiation, thereby blocking differentiation.

EVI1 Fails to Bind to Known GATA Sites—It is clear that GATA-1 can bind to and transactivate via the EVI1 binding site, and that EVI1 can block this transactivation. This raises the possibility that EVI1 may interfere with the action of GATA-1 on its own targets, that is, erythroid-specific genes that bear the GATA consensus motif. One way to address this is to see if EVI1 binds to known GATA sites that exist adjacent to erythroid-specific genes. In addition, by searching the GenBank™ and EMBL DNA sequence data bases, we had uncovered other potential genomic targets of EVI1, including sequences from the proliferating cell nuclear antigen promoter, the gelatinase promoter, and the transcobalamin promoter (Table I). To test the affinity of the first zinc finger domain of EVI1 for these sites, we performed competitive gel shift experiments, using a radiolabeled "wild type" EVI1 binding site (TGACAAGATAA) and MBP-EVI1(1–254) protein purified from *E. coli*. The MBP-EVI1(1–254) protein used in these studies contains amino acids 1–254 of EVI1, which comprises zinc fingers 1–7 of EVI1, and shows essentially identical DNA binding for the TGACAAGATA as full-length protein (data not shown). It was used in these studies because of its ease of purification and its stability.

The binding of MBP-EVI1(1–254) for the radiolabeled wild type oligonucleotide was competed with varying amounts of different cold competitor oligonucleotides, each corresponding to a candidate target site, bearing the GATA core together with the surrounding sequences from the indicated genes. Surprisingly, none of these sequences bound to EVI1 with appreciable affinity (Table I). The chicken α globin GATA site, having sequence GATAAGATAA, deviates from the EVI1 core only at position 3. Despite this degree of similarity to the EVI1 consensus, the chicken α globin sequence bound with poor avidity to EVI1 (Table I). However, when a C was substituted for that T in the chicken α globin sequence, it bound with wild type affinity (Table I), indicating the importance of C at position 3.

These experiments indicate that none of the reported GATA target sites adjacent to either erythroid-specific genes, or other genes that we tested, are likely to be high affinity targets of EVI1 binding. To assure accurate determination of EVI1-DNA affinities, each value for relative affinity was derived from the data of binding reactions done in triplicate with three competitor concentrations. In addition, both protein and DNA titra-

tions were done prior to competitive gel shift to assure that reactions were performed in DNA excess, and most of the affinity measurements were performed on separate occasions with essentially the same results.

Identification of Base Contacts for EVI1 Zinc Fingers 1–7—The failure of EVI1 to bind to GATA sites that differ by only a single base pair change in the TGACAAGATAA motif suggested a high degree of specificity in its interaction with DNA. Given this apparent specificity, we were interested in defining the minimum requirements for high affinity binding of EVI1 to the first recognition site, TGACAAGATAA. To that end, we performed a series of "missing contact probing" (42) and methylation interference (41) assays. End-labeled ^{32}P -DNA containing an EVI1 binding sequence (TGACAAGATAA; numbered in the text as T⁰, G¹, A², C³, etc.) was modified with various chemicals, incubated with MBP-EVI1(1–254), fractionated into binding and nonbinding pools by nondenaturing gels, and cleaved at sites of modification with piperidine. Analysis of the piperidine cleavage products on a sequencing gel allows the identification of the specific bases modified in the two pools. The reagents used were dimethylsulfate, which preferentially methylates guanine residues; formic acid, which depurinates the DNA; and hydrazine, which depyrimidates the DNA. We used bacterially expressed EVI1, comprising zinc fingers 1–7 (MBP-EVI1(1–254)) to isolate the bound and free fractions. Modifications that interfere with binding are enriched in the nonbound pool of fragments and underrepresented in the bound pool. On the top strand, both G¹ and G⁶ were determined to form important contacts, since the unbound DNA (Fig. 2A, *second lane from left*) was greatly enriched for DNAs methylated at these sites. Depurination of these sites also interfered with binding (Fig. 2A, *central lanes*), as did depurination of A², A⁴, A⁵, and A⁷. Modification of A⁹ or A¹⁰, however, did not interfere with binding. Analysis of contact points on the top strand using hydrazine revealed that depyrimidation at T⁰, C³, or T⁸ was deleterious to binding. By labeling the bottom strand (Fig. 2B), contacts were evident at the G opposite C³, at the A opposite T⁸, as well as the T nucleotides across from A⁴ and A⁵. Minor effects on binding are seen with depyrimidation of the C and T opposite G⁶ and A⁷, respectively.

In Vitro Binding of EVI1 to Mutant Binding Site Oligonucleotides—The results of the methylation interference and base-removal analyses are summarized in Fig. 3, in which the size of the *arrow* is proportional to the relative strength of the EVI1-DNA interaction. It is clear that the protein makes extensive contact with the motif, and argues that nearly all of the specified bases are required for high affinity binding. We wished to confirm these results and to examine the relative importance of these bases to high affinity binding. To that end,

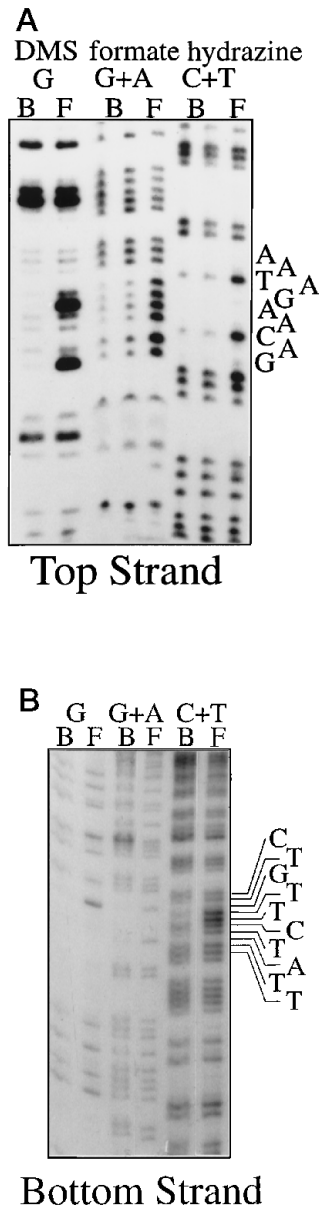


FIG. 2. Methylation interference and missing contact probing studies reveal multiple contact points between EVI1 and its binding site. Depicted are denaturing PAGE analysis of the piperidine cleavage products following chemical modification and EVI1 binding. To the right of each panel is the EVI1 binding motif. Panel A shows results obtained with the top strand radiolabeled. The type of modification is as indicated, along with the specific base(s) modified. B and F denote the fractions that bound to EVI1, or remained free, respectively. For the formate and hydrazine reactions, there are two bound lanes presented. Panel B shows results obtained with the bottom strand radiolabeled.

we performed competitive gel shift experiments, in which binding of a radiolabeled wild type EVI1 binding site (TGACAA-GATAA) to MBP-EVI1(1–254) was competed by varying amounts of cold competitor oligonucleotides, each bearing a single base change relative to the wild type motif. From the degree of competition observed with increasing concentrations of competitor oligonucleotide, one can calculate the affinity of the protein for the competitor relative to the wild type oligonucleotide (see “Materials and Methods” for details of relative affinity calculation). Given the length of the binding motif and the number of contacts, we thought it likely that a change of 1 base out of the 10-bp motif would incrementally diminish the affinity of binding to EVI1. The results of these studies were



FIG. 3. Summary of the methylation interference and missing contact probing analysis. The arrows indicate the points of contact between EVI1 and the 11-bp binding site. The size of the arrow depicts the relative importance of the contact, as assessed by visual assessment of band intensities shown in Fig. 2.

rather surprising, revealing that changes at G¹, A², C³, A⁵, G⁶, or T⁸ virtually eliminated binding to EVI1 *in vitro* (Fig. 4A). These and our previous binding site selection experiments (14) indicated that the base at position 10 could be A or T, and the missing contact probing experiments described above indicated no contact of EVI1 with A¹⁰. Consistent with these findings, alteration at A¹⁰ had a less dramatic effect (27% the affinity of the wild type site) than changes at other positions (Fig. 4A). A longer recognition site for zinc fingers 1–7 of EVI1 was identified, having sequence GACAAGATAAGATAA (13). We assessed the relative affinity of this 15-bp EVI1 binding motif by quantitative gel shift analysis, and found that it had identical affinity to the 11-bp motif that we had identified (TGACAA-GATAA) (Fig. 4A).

These data support and extend the missing contact probe and methylation interference experiments, and argue that changes at positions 1, 2, 3, 5, 6, or 8 are not compatible with high affinity binding to EVI1. Together with *in vitro* binding selection studies (13, 14), the data presented here indicate a minimum core EVI1 binding site of GACAAGATA, which extends by 4 bases on the 5' side the GATA consensus sequence, (A/T)GATA(A/G).

The in Vivo Binding of EVI1 to Mutant Sites Parallels in Vitro Binding—We now sought to determine if the EVI1 binding data obtained *in vitro* from competitive gel shift experiments was an accurate reflection of *in vivo* binding. Since the nuclear milieu is distinct from that of the *in vitro* conditions, one can imagine that different binding results may be obtained. Additionally, accessory proteins that bind to EVI1 may stabilize binding to sites that appear to be low affinity sites *in vitro*, when only EVI1 is present. To assess EVI1 binding to oligonucleotides *in vivo*, we created CAT reporters in which either wild type or various mutant binding site oligonucleotides were placed upstream of a minimal promoter from the herpes simplex virus thymidine kinase gene, and cotransfected these reporters into NIH 3T3 cells along with an expression plasmid for *Evi1*. The reporter constructs tested contained one of the following binding motifs: the wild type motif (either the 11-bp or the 15-bp motif), or motifs with a single base pair deviation from the wild type 11-bp motif (G¹ → T, C³ → A, or G⁶ → T). Remarkably, *Evi1* had no effect on any of these plasmids (Fig. 4B), indicating that *Evi1* does not contain a transcriptional activator, at least not one that is active with these binding sites in NIH 3T3 cells. That EVI1 protein was overproduced as a consequence of plasmid transfection in these experiments was shown by immunoblotting and gel shift experiments (data not shown).

Since *Evi1* proved ineffectual on these promoters, we constructed a dominantly acting chimeric *Evi1* activator, comprised of the VP16 activation domain of herpes simplex virus linked downstream of the first zinc finger region of *Evi1* (fin-

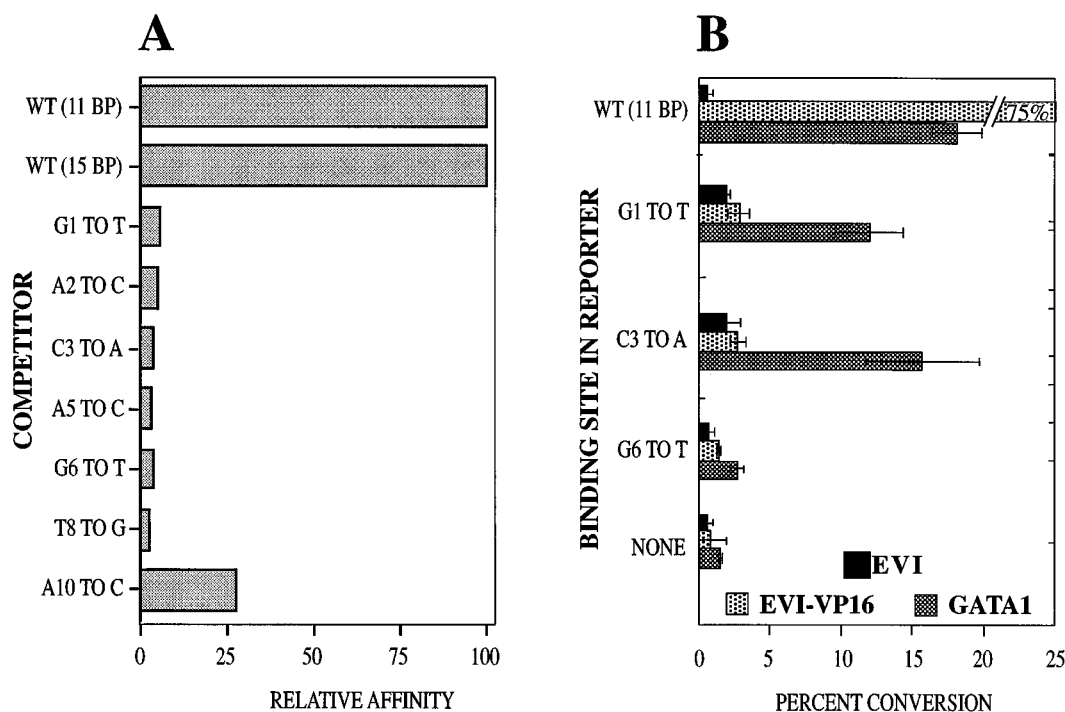


FIG. 4. **A**, quantitative gel shift studies reveal the affinity of MBP-EVI1(1–254) for various oligonucleotides with single base changes to the GACAAGATAA motif, relative to the wild type (WT) 11-bp motif. To the left is indicated the mutation in the competing oligonucleotide. The bars indicate the affinity expressed as a percentage of that for the 11-bp wild type oligonucleotide. Also shown (WT(15 BP), second bar from top) is the relative affinity for the 15-bp motif identified by Delwel *et al.*¹³: (GACAAGATAAGATAA). **B**, cotransfection studies to show the *in vivo* activity of the wild type (WT (11 BP)) or mutant binding sites (as indicated). Plates (10 cm) of NIH 3T3 cells (in triplicate) were cotransfected with tk-CAT reporters (5 μ g/plate) containing the binding site indicated (inserted at bp –109 relative to the start of transcription), along with either an *Evi1*, *Evi1-VP16*, or GATA-1 expression vector, as indicated. The CAT activity in cell lysates is expressed as the percent acetylation of the [¹⁴C]chloramphenicol substrate, plus and minus the standard error.

gers 1–7, amino acids 1–250). This construct encoded a protein with potent transactivating capacity that was specific for reporters bearing a wild type EVI1 binding motif (Fig. 4B). The reporters having mutant binding sites (G¹ → T, C³ → A, or G⁶ → T) were not responsive to the *Evi1-VP16* chimera, indicating that EVI1 had little if any affinity for these sites *in vivo*. As expected, GATA-1 was able to activate reporters containing the wild type motif, as well as the G¹ → T and C³ → A mutations, but not the G⁶ → T mutation. Since G⁶ is part of the GATA binding motif, it is not surprising that reporters bearing this mutant fail to be activated by GATA-1. These data show that the mutant reporters are functional and respond appropriately to GATA-1.

Genomic Binding Sites for EVI1 Contain GACAAGATAA Sequence—To complement the studies described above, we wanted to determine the sequences to which EVI1 would bind with high affinity within the mouse genome. This was of interest, since all of the published studies concerning the DNA binding specificity of EVI1 have been performed with motifs derived from selections of random oligonucleotide pools (13, 14). The binding sites within the mouse genome may include contextual features important for binding not apparent in the selected oligonucleotides, and thus their elucidation may give further insight into the specificity of the protein for DNA. To accomplish this, purified bacterially expressed MBP-EVI1 (1–254) was allowed to bind to a plasmid-based genomic library of mouse DNA sequences (average insert size 3.5 kilobases), and the resulting DNA-protein complexes were isolated by filtration of the mixture through nitrocellulose filters. The bound DNA was then eluted and subjected to subsequent rounds of selection and filtration. To identify binding sites within the selected plasmids, individual plasmids were digested with *Sau3A1* and cloned into pBluescript, and then a pool of sub-

clones for each selected plasmid was again selected by binding to MBP-EVI1(1–254) and nitrocellulose filtration. This sub-selection yielded plasmids containing the *Sau3A1* fragment of the parentals that contained the EVI1 binding site. Since these fragments were relatively short (average insert size 250 bp), they were sequenced to find the EVI1 binding site. Sequence data from five such sites is shown in Fig. 5. All five contain the GACAAGATA core; none contain the extra GATAA at the 3' end, which was identified in the cycles of selection with random oligonucleotides (13). Four out of five contain a T preceding the first G, and, interestingly, all contain an A at the –2 position. However, since our probe for the methylation interference studies contained a T at the –2 position, we cannot assess potential contacts between the protein and the base at this site.

DISCUSSION

The molecular role of EVI1 in leukemogenesis is not clear. In certain studies, it appears to interfere with differentiation program induced by certain cytokines in cultured and primary hematopoietic cells (34, 44). EVI1 is a sequence-specific DNA-binding protein possessing 10 zinc fingers in two separate domains. In our previous studies, we found that EVI1 binds with high affinity to the sequence TGACAAGATAA (14). Given the DNA binding properties of EVI1, a possible mechanism for its role in leukemogenesis is to dominantly interfere with the expression of genes required for normal hematopoiesis. To arrive at a mechanistic understanding of the role that EVI1 plays in leukemia, we are conducting studies to determine its genetic targets within the cell, and its action on those targets. One possible set of genetic targets in this mechanism is the GATA target genes, whose proper regulation is essential for the erythroid lineages. The GATA motif (T/A)GATA(A/G) is a *cis*-acting element that plays an important role in the regulation of

Clone JK1	5'TCTAGGTTAT GACAAGATA ACACATGATGA3'
Clone JK2	5'ATTCAGAAAT GACAAGATA GCAGAATAATG3'
Clone JK3	5'GCATACATAT GACAAGATA AATAGATTATT3'
Clone JK4	5'GCTTGTGCAT GACAAGATA GCAACTGCAAC3'
Clone JK5	5'GGTTGGAAG GACAAGATA AATGATGTTAT3'

FIG. 5. **Sequence analysis of the EVI1 binding sites isolated from mouse genomic DNA.** Depicted are five EVI1 binding sites, each from the clone indicated, identified within the *Sau3A1* subclones of plasmids selected for EVI1 binding. The core region of identity between the sequences is indicated in *bold*.

erythroid-specific genes. It is present adjacent to a variety of globin (16) and non-globin (18, 45) erythroid-specific genes, and is present within a minimal 3'- β -globin gene enhancer (17). Functional importance of the GATA motif in the regulation of these genes has emerged from numerous studies (20, 21, 46–48). The presence of a (T/A)GATA(A/G) motif within the EVI1 binding site suggested to us the possibility that EVI1 may play a role in the regulation of genes that are transcriptionally controlled via this motif. In addition, by searching the DNA data bases, we identified near matches of the EVI1 binding site adjacent to several genes, including proliferating cell nuclear antigen.

In this paper we report our studies on the ability of the first set of zinc fingers of EVI1 (numbers 1–7) to bind to various GATA binding sites, and found that EVI1 binds poorly if at all to known GATA sites, despite only single base pair differences between the EVI1 motif (GACAAGATA) and certain GATA target sites, such as that for the chicken α globin gene (GATAA-GATA) (Table I). From our quantitative studies presented here, it is clear that EVI1 cannot bind with any appreciable affinity to a AGATAA monomer. This makes it very likely that, without accessory factors that could increase affinity, EVI1 does not play any significant role in the regulation of genes that contain AGATAA but lack the additional bases that would yield a high affinity site for EVI1. This includes most of the (T/A)G-ATA(A/G) sequences that have been found to be important for erythroid-specific gene transcription. It is likely that GATA proteins can bind to endogenous EVI1 sites, since they are likely to contain the (T/A)GATA(A/G) motif.

The high degree of specificity that EVI1 shows for its binding motif was probed further by performing methylation interference and missing base contact probing experiments. These revealed that EVI1 makes contacts at multiple bases in its recognition sequence (Fig. 3). These findings were confirmed and extended by studying the binding both *in vitro* and *in vivo* of EVI1 to a series of oligonucleotides bearing single base changes to the GACAAGATA motif (Fig. 4, *A* and *B*). Interestingly, we observed the same affinity between EVI1 and the short motif that we described previously (GACAAGATAA) and the longer one described by Delwel *et al.* (13) (GACAAGATAA-GATAA). Since both of these motifs were identified by *in vitro* binding to random oligonucleotides, and may not be representative of EVI1 binding sites in the nucleus, we examined what sites EVI1 binds to within genomic mouse DNA, by performing binding selection experiments with genomic fragments (average size 3.5 kilobases). This led to the identification of EVI1 binding sites with the shorter rather than the longer motif. Indeed, in a larger scale selection performed with full-length EVI1 protein produced in insect cells, we have isolated 16 more EVI1 binding sites in genomic mouse DNA, and have never found the longer motif.² These findings suggesting that the selection of the longer 15-bp motif from the pool of random oligonucleotides was a function of the method used, and may not represent a physiologically relevant motif.

One consistent finding in our studies was that *Evi1* fails to activate transcription of any reporter, despite the presence of an acidic region, which commonly acts as an activation domain in transcription factors. Reporters that contained multimers of the GACAAGATAA motif exhibited only basal levels of transcription in most cell lines tested, including NIH3T3, Ltk⁻ cells, choriocarcinoma cells (JEG), hepatoma cells (HepG2), and HeLa cells (Fig. 4, data not shown). In these cells, cotransfection of *Evi1* had little effect (Fig. 4, data not shown). When these reporters are activated by GATA-1, then addition of *Evi1*-expressing plasmids results in transcriptional repression. Similarly, in WEHI cells, a myelomonocytic leukemic cell line that does not express *Evi1*, the GACAAGATAA-containing reporter was expressed at significantly higher levels; in these cells, cotransfection of *Evi1* resulted in transcriptional repression (data not shown). Thus, while *Evi1* did not appear to act as a repressor of basal transcription, it did appear to repress higher level transcription, suggesting that EVI1 does not interact with the basal transcription machinery, but rather with other transcription factors or coactivator proteins. This repressive effect of *Evi1* appeared to depend on the presence of a GACAAGATAA sequence; no effect was seen on viral promoters such as the SV40 early region promoter or the metallothionein promoter (data not shown). Tanaka *et al.* (49) have reported that *Evi1* overexpression in NIH 3T3 cells leads to increased AP-1 activity, and that this activity is dependent on the presence of the second set of zinc fingers on the protein. This suggests that EVI1 may be acting as a transcriptional activator; the effect may be indirect, however, since, although *Evi1* appears to increase *c-fos* transcription, there is no evidence of EVI1 binding to the *c-fos* promoter.

Structure-function studies on EVI1 have shown that high affinity binding to the GACAAGATA sequence is mediated by fingers 4–7 (13). The structural studies on the zinc finger have showed that each finger interacts with three nucleotides (50), which argues that the key fingers (4–7) would interact with 12 bp. The work we present here, specifically the methylation interference and missing base contact probing, indicates an essential core of 9 base pairs: GACAAGATA, which might interact with only three of these fingers. Based on a comparison of the zinc finger amino acid sequences with their cognate DNA binding sites, as well as the information from structural studies, a set of rules to define these interactions has been proposed (50–53). It has been difficult to use these rules to assign specific zinc fingers in the first domain of EVI1 with base pairs in the binding site. As suggested previously, the RN residues in fingers 4 and 6 may interact with the GA residues of the EVI1 binding site (13). In this regard, it is informative to consider the crystal structure of the Tramtrack (TTK) protein, the finger 1 of which has a recognition helix sequence very similar to zinc fingers 4 and 6 of EVI1 (given N-terminal to C-terminal): HISNFCR for TTK, QFSNLCR for finger 6, and DPSNLQR for finger 4. The TTK finger binds to the sequence GAT, with the R contacting the G, the N contacting the A, and the S interacting with T. Interestingly, the GAT is followed by A in the TTK binding site, and the A-T-A triplet is readily deformable (54), allowing the short side-chained S to make contact with T. Even with this added insight, however, it is difficult to phase the interactions of specific fingers with specific triplets to satisfy the methylation interference studies and the predictions made by the recognition code. What this suggests is that further analysis of EVI1-DNA interactions will provide information on novel types of zinc finger-DNA interactions and allow a better understanding of this important class of DNA binding motif. Necessary additional information may come from ongoing mutagenesis studies.

² J. Y. Aycocck, P. Hui, A. R. Bjoring, J. H. Kim, and A. S. Perkins, submitted for publication.

Our data clearly indicate that in a variety of settings EV11 can act as a transcriptional repressor. Other leukemogenic nuclear proteins that appear to function as transcriptional repressors in their oncogenic state are ErbA, a truncated and mutated form of the thyroid hormone receptor, and PML-RAR α , the protein encoded by the t(15;17) translocation of acute promyelocytic leukemia. The target genes for none of these three leukemogenic oncoproteins is known with certainty, but it appears that ErbA functions in oncogenesis by blocking the effects of retinoic acid (55). Likewise, PML-RAR α may interfere with the action of RAR α on its normal targets (56). Alternatively, PML-RAR α may act in oncogenesis by interfering with intranuclear localization of PML (57, 58). It is intriguing that these three leukemogenic zinc finger proteins all potentially act as transcriptional repressors, and may act to block normal cellular differentiation (44, 59–62). While the genetic targets of these factors that are relevant to leukemogenesis are not known, it could be that similar pathways of differentiation are blocked by the different proteins, but via a different target with each factor. Future studies aimed at the identification of target genes for these proteins should yield insight into this possibility.

Acknowledgments—We thank N. G. Copeland and M. C. Lopingco for helpful suggestions.

REFERENCES

- Mucenski, M. L., Taylor, B. A., Ihle, J. N., Hartley, J. W., Morse, H. C., III, Jenkins, N. A., and Copeland, N. G. (1988) *Mol. Cell. Biol.* **8**, 301–308
- Bergeron, D., Poliquin, L., Houde, J., Barbeau, B., and Rassart, E. (1992) *Virology* **191**, 661–669
- Oval, J., Smedsrud, M., and Taetle, R. (1992) *Leukemia* **6**, 446–451
- Morishita, K., Parganas, E., Willman, C. L., Whittaker, M. H., Drabkin, H., Oval, J., Taetle, R., Valentine, M. B., and Ihle, J. N. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 3937–3941
- Fichelson, S., Dreyfus, F., Berger, R., Melle, J., Bastard, C., Miclea, J., and Gisselbrecht, S. (1992) *Leukemia* **6**, 93–99
- Nucifora, G., Begy, C. R., Kobayashi, H., Roulston, D., Claxton, D., Pedersen-Bjergaard, J., Parganas, E., Ihle, J. N., and Rowley, J. D. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4004–4008
- Levy, E. R., Parganas, E., Morishita, K., Fichelson, S., James, L., Oscier, D., Gisselbrecht, S., Ihle, J. N., and Buckle, V. J. (1994) *Blood* **83**, 1348–1354
- Suzukawa, K., Parganas, E., Gajjar, A., Abe, T., Takahashi, S., Tani, K., Asano, S., Asou, H., Kamada, N., Yokota, J., Morishita, K., and Ihle, J. N. (1994) *Blood* **84**, 2681–2688
- Mitani, K., Ogawa, S., Tanaka, T., Miyoshi, H., Kurokawa, M., Mano, H., Yazaki, Y., Ohki, M., and Hirai, H. (1994) *EMBO J.* **13**, 504–510
- Morishita, K., Parker, D. S., Mucenski, M. L., Jenkins, N. A., Copeland, N. G., and Ihle, J. N. (1988) *Cell* **54**, 831–840
- Morishita, K., Parganas, E., Douglass, E. C., and Ihle, J. N. (1990) *Oncogene* **5**, 963–971
- Bordereaux, D., Fichelson, S., Tambourin, P., and Gisselbrecht, S. (1990) *Oncogene* **5**, 925–927
- Delwel, R., Funabiki, T., Kreider, B., Morishita, K., and Ihle, J. (1993) *Mol. Cell. Biol.* **13**, 4291–4300
- Perkins, A. S., Fishel, R., Jenkins, N. A., and Copeland, N. G. (1991) *Mol. Cell. Biol.* **11**, 2665–2674
- Funabiki, T., Kreider, B. L., and Ihle, J. N. (1994) *Oncogene* **9**, 1575–1581
- Evans, T., Reitman, M., and Felsenfeld, G. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 5976–5980
- Wall, L., deBoer, E., and Grosveld, F. (1988) *Genes & Dev.* **2**, 1089–1100
- Mignotte, V., Eleouet, F., Raich, N., and Romeo, P. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6548–6551
- Reitman, M., and Felsenfeld, G. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6267–6271
- Plumb, M., Frampton, J., Wainwright, H., Walker, M., Macleod, K., Goodwin, G., and Harrison, P. (1989) *Nucleic Acids Res.* **17**, 73–93
- Martin, D., Tsai, S., and Orkin, S. (1989) *Nature* **338**, 435–438
- Evans, T., and Felsenfeld, G. (1989) *Cell* **58**, 877–885
- Tsai, S.-F., Martin, D., Zon, L., D'Andrea, A., Wong, G., and Orkin, S. (1989) *Nature* **339**, 446–451
- Dorfman, D., Wilson, D., Burns, G., and Orkin, S. (1992) *J. Biol. Chem.* **267**, 1279–1285
- Lee, M., Temizer, D., Clifford, J., and Quertermous, T. (1991) *J. Biol. Chem.* **266**, 16188–16192
- Yamamoto, M., Ko, L., Leonard, M., Beug, H., Orkin, S., and Engel, J. (1990) *Genes & Dev.* **4**, 1650–1662
- Joulin, V., Bories, D., Eleouet, J., Labastie, M., Chretien, S., Mattei, M., and Romeo, P. (1991) *EMBO J.* **10**, 1809–1816
- Ho, I., Vorhees, P., Marin, N., Karpinski Oakley, B., Tsai, S., Orkin, S., and Leiden, J. (1991) *EMBO J.* **10**, 1187–1192
- Ko, L., Yamamoto, M., Leonard, M., George, K., Ting, P., and Engel, J. (1991) *Mol. Cell. Biol.* **11**, 2778–2784
- Arceci, R., King, A., Simon, M., Orkin, S., and Wilson, D. (1993) *Mol. Cell. Biol.* **13**, 2235–2246
- Kelly, C., Blumberg, H., Zon, L., and Evans, T. (1993) *Development* **118**, 817–827
- Laverriere, A., MacNeill, C., Mueller, C., Poelmann, R., Burch, J., and Evans, T. (1994) *J. Biol. Chem.* **269**, 23177–23184
- Nagai, T., Harigae, H., Ishihara, H., Motohashi, H., Minegishi, N., Tsuchiya, S., Hayashi, N., Gu, L., Andres, B., and Engel, J. (1994) *Blood* **84**, 1074–84
- Kreider, B., Orkin, S., and Ihle, J. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6454–6458
- Pevny, L., Simon, M. C., Robertson, E., Klein, W. H., Tsai, S. F., D'Agati, V., Orkin, S., and Costantini, F. (1991) *Nature* **349**, 257–260
- Sambrook, J., Fritsch, E., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Graham, F., and Vander Eb, A. (1973) *Virology* **52**, 456–467
- Maina, C. V., Riggs, P., Grandea, A. G., III, Slatko, B. E., Moran, L. S., Tagliamonte, J. A., McReynolds, L. A., and di Guan, C. (1988) *Gene (Amst.)* **74**, 365–373
- Taylor, J. W., Ott, J., and Eckstein, F. (1985) *Nucleic Acids Res.* **13**, 8765–8785
- Fearon, E., Finkel, T., Gillison, M., Kennedy, S., Casella, J., Tomaselli, G., Morrow, J., and Van Dang, C. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7958–7962
- Seibenlist, U., and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 122–126
- Brunelle, A., and Schleif, R. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6673–6676
- Glass, C., Holloway, J., Devary, O., and Rosenfeld, M. (1988) *Cell* **54**, 313–323
- Morishita, K., Parganas, E., Matsugi, T., and Ihle, J. N. (1992) *Mol. Cell. Biol.* **12**, 183–189
- Mignotte, V., Wall, L., deBoer, E., Grosveld, F., and Romeo, P.-H. (1989) *Nucleic Acids Res.* **17**, 37–54
- Gumucio, D., Rood, K., Gray, T., Riordan, M., Sartor, C., and Collins, F. (1988) *Mol. Cell. Biol.* **8**, 5310–5322
- Watt, P., Lamb, P., Squire, L., and Proudfoot, N. (1990) *Nucleic Acids Res.* **18**, 1339–1350
- Gong, Q., Stern, J., and Dean, A. (1991) *Mol. Cell. Biol.* **11**, 2558–2566
- Tanaka, T., Nishida, J., Mitani, K., Ogawa, S., Yazaki, Y., and Hirai, H. (1994) *J. Biol. Chem.* **269**, 24020–24026
- Pavletich, N., and Pabo, C. (1991) *Science* **252**, 809–817
- Desjarlais, J., and Berg, J. (1992) *Proteins* **12**, 101–104
- Desjarlais, J., and Berg, J. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7345–7349
- Klevit, R. (1991) *Science* **253**, 1367
- Travers, A., and Klug, A. (1990) in *DNA Topology and Its Biological Effects* (Cozzarelli, N., and Wang, J., ed) pp. 57–106, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Sharif, M., and Privalsky, M. (1991) *Cell* **66**, 885–893
- Kastner, P., Perez, A., Lutz, Y., Rochette-Egly, C., Gaub, M., Durand, B., Lanotte, M., Berger, R., and Chambon, P. (1992) *EMBO J.* **11**, 629–642
- Dyck, J., Maul, G., Miller, W., Jr., Chen, J., Kakizuka, A., and Evans, R. (1994) *Cell* **76**, 333–343
- Weis, K., Rambaud, S., Lavau, C., Jansen, J., Carvalho, T., Carmo-Fonesca, M., Lamond, A., and Dejean, A. (1994) *Cell* **76**, 345–356
- Grignani, F., Ferrucci, P., Testa, U., Talamo, G., Fagioli, M., Alcalay, M., Mencarelli, A., Grignani, F., Peschle, C., Nicoletti, I., and Pelicci, P. (1993) *Cell* **74**, 423–431
- Gandrillon, O., Jurdic, P., Pain, B., Desbois, C., Madjar, J., Moscovici, M., Moscovici, C., and Samarut, J. (1989) *Cell* **58**, 115–121
- Schroeder, C., Gibson, L., and Beug, H. (1992) *Oncogene* **7**, 203–216
- Schroeder, C., Gibson, L., Zenke, M., and Beug, H. (1992) *Oncogene* **7**, 217–227
- Kemper, B., Jackson, P., and Felsenfeld, G. (1987) *Mol. Cell. Biol.* **7**, 2059–2069
- Travali, S., Ku, D., Rizzo, M., Ottavio, L., Baserga, R., and Calabretta, B. (1989) *J. Biol. Chem.* **264**, 7466–7472
- Johnston, J., Yang-Feng, T., and Berliner, N. (1992) *Genomics* **12**, 459–464