We have targeted the activation of an endogenous chromosomal locus including the human erythropoietin gene using synthetic transcription factors. These transcription factors are targeted to particular DNA sequences in the 5′-flanking region of the erythropoietin gene through engineering of a zinc finger DNA binding domain. The DNA binding domain is linked to a VP16 transcriptional activation domain. We find that these synthetic transcription factors invariably activate transiently transfected templates in which sequences within the 5′ flank of the erythropoietin gene are fused to a luciferase reporter. The efficiency of activation under these circumstances at a defined site is dependent on DNA binding affinity. In contrast, only a subset of these zinc finger proteins is able to activate the endogenous chromosomal locus. The activity of these proteins is influenced by their capacity to gain access to their recognition elements within the chromatin infrastructure. Zinc finger transcription factors will provide a powerful tool to probe the determinants of chromatin accessibility and remodeling within endogenous chromosomal loci.

The enormous progress in our understanding of gene control in eukaryotes using model systems presents substantial opportunities to apply this knowledge for therapeutic benefit in man. The rational design and engineering of components of the transcriptional machinery provide a powerful means to test conventional paradigms for the roles of protein-DNA and protein-protein interactions in gene regulation. These designer transcription factors may also provide novel means of regulating endogenous chromosomal loci for a variety of beneficial purposes. Over the past decade, the primary structural determinants of DNA recognition by zinc fingers of the Cys2-His2 type have been elucidated (1–8). Designer transcriptional regulators containing three or more zinc finger domains have been used in isolation (9, 10) or following linkage to transcriptional activation (9, 11–13) or repression domains (12, 13). These novel proteins control the transcription of reporter genes both transiently transfected into human cells (9, 11–13) and endogenous chromosomal loci (9, 13). Exactly how these regulatory functions are exerted remains to be resolved. An important issue in considering transcription factor function in eukaryotes is the capacity of the regulator to gain access to specific sites in chromatin and recruit transcriptional co-activators and co-repressors that modify the chromatin environment (14). These issues have been investigated for the archetypal Cys2-His2 zinc finger protein and transcriptional regulator TFIIIA (15–26). There is general agreement that the nucleosome can impede recognition of specific promoter elements by TFIIIA (15, 17–26) and that modification of histone-DNA interactions through nucleosome repositioning (18, 19), histone depletion (20, 21), and removal of the histone tails (22, 23) can promote TFIIIA binding to a nucleosomal infrastructure. Accumulation of histone H1 in chromatin can specifically interfere with TFIIIA function in vivo (24, 25) in a process that involves the repositioning of nucleosomes (25, 26). These studies demonstrate the role of chromatin infrastructure access by transcription regulators containing zinc finger DNA binding domains of the Cys2-His2 type.

In this work, we first designed 10 novel zinc finger DNA binding domains to recognize specific 9-bp sequences in the 5′ flank of the erythropoietin gene and characterized their unique DNA recognition selectivities and variable affinities for DNA. These zinc finger domains were then linked to the VP16 transcriptional activation domain (27) and tested for their capacity to activate transcription in both transient transfections and from endogenous chromosomal loci. Our results indicate that all of our synthetic regulators that bind DNA in vitro with dissociation constants <10 nM can activate transiently transfected templates. This activation of a particular site is dependent on DNA binding affinity. In contrast, only a subset of the synthetic regulators can activate the endogenous chromosomal locus.

We find that the regulators that work in the endogenous chromosome can bind within the chromatin infrastructure, but that the differential binding of distinct regulators at a particular site is largely independent of primary DNA binding affinity. These studies indicate that chromosome and chromatin organization is a determinant of zinc finger transcription factor function within endogenous chromosomal loci.

**EXPERIMENTAL PROCEDURES**

**Design, Synthesis, and Purification of Zinc Finger Proteins**—The foundation for our design strategy was to scan 1000 bp of the 5′-
flanking sequences of the human erythropoietin (EPO) gene to choose the best candidate recognition elements for which to design zinc finger DNA binding domains. The human transcription factor Sp1 (amino acids 532–624) was used as the backbone to assemble 10 distinct zinc finger DNA binding domains. The design determinants of these proteins will be described elsewhere; however, the amino acid sequences chosen to recognize particular sequences are illustrated in Table I. Our strategy to synthesize the zinc finger DNA binding domains is outlined in Fig. 1. To create the zinc finger proteins (EPOZFPs), a polymerase chain reaction (PCR)-based assembly procedure was applied using six overlapping oligonucleotides. Three oligonucleotides for the zinc finger coding sequences encode portions of the DNA binding domain containing the β sheet and linker regions between the α-helical DNA recognition sequences of the Sp1 zinc finger DNA binding domain scaffold (Fig. 1A, oligos 1, 3, and 5). The other three oligonucleotides encode the recognition helices (oligos 2, 4, and 6). The overlap between adjacent oligonucleotides is 15 base pairs. The PCR synthesis was carried out in two steps (Fig. 1B). First, the double-stranded DNA template was created by combining the six oligos and filling the gaps with a four-cycle PCR reaction (using Taq and Pfu thermostable DNA polymerases). The annealing temperature was 25 °C, a temperature at which the six oligos would anneal to form a DNA scaffold. In the second phase of construction, the template was amplified using a pair or external primers containing KpnI and BamHI restriction sites. The PCR products were directly cloned into the Tac pMalC2 bacterial expression vector. In the second phase of construction, the template was amplified using a pair or external primers containing KpnI and BamHI restriction sites. The PCR products were directly cloned into the Tac pMalC2 bacterial expression vector. After 2 weeks of selection in medium containing Zeocin (Invitrogen) cell line using LipofectAMINE (Life Technologies, Inc.), stable clones were isolated and analyzed for doxycycline (Dox)-dependent transcriptional activation of ZFP expression.

33851

Transcriptional Activation of Chromatin

**TABLE I**

| Name              | Target sequence | Subsites | Designs |
|-------------------|-----------------|----------|---------|
| EPOZFP–862a       | GCGGTGGTCTc     | F1GGTg   | QSSDLQ  |
|                   |                 | F2GGTg   | RSDALSR |
|                   |                 | F3CGGg   | RSDERKR |
| EPOZFP–862b       | GCGGTGGTCTc     | F1GGTg   | QSSDLQ  |
|                   |                 | F2GGTg   | RSDALSR |
|                   |                 | F3CGGg   | RSDERKR |
| EPOZFP–862c       | GCGGTGGTCTc     | F1GGTg   | QSSDLQ  |
|                   |                 | F2GGTg   | RSDALSR |
|                   |                 | F3CGGg   | RSDERKR |
| EPOZFP–545a       | GGTGAGGAGGt     | F1GAGg   | RSDNLAR |
|                   |                 | F2GAGg   | RSDNLAR |
|                   |                 | F3GTTg   | RSDNLAR |
| EPOZFP–545b       | GGTGAGGAGGt     | F1GAGg   | RSDNLAR |
|                   |                 | F2GAGg   | RSDNLAR |
|                   |                 | F3GTTg   | RSDNLAR |
| EPOZFP–535        | GAGGTGGGAGa     | F1GCGg   | RADTLR  |
|                   |                 | F2GGTg   | RSDALSR |
|                   |                 | F3CGGg   | RSDERKR |
| EPOZFP–233        | TGGTCCGGGg      | F1GAGg   | RSDNLAR |
|                   |                 | F2GAGg   | RSDNLAR |
|                   |                 | F3GTTg   | RSDNLAR |
| EPOZFP–82         | GTGGGAGGAg      | F1GAGg   | RSDNLAR |
|                   |                 | F2GGGg   | RSDDLR  |
|                   |                 | F3GTGg   | RSDALSR |
| EPOZFP–72         | GGGGTGGGGg     | F1GAGg   | RSDNLAR |
|                   |                 | F2GGGg   | RSDDLR  |
|                   |                 | F3GTTg   | RSDNLAR |
| Sp1               | GGGGGGAGGt      | F1GAGg   | KTSHLRA |
|                   |                 | F2GAGg   | RSDDLR  |
|                   |                 | F3GTTg   | RSDNLAR |
| Zf268             | GCGGTGGCTc      | F1GAGg   | RSDNLAR |
|                   |                 | F2GAGg   | RSDDLTT |
|                   |                 | F3GCGg   | RSDERKR |

was removed. The pBS579 construct, which was used as a negative control, encodes a nonspecific ZFP gene. **Cell Culture: Stable Inducible Cell Lines Expressing ZFPs, Luciferase Reporter Assays, Northern Blots, Taqman Analysis, ELISA, and Western Blots—Human embryonic kidney cells (HEK293) were grown in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum. To generate stable Tet-inducible EPOZFP cell lines, the coding region from the pEPOZFP–862 cDNA was subcloned into pcDNA4/TO (Invitrogen) using AflIII and HindIII restriction sites. The resulting plasmid pTO-EPOZFP–862 construct was transfected into the T-Rex-293TM (Invitrogen) cell line using LipofectAMINE (Life Technologies, Inc.). After 2 weeks of selection in medium containing ZeocinTM (Invitrogen), stable clones were isolated and analyzed for doxycycline (Dox)-dependent activation of ZFP expression. The luciferase reporter constructs were generated by annealing two complementary oligonucleotides in an annealing buffer containing 50 mM NaCl, 10 mM Tris-HCl, 1 mM MgCl₂, 1 mM dithiothreitol. The oligo mixture was heated at 95 °C for 5 min and allowed to cool down slowly to room temperature. The annealed oligonucleotides containing three tandem repeats of the ZFP target sequences were inserted into the pGL3 promoter vector (Promega) between the MluI and BglII sites.
upstream of the SV40 promoter. All constructs were confirmed by DNA sequencing.

Transient transfection was carried out using LipofectAMINE. Luciferase reporter assays were performed by co-transfection of ZFP effector DNA (50 ng), luciferase reporter DNA (900 ng), and pCMV-β-gal (100 ng, used as an internal control) into HEK293 cells seeded in six-well plates. Cell lysates were harvested 40 h after transfection, and the luciferase activities were measured by the Dual-Light luciferase and β-galactosidase reporter assay system (Tropix). To assay the activation of the endogenous chromosomal EPO gene, we made use of established procedures to carry out Northern analysis of EPO mRNA. Briefly, poly(A)1 RNA was isolated from either mock-transfected or pcEPOZFP862-transfected HEK293 cells using the Oligotex kit (Qiagen, Valencia, CA). 7 μg were resolved on a 2.4% agarose gel containing 2.4 M formaldehyde and blotted onto Nytran SuPerCharge membrane using 20X SSC. The membrane was hybridized at 65 °C for 1 h in Rapid-Hyb Buffer (Amersham Pharmacia Biotech) containing 32P-labeled EPO cDNA probe. The same membrane was re-hybridized with a 32P-labeled GAPDH DNA probe after stripping the EPO probe. The EPO cDNA construct, pcEPO was generated by inserting a human EPO cDNA fragment obtained by PCR amplification into the pcDNA3.1 vector (Invitrogen) at the XbaI and EcoRI sites. The clone was confirmed by sequencing. The pTBAHVP16 plasmid, which was used to generate riboprobes for detection of both human β-actin and ZFP genes, was generated by inserting the VP16 fragment from the pcDNA-NVF vector into the pTRI-β-actin-125-human vector (Ambion, Austin, TX). The pcDNA-NVF DNA was digested with XhoI, repaired with Klenow enzyme, and digested again with BamHI. The TRI-β-actin vector was

Fig. 2. The properties of EPOZFPs. A, schematic representation of the human EPO gene, showing key structural features. Horizontal arrow, start of transcription; open rectangle, coding region; hatched boxes, Alu elements; black boxes, targeted sequences, sites 1, 2, 3, and 4; wavy lines, CpG island. Abbreviations are as follows: Ba, BamHI; X, XbaI; Bg, BglII; PsI, HRE, hypoxic response element. Numbering is relative to the start site of transcription (+1). The targeted sites and the positions of the first nucleotide in the sequences is shown at the bottom. B, the protein gel shows all the purified EPOZFP, Sp1, and Zif268 proteins. The proteins were expressed as maltose-binding fusion proteins and purified as described (see "Experimental Procedures"). The leftmost lane contains size markers. C, gel-shift assays using the various EPOZFPs binding to their target sites. The name of each EPOZFP is indicated by the number in the top left-hand corner of each gel panel. Proteins are used to bind to their targets in 2-fold serial dilutions, with the highest protein concentration in lane 2, and the lowest concentration in lane 16 from left to right. Lane 1 is a control lane containing radiolabeled DNA alone. The equilibrium dissociation constants of each EPOZFP are indicated at the top of each panel.
digested with Smal and BamHI.

For Taqman analysis of mRNA abundance, total cellular RNA from transfected HEK293 cells was isolated using the RNeasy kit (Qiagen, Valencia, CA). Real time PCR analysis (Taqman) was performed in a 96-well format on an ABI 7700 SDS machine (Perkin Elmer) and analyzed with SDS version 1.6.3 software. RNA samples (20 ng) were mixed with 0.3 μM each primer, 0.1 μM probe, 5.5 mM MgCl2, and 0.3 mM each dNTP, 0.625 unit of AmpliTaq Gold RNA polymerase, 6.25 units of Multiscribe reverse transcriptase, and 5 units of RNase inhibitor in Taqman buffer A from Perkin Elmer. The reverse transcription was performed at 48 °C for 30 min. After denaturing at 95 °C for 10 min, PCR amplification reactions were conducted for 40 cycles at 95 °C for 15 s and at 60 °C for 1 min. The EPO primer and probe set (GACAGGTGGTTCTCCT, FAM-TGAGCCAAGGACAGGAAGAC, TAMRA) were used to measure the human EPO and ZFP expression levels, respectively. The GAPDH primer and probe set (GACAGATGTCCAGGATTGCT, TAMRA) were used to monitor the internal control GAPDH mRNA. The abundance of expressed ZFPs was controlled for at both the RNA level as described above, but also by Western blotting. This analysis was performed by resolving 10 μg of whole cell lysates on a 10–20% Tris-HCl polyacrylamide gel containing SDS. Proteins were transferred onto a nitrocellulose membrane using 1× SDS, 20% Methanol, and then the filter was blocked using 5% nonfat dry milk for 1 h at room temperature. Blotting was done with anti-Flag M2 monoclonal antibody (Sigma) diluted 1:1000 in 5% (w/v) nonfat dry milk, 0.1% PBS-Tween for 1 h at room temperature. Subsequently, an anti-mouse horseradish peroxidase conjugate (Amersham Pharmacia Biotech) was used at a 1:3000 dilution in 5% (w/v) nonfat dry milk, 0.1% PBS-Tween for 1 h at room temperature. All washes were done in 0.1% PBS-Tween. The protein bands were detected by the ECL system (Amersham Pharmacia Biotech).

Endogenous EPO expression was assayed either in response to transient transfection with ZFP effectors or in response to the induction of pEPOZFP–862a following stable transfection. Assays were performed at the indicated time after transfection. Either RNA was extracted as indicated above, or the culture medium was harvested for measurement of secreted EPO protein using the human EPO ELISA kit (R&D Systems, Minneapolis, MN).

Chromatin Analysis, Chromatin Immunoprecipitation, DNase I, and Micrococcal Nuclease Mapping—Chromatin immunoprecipitation was performed using a ChIP assay kit according to the instructions from the manufacturer (Upstate Biotechnology, Inc., Lake Placid, NY). Approximately 2 million cells were cross-linked with 1% formaldehyde for 10 min, washed with PBS, and lysed in lysis buffer. The cell lysate was sonicated on ice, resulting in a DNA fragment length of approximately 200 bp. After removing cell debris by centrifugation, immunoprecipitation was performed in ChIP dilution buffer overnight with 3 μg of VP16 (Cruz) or 10 μg of anti acetylated H3 antibody (Upstate Biotechnology). The antibodies were collected with Protein A-Sepharose (Sigma) and washed with PBS-Tween.

RESULTS

Characterization of Synthetic ZFPs—We used mouse embryos to perform experiments concerning the recognition of specific sequences by zinc finger domains to design 10 proteins that we predicted would recognize sequences in the 5′ flank of the human EPO gene. The strategy for these experiments is as follows: (1) we isolated the 5′ flank of the human EPO gene and determined its sequence; (2) we designed a range of zinc finger DNA binding domains that exhibit specific DNA sequence binding motifs, and consequently bind to our preferred 5′ flank of the human EPO gene; (3) we synthesized the DNA sequences corresponding to each zinc finger domain; and (4) we characterized the zinc finger domains by gel retardation assays to quantitate their ability to bind to DNA. These results are tabulated in Table II. Our results indicate that we have designed a range of zinc finger DNA binding domains that exhibit specific DNA sequence binding motifs in the 5′ flank of the EPO gene. These affinities range from dissociation constants of 0.23 nM (EPOZFP–535) to greater than 20 nM (EPOZFP–72). We next wished to relate these zinc finger domains and

| Name | Target sequence | Kd (nM) |
|------|-----------------|---------|
| EPOZFP–862a | GCGTGCTTCCTc | 6 |
| EPOZFP–862b | GCGTGCTTCCTc | 9 |
| EPOZFP–862c | GCGTGCTTCCTc | 2 |
| EPOZFP–545a | GTGAGGACAG | 7.5 |
| EPOZFP–545b | GTGAGGACAG | 7 |
| EPOZFP–535a | GAGTGCTGAGc | 0.23 |
| EPOZFP–233 | GAGTGCTGAGc | 12 |
| EPOZFP–82 | GTGGGGGAGGa | 12 |
| EPOZFP–72 | GTGGGGGAGGa | 22.5 |
| 5h | GGCGGCTGGg | 40 |
| ZF868 | CTCTGGCGGCL | 2 |

For 5 min at 22 °C, DNase I or micrococcal nuclease (Worthington) was added at 0, 5, 10, 20, or 40 units/ml for nuclei or 100-fold lower concentrations for naked DNA. The reaction was stopped by adding Qiagen buffer AL, followed by proteinase K treatment and purification (DNaseasy kit, Qiagen). The recovered DNA was digested to completion overnight at 37 °C with the indicated restriction endonuclease and then concentrated by ethanol precipitation. The entire sample is loaded onto an agarose gel of 1–2%, before electrophoresis and transfer to a Nitrocellulose membrane (Schleicher & Schuell). Indirect end labeling was used to detect the sites of nuclease cleavage. Hybridization was carried out in Rapid Hyb buffer (Amersham Pharmacia Biotech) using PCR-amplified genomic DNA fragments as indicated. The DNA probes were radiolabeled with [α-32P]dCTP to a specific activity of 10 000 cpm/mg DNA. Hybridization and subsequent washings were performed at 65 °C. Membranes were visualized in the PhosphorImager.
FIG. 3. The transcriptional activation properties of the EPOZFPs. 

A, luciferase reporter assays involved the co-transfection of the ZFP expression construct (50 ng) together with the respective luciferase reporter DNA (900 ng) (see "Experimental Procedures") and CMV β-gal as an internal control into HEK293 cells. After 40 h, cell lysates were harvested and the luciferase activity induced by each EPOZFP measured as indicated. The pcDNAEmpty and −862c-NF provide controls using proteins lacking the zinc fingers and VP16 activation domain, respectively. B, ELISA assays on endogenous EPO expression. EPO protein synthesis was assayed at 40 h after expression of the various EPOZFPs indicated as described (see "Experimental Procedures"). SBS 579 contains an expression vector for a nonspecific ZFP. C, quantitation of EPO mRNA by real time PCR (Taqman). EPO mRNA was assayed by real time PCR as described (see "Experimental Procedures") after expression of the various EPOZFPs for 40 h as described. SBS 579 contain an expression vector for a nonspecific ZFP. D, quantitation of the various EPOZFP mRNAs by real time PCR. The mRNA synthesized from the various transiently transfected EPOZFP expression constructs and their controls is quantitated. E, immunoblot of protein synthesized from the various transfected EPOZFPs and their controls.
their recognition of these diverse sequences to their capacity to activate transcription from the EPO promoter when fused to a luciferase reporter.

**Transcriptional Regulation of the EPO Promoter Using ZFPs**—Intensive analysis has defined the key cis-acting elements for regulation of the human EPO promoter. Erythropoietin is a protein hormone produced primarily in the kidneys and liver which controls the biogenesis of erythrocytes. Expression of the EPO gene is primarily controlled in a tissue-specific manner that is sensitive to oxygen tension. Hypoxia induces EPO gene expression at the transcriptional level through the action of a hypoxia response element, which interacts with HIF-1, a heterodimer of the PAS-family transcription factor HIF-1α and ARNT, the aryl hydrocarbon receptor nuclear translocator (31). The hypoxia response element is located at the 3' flank of the EPO gene (31–33). A homodimer of HNF-4 also binds at this site (32, 33). The 5' flank of the EPO gene has also been determined to influence transcription; however, the only known regulatory elements that contribute to EPO gene activity within this DNA segment are located proximal to the start site of transcription (between −61 and −45 relative to the start at +1) (30). All ZFP transcription factors activate the luciferase reporter gene (Fig. 3A). These constructs contain three reiterations of the ZFP recognition site upstream of the
SV40 minimal promoter. Expression of a ZFP that lacks the VP16 activation domain does not activate luciferase (Fig. 3A, NVF). The range of activation varies between 2-fold (EPOZFP–862) and 16-fold (EPOZFP–862c). Surprisingly, in the data shown in Fig. 3A and in all subsequent repetitions of this experiment, there is no strong correlation in these transient assays between the capacity of a ZFP to activate transcription and the measured binding affinity (Table II). Both the strong (EPOZFP–535), \( K_d = 0.23 \text{ nM} \) and weak (EPOZFP–72, \( K_d = 22.5 \text{ nM} \)) binders activate transcription with comparable efficiency. Although these results indicate that our basic design of ZFP transcription factors is robust, there are several variables to interpretation of these data. The newly synthesized ZFPs might have varying stability, or be deficient in entry into the nucleus, or might be engaged in non-productive protein-protein interactions with inhibitory factors. We have controlled for these possibilities in various ways. Western blotting demonstrates that all of the ZFPs accumulate to high levels in cells (see Fig. 3D) and the cells accumulate comparable levels of ZFP encoding mRNA (see Fig. 3D). In other experiments, we have established that the ZFPs accumulate uniformly in nuclei and that they retain identical capacities within nuclear extracts prepared from these cells to recognize that DNA sequences to which they bind in vitro (data not shown).

In order to further test the relationship between DNA binding affinity and the capacity of the ZFPs to activate the EPO gene by virtue of their binding at a particular site, we established an inducible cell line expressing the most effective activating ZFP in our transient expression assays, EPOZFP–862. Our strategy was to transfect pTO-EPOZFP–862c into the T-Rex–292TM cell line (see “Experimental Procedures”), which stably expresses the Tet repressor and allows for the regulated expression of a gene of interest under the control of a Tet-responsive promoter when doxycycline is added to the culture medium. Twenty-one stable clones were isolated and analyzed for ZFP expression and erythropoietin expression in a Dox-dependent manner (Fig. 4, A and B). We used Northern blotting to confirm that the endogenous human EPO chromosomal locus was activated to synthesize a full-length poly(A) mRNA in response to either transient or stable expression of EPOZFP–862 (Fig. 4C).

We next investigated the activation of a luciferase reporter to which was fused at the same site relative to the SV40 minimal promoter three reiterations of the EPOZFP–862 recognition sequence and mutant forms of this sequence to which EPOZFP–862 has reduced binding affinity (Fig. 5A). Under these defined conditions, there is a strong correlation between the ZFP affinity for a particular sequence and the capacity to activate transcription (Fig. 5B). We suggest that, within one defined chromosomal or chromatin context, a particular ZFP will regulate transcription in a manner that largely reflects relative binding affinity for the cognate sequence. Transcriptional activation occurs most efficiently when the \( K_d \) is progressively reduced from 10 nM and is relatively inefficient when the \( K_d \) exceeds 30 nM (Fig. 5B). This contrasts with the lack of correlation between binding affinity and transcriptional response apparent when different ZFP binding sites relative to the transcription start site are used in luciferase reporter as-
says and in activation experiments on the endogenous chromosomal locus (Table II, Fig. 3, A–C).

We next compared the capacity of ZFPs that would recognize sequences present in the 5′ flank of the EPO gene in vitro on naked DNA and in vivo on transiently transfected DNA, to activate the endogenous chromosomal locus. In Fig. 3 (A–C), we compare the capacity of ZFPs to function in the luciferase reporter assay with their capacity to activate the endogenous EPO gene as assayed both by quantitative real time PCR (Taqman) and by ELISA assays. We find that, in contrast to the pleiotropic activation of the transiently transfected reporter construct (Fig. 3A), activation of the endogenous chromosomal EPO promoter is found with those ZFPs targeted to one selected target in the promoter (Fig. 3, B and C). This is the binding site for EPOZFP−862. Three individual protein designs with varying affinity for DNA (EPOZFP−862a, −b, and −c) activate the endogenous chromosomal EPO gene.

Our evidence for the EPOZFP−862a, −b, and −c proteins functioning at Site 1 (Figs. 2 and 3) led us to employ a more direct strategy to test our hypothesis that these ZFPs were gaining access to the chromatin infrastructure at this site. Chromatin immunoprecipitation is a powerful method to demonstrate selective enrichment of a particular regulatory protein at a particular site in chromatin (34). We chose to test for enrichment of the EPOZFP−862 protein both at site 1 and at a second site of perfect identity located at +6826 relative to the transcription start site. After cross-linking and immunoprecipitation using antibodies against VP16 (see “Experimental Procedures”), we made use of three primer pairs centered on 5′-flanking regions −1838, −927, −363 and two sets of primer pairs centered on sequences downstream from the start site of transcription at +538 and the other site of perfect identity.

FIG. 6. The promoter of the human EPO gene is not hypersensitive, but the 3′-flanking sequence is hypersensitive to DNase I cleavage. A, a scheme showing restriction sites (X = XbaI), probes (hatched boxes), coding region (solid rectangle), and hypoxia response element (gray box). B, nuclei were treated with DNase I (see “Experimental Procedures”) at 0, 5, 10, or 20 units/ml (lanes 1–4), then digested with XbaI and probed with a fragment abutting −3673. Positions of DNA sizes relative to the transcription start are shown. C, the blot shown in panel B was stripped of probe and rehybridized to detect DNase I-hypersensitive sites at the 3′-distal region. The DNA fragment that anneals near the +791 XbaI site was used as a probe. Relative positions of DNase I-hypersensitive sites (+4591 and +4091) plus the position of the hypoxic response element are indicated.
The abundance of these DNA sequences was quantified relative to a GAPDH internal control. Our results show a 30–70-fold enrichment between the $-927$ and $-363$ sites for the EPOZFP$^{2862}$ protein (Fig. 7A). This distribution might be expected due to our use of DNA fragments of 500 bp in size in our chromatin immunoprecipitation assays. Both the $-927$ and $-363$ fragments would contain the $-862$ site within this size distribution. An additional important control is the loss of recovery of DNA fragments at both the 5′ ($-1838$) and 3′ ($+538$) regions flanking the $-862$ site (Fig. 7A). Finally, it is noteworthy that an identical sequence, GCGGTGGCT, for recognition by EPOZFP$^{2862}$ positioned at $+6839$ is not recovered at all in this experiment. (Fig. 7A). We conclude that parameters other than DNA sequence per se determine the access of the EPOZFP$^{2862}$ protein to the cognate sequence.

In order to better understand the access of EPOZFP$^{−862}$ protein to the endogenous chromosomal EPO locus, we made use of micrococcal nuclease digestion of chromatin. This enzyme is considerably smaller than DNase I and gains efficient access to the linker DNA between nucleosomes (35). Sequence analysis of the 5′ flank of the EPO gene indicates the presence of Alu elements, which are known to position nucleosomes (28, 29). The EPOZFP$^{−862}$ site is at the 5′ of the Alu element (Fig. 7A), which we would predict would lie in linker sequence adjacent to the nucleosome core. To test this hypothesis, we digested chromatin with micrococcal nuclease and then used indirect end labeling to determine the positions of nucleosomes. We find that nucleosomes are positioned such that linker DNA is accessible at $-490$ and $-670$ with a major region of micrococcal nuclease sensitivity beginning at $-852$. Thus, a domain of micrococcal nuclease sensitivity corresponds to a site of responsiveness to EPOZFP$^{−862}$.

As a test of the capacity of the EPOZFP$^{−862}$ protein to initiate events leading to chromatin remodeling, we examined the chromatin structure of the endogenous EPO gene in the cell line that induced expression of EPOZFP$^{−862}$ in response to doxycycline. In the absence of EPOZFP$^{−862}$ expression, there is no significant expression of the endogenous EPO gene (Figs. 3 and 4) and there is no DNase I hypersensitivity over the promoter (Fig. 6, A and B). In the presence of doxycycline, the endogenous EPO locus is transcriptionally activated (Figs. 3 and 4) and a DNase I-hypersensitive site appears over the promoter (Fig. 8, A and B, lanes 1 and 2). In the absence of doxycycline, no DNase I-hypersensitive site appears (Fig. 8, A and B, lanes 3 and 4). These results confirm the capacity of the synthetic transcriptional regulators to remodel chromatin in a targeted manner.
The primary conclusion from these experiments is that the activity of synthetic zinc finger transcription factors at an endogenous chromosomal locus (Figs. 3 (B and C) and 4) differs significantly from the determinants of protein binding to naked DNA (Fig. 2 and Table II) and from the capacity of these diverse proteins to activate transiently transfected templates (Figs. 3A and 5). Although there is little difficulty in activating transcription from transiently transfected templates (Fig. 3A), there are constraints in gaining access to sites within endogenous chromosomes in order to regulate expression (Figs. 3 (B and C) and 6–8). These observations are consistent with earlier work on the mouse mammary tumor virus promoter in which the access of the transcriptional regulators NF1 and Oct 1 is dependent on the nature of the chromatin infrastructure (36–38). The prior assembly of nucleosomes on the NF1 and Oct 1 binding sites precludes their access to the MMTV promoter, the chromatin assembled on transiently transfected templates poses no measurable impediment to the transcription of synthetic zinc finger transcription factors at an endogenous chromosomal locus (Figs. 3 and 7). There are no known transcriptional regulatory elements that normally contribute to EPO gene expression at this site (30–34). In fact, this region of the EPO gene is refractory to DNase I cleavage consistent with the lack of regulatory DNA (Figs. 6 and 8). Thus the targeting of EPOZFP–862 to a linker region flanking a nucleosome positioned on an Alu element generates a novel cis-acting element (Figs. 3 and 7).

These observations suggest an improved strategy for dissecting the requirements for eukaryotic gene control. First, it is important to determine the underlying chromatin infrastructure and identify linker DNA segments between positioned nucleosomes. Second, it should be possible to design ZFPs that will recognize these accessible sequences such that they bind stably in chromatin (Figs. 7 and 8). Third, it would then be possible to systematically recruit a variety of transcriptional activation or repression domains to that particular site in order to remodel chromatin and regulate transcription. In this way, it should be possible to manipulate human genes in their endogenous chromosomal context and more effectively design ZFPs with therapeutic potential.

Acknowledgments—We thank Michelle Ha, Mike Kunis, Yolanda Santiago, Priya Sreenivasan, Xiaohong Zhong, and Susanne Zhu for expert assistance. We are grateful to Edward Lanphier, Peter Bluford, and Carl Pabo for encouragement.

REFERENCES

1. Desjarais, J. R., and Berg, J. M. (1992) Proteins Struct. Funct. Genet. 12, 101–104
2. Desjarais, J. R., and Berg, J. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2256–2260
3. Rebar, E. J., and Pabo, C. O. (1994) Science 263, 671–673
4. Jamieson, A. C., Kim, S.-H., and Wells, J. A. (1994) Biochemistry 33, 5689–5695
5. Choo, Y., and Klug, A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 11163–11167
6. Wu, H., Yang, W.-P., and Barbas, C. F., III (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 344–348
7. Jamieson, A. C., Wang, H., and Kim, S.-H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 12834–12839
8. Greisman, H. A., and Pabo, C. O. (1997) Science 275, 657–661
9. Choo, Y., Sanchez-Garcia, I., and Klug, A. (1994) Nature 372, 642–645
10. Kim, J.-S., and Pabo, C. O. (1997) J. Biol. Chem. 272, 29795–29800
11. Kim, J.-M., Kim, J., Cepek, K. L., Sharp, P. A., and Pabo, C. O. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 3616–3620
12. Beeri, R. R., Segal, D. J., Dreier, B., and Barbas, C. F., III (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 14628–14633
13. Beeri, R. R., Dreier, B., and Barbas, C. F., III (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 1495–1500
14. Wolfe, A. P., and Hayes, J. J. (1999) Nucleic Acids Res. 27, 711–720
15. Gottesfeld, J., and Bloomer, L. S. (1982) *Cell* **28**, 781–791
16. Rhodes, D. (1985) *EMBO J.* **4**, 3473–3482
17. Tremethick, D., Zucker, D., and Worcel, A. (1990) *J. Biol. Chem.* **265**, 5014–5023
18. Ura, K., Hayes, J. J., and Wolffe, A. P. (1995) *EMBO J.* **14**, 3752–3765
19. Panetta, G., Buttinelli, M., Flaus, A., Richmond, T. J., and Rhodes D. (1998) *J. Mol. Biol.* **282**, 683–697
20. Hayes J. J., and Wolffe A. P. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 1229–1233
21. Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1990) *Cell* **72**, 73–84
22. Vitolo, J. M., Thiriet, C., and Hayes, J. J., (2000) *Mol. Cell. Biol.* **20**, 2167–2175
23. Bouret, P., Dimitrov, S., and Wolffe, A. P. (1994) *Genes Dev.* **8**, 1147–1159
24. Chipev, C. C., and Wolffe, A. P. (1992) *Mol. Cell. Biol.* **12**, 45–55
25. Sera, T., and Wolffe, A. P. (1998) *Mol. Cell. Biol.* **18**, 3668–3680
26. Tavazoie, S. M., Kingsbury, R. C., and McKnight, S. L. (1996) *Genes Dev.* **2**, 718–729
27. Englander, E. W., Wolffe, A. P., and Howard, B. H. (1993) *J. Biol. Chem.* **268**, 19565–19573
28. Reeves, R., Gorman, C. M., and Howard, B. H. (1996) *J. Biol. Chem.* **271**, 10405–10412
29. Fryer, C. J., and Archer, T. K. (1998) *Trends Genet.* **14**, 477–480
30. Drew, H. R. (1984) *J. Mol. Biol.* **176**, 535–557
31. Archer, T. K., Lefebvre, P., Woldford, R. G., and Hager, G. L. (1992) *Science* **255**, 1573–1576
32. Archer, T. K., Cordingley, M. G., Woldford, R. G., and Hager, G. L. (1991) *Mol. Cell. Biol.* **11**, 688–698
33. Reeves, R., Gorman, C. M., and Howard, B. H. (1985) *Nucleic Acids Res.* **13**, 3599–3615
34. Englemeier, E. W., and Wolffe, A. P. (1995) *Cell* **72**, 73–84
35. Vitolo, J. M., Thiriet, C., and Hayes, J. J., (2000) *Mol. Cell. Biol.* **20**, 2167–2175
36. Bouret, P., Dimitrov, S., and Wolffe, A. P. (1994) *Genes Dev.* **8**, 1147–1159
37. Chipev, C. C., and Wolffe, A. P. (1992) *Mol. Cell. Biol.* **12**, 45–55
38. Tavazoie, S. M., Kingsbury, R. C., and McKnight, S. L. (1996) *Genes Dev.* **2**, 718–729
39. Englander, E. W., Wolffe, A. P., and Howard, B. H. (1993) *J. Biol. Chem.* **268**, 19565–19573
40. Reeves, R., Gorman, C. M., and Howard, B. H. (1996) *J. Biol. Chem.* **271**, 10405–10412
41. Fryer, C. J., and Archer, T. K. (1998) *Trends Genet.* **14**, 477–480
42. Drew, H. R. (1984) *J. Mol. Biol.* **176**, 535–557
43. Archer, T. K., Lefebvre, P., Woldford, R. G., and Hager, G. L. (1992) *Science* **255**, 1573–1576
44. Archer, T. K., Cordingley, M. G., Woldford, R. G., and Hager, G. L. (1991) *Mol. Cell. Biol.* **11**, 688–698
45. Fryer, C. J., and Archer, T. K. (1998) *Nature* **393**, 88–91