Ligand-dependent transcriptional regulators were generated by fusion of designed Cys2-His2 zinc finger proteins and steroid hormone receptor ligand binding domains. To produce novel DNA binding domains, three-finger proteins binding specific 9-base pair sequences were constructed from modular building blocks. Fusion of these zinc finger proteins to a transcriptional activation domain and to modified ligand binding domains derived from either the estrogen or progesterone receptors yielded potent ligand-dependent transcriptional regulators. Together with optimized minimal promoters, these regulators provide 4-hydroxytamoxifen- or RU486-inducible expression systems with induction ratios of up to 3 orders of magnitude. These inducible expression systems are functionally independent, and each can be selectively switched on within the same cell. The potential use of zinc finger-steroid receptor fusion proteins for the regulation of natural promoters was also explored. A gene-specific six-finger protein binding an 18-base pair target sequence was converted into a ligand-dependent regulator by fusion with either two estrogen receptor ligand binding domains or one ecdysone receptor and one retinoid X receptor ligand binding domain. These single-chain receptor proteins undergo an intramolecular rearrangement, rather than intermolecular dimerization and are functional as monomers. Thus, the ability to engineer DNA binding specificities of zinc finger proteins enables the construction of ligand-dependent transcriptional regulators with potential for the regulation of virtually any desired artificial or natural promoter. It is anticipated that the novel chemically regulated gene switches described herein will find many applications in applied and basic research, where the specific modulation of gene expression can be exploited.

Designed transcription factors with defined target specificity and regulatory function could provide invaluable tools for basic and applied research and for gene therapy. Accordingly, the design of sequence-specific DNA binding domains has been the subject of intense interest for the last 2 decades. Of the many classes of DNA-binding proteins studied, the modular Cys2-His2 zinc finger DNA binding motif has shown the most promise for the production of proteins with tailored DNA binding specificity (1–7). The novel architecture of this class of proteins provides for the rapid construction of gene-specific targeting devices. Polydactyl zinc finger proteins are most readily prepared by assembly of modular zinc finger domains recognizing predefined three-nucleotide sequences (6, 8, 9). Polydactyl proteins may be assembled using variable numbers of zinc finger domains of varied specificity providing DNA-binding proteins that not only recognize novel sequences but also sequences of varied length. By combining six zinc finger domains, proteins have been produced that recognize 18 contiguous base pairs of DNA sequence, a DNA address sufficiently complex to specify any locus in the 4 billion-base pair human genome (or any other genome). Fusion of polydactyl zinc finger proteins of this type to activation or repression domains provides transcription factors that efficiently and specifically modulate the expression of both transgenes and endogenous genes (8, 9). While the availability of designed transcription factors with tailored DNA binding specificities provides novel opportunities in transcriptional regulation, additional applications would be available to ligand-dependent transcriptional factors. Designer zinc finger proteins dependent on small molecule inducers would have a number of applications, both for the regulation of endogenous genes and for the development of inducible expression systems for the regulation of transgenes.

Natural transcription factors are regulated by a number of different mechanisms, including posttranslational modification such as phosphorylation (10, 11), or by ligand binding. The prototype ligand-activated transcription factors are members of the nuclear hormone receptor family, including the receptors for sex steroids or adrenocorticoids (12, 13). These receptors are held inactive in the absence of hormone, by association with a number of inactivating factors including hsp90 (14). Upon ligand binding, nuclear hormone receptors dissociate from the inactivating complex, dimerize, and become able to bind DNA and activate transcription (12–14). Significantly, not only hormone binding but also inactivation and dimerization functions reside within the ligand binding domain (LBD)1 of these proteins (15). This fact has been exploited experimentally and steroid hormone receptor LBDs have found wide use as tools to render heterologous proteins hormone-dependent. In particular, the estrogen receptor (ER) LBD has been used to render the functions of c-Myc (16), c-Fos (17), and even the cytoplasmic kinase c-Raf (18) hormone-dependent.

To develop an inducible expression system for use in basic

1 The abbreviations used are: LBD, ligand binding domain; 4-OHT, 4-hydroxytamoxifen; DBD, DNA binding domain; ErR, ecdysone receptor; PR, progesterone receptor; RXR, retinoid X receptor; bp, base pair(s); aa, amino acid(s); BSA, bovine serum albumin; ELISA, enzyme-linked immunosorbent assay.
research and gene therapy, the availability of ligand-dependent transcriptional regulators is a prerequisite. Preferentially, these regulators would be activated by a small molecule inducer with no other biological activity, bind specific sequences present only in the target promoter, and have low immunogenicity. A number of ligand-regulated artificial transcription factors have been generated by various means, using functional domains derived from either prokaryotes (19–22) or eukaryotes (23–29). Of the functional domains derived from eukaryotic proteins, nuclear hormone receptor LBDs have been the most widely used. In particular, regulators based on the Gal4 DNA binding domain (DBD) fused to a human ER (27, 28) or progestosterone receptor (PR) LBD (25, 26), as well as the ecdysone-inducible system based on the Drosophila ecdysone receptor (EcR) and the mammalian retinoid X receptor (RXR) (23, 24) allows ligand-dependent regulators to be directed to any desired artificial or natural promoter. Here we explore the utility of these regulators would be activated by a small molecule in-
Chemically Regulated Zinc Finger Transcription Factors

Single-Chain Switch Constructs—For construction of single-chain fusions with two ER LBDs, the point-mutated mouse ER LBD was amplified from pBSKS+ER (30) either using the primers ERFse-F and ERSpe-B (5'-GAG GAG GAG GAG GAG AGT AGT AGT ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC ACC AGA GGA GAT GGC TGG GCG AGT CGA CTC TAG AGG GTA TAT AAT GG-3') or ERNhne-F1 (for the 18-aa linker construct; 5'-GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GGT GGC GGT GGC TCT TCT TCC AAT GAA ATG GGT TCT GCA GAG GAC-3') or ERNhne-F2 (for the 30-aa linker construct; 5'- GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GGT GGC GGT GGC TCT TCT TCC AAT GAA ATG GGT TCT GCA GAG GAC-3') and ERAsc-B. The PCR products were then digested with, respectively, FseI and SpeI, or Nhel and AscI, and inserted into FseI–AscI-linearized pcDNA3/E2C-VP64 (8).

For construction of RXR-EcR single-chain fusions, the ligand binding domain of the human retinoid X receptor (hRXRα; aa 373–654) was PCR-amplified from pVgRXR (Invitrogen) using the primers RXRFse-F and RXRSpe-B (5'-GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GAG GGT GGC GGT GGC TCT TCT TCC AAT GAA ATG GGT TCT GCA GAG GAC-3'). The RXR LBD was amplified from the plasmid pGL3-Basic (CLONTECH) and used to replace the luciferase protein, a series of reporter plasmids was constructed. Each protein contains an N-terminal VP64 activation domain (8) and a transcriptional activation domain plus the LBD of either of the two nuclear hormone receptors. The VP64-C7-PR fusion construct was kindly provided by Dr. Price. The VP64-C7-ER fusion construct was kindly provided by Dr. Price. The VP64-C7-ER fusion construct was kindly provided by Dr. Price.

Reporter Constructs for Determining the Optimal Spacing and Orientation of the Two Half-sites—C7 dimer-TATA fragments were generated by PCR amplification with C7 dimer-TATA primers (5'-CGT CCA GAT CTC CTC GAG-3') and 10xN1-TATA fragments were assembled from two pairs of complementary oligonucleotides each and cloned into 5'-GAG GCT AGC TCT TCC GGT GGC GGC CAA GAC TTT GTT AAG AAG GTC CCA GAT TTG GCG CTC CTC GAG C-3'). The PCR products were then digested with, respectively, FseI and SpeI, or Nhel and AscI, and inserted into FseI–AscI-linearized pcDNA3/E2C-VP64 (8). DNA binding domains were exchanged via SfiI digestion, and effector domains were exchanged via BglII digestion.

To generate the 36-aa linker E2C-RLLE-VP64 fusion construct, the site of insertion was engineered into the vector by PCR amplification with C7 dimer-TATA primers (5'-CGT CCA GAT CTC CTC GAG-3') and 10xN1-TATA fragments were assembled from two pairs of complementary oligonucleotides each and cloned into 5'-GAG GCT AGC TCT TCC GGT GGC GGC CAA GAC TTT GTT AAG AAG GTC CCA GAT TTG GCG CTC CTC GAG C-3'). The PCR products were then digested with, respectively, FseI and SpeI, or Nhel and AscI, and inserted into FseI–AscI-linearized pcDNA3/E2C-VP64 (8). DNA binding domains were exchanged via SfiI digestion, and effector domains were exchanged via BglII digestion.

Generation of Hormone-regulated Zinc Finger-Steroid Receptor Fusion Proteins—Previous studies have demonstrated the potential of engineered C9-H2 zinc finger proteins for the regulation of target gene expression (8, 9, 34, 35). However, to fully realize the potential of engineered zinc finger proteins, it is desirable that their otherwise constitutive DNA binding activity be rendered ligand-dependent. The LBDs of the human PR and the murine ER have previously been used for the regulation of heterologous proteins, after having been modified to remove their responsiveness to their natural hormone inducers while retaining activity with synthetic antagonists (26, 30). In our initial study, the Zif268 variant C7 (4) was fused to a transcriptional activation domain plus the LBD of either of the two nuclear hormone receptors. The VP64-C7-PR fusion protein contains an N-terminal VP64 activation domain (8) and a C-terminal human PR LBD (aa 645–914) lacking amino acids 915–933. This LBD is responsive to the progesterone antagonist RU486/mifepristone but not to progesterone (26). The VP64-C7-ER fusion protein contains a C-terminal murine ER LBD (aa 262–599) with a single amino acid substitution (G525R) and is responsive to the estrogen antagonist 4-OHT but not to estrogen (30) (Fig. 1A). Preliminary experiments confirmed that fusion to either LBD rendered the C7 protein hormone-dependent (data not shown).

Determination of the Optimal Response Element for Zinc Finger-Steroid Receptor Fusion Proteins—Naturally occurring steroid receptors bind DNA as dimers and typically recognize response elements consisting of palindromic sequences (12, 13). Furthermore, it was demonstrated that in some cases direct repeats can also serve as binding sites for receptor dimers (36). Given this obvious flexibility in DNA recognition by the naturally occurring receptor dimers, the optimal structure of a response element for an artificial, zinc finger-based transcriptional switch was not known. In order to develop an efficient, hormone-inducible system for the regulation of target gene expression, a detailed knowledge of the binding site architecture is required.

To determine the optimal orientation and spacing of the two half-sites of a response element for a zinc finger-LBD fusion protein, a series of reporter plasmids was constructed. Each contained two C7 binding sites upstream of a TATA box and a firefly luciferase coding region (Fig. 1B). The two C7 binding sites were introduced in different orientations (direct, inverted or everted repeat) and with various spacings (no spacing or 1–5 bp spacing). Plasmids directing expression of VP64-C7-PR or VP64-C7-ER fusion constructs were then co-transfected with the various reporter plasmids and assayed for hormone-induced luciferase expression (Fig. 2). Significantly, each of the C7 dimer binding sites was able to act as a response element for...
tors and the indicated reporter plasmids. 24 h after transfection, cells of the half-sites of an artificial response element. co-transfected with VP64-C7-PR (\(6^{\text{th}}\) value was measured 48 h after transfection. Each 

direct repeat; with 4–5-bp spacing or no spacing at all (Fig. 2

element; everted repeat.
both PR- and ER-based proteins, albeit with varying efficiency. In contrast, a reporter plasmid with a single C7 binding site was not activated, indicating that hormone-induced activation of transcription was mediated by dimerized transcription factors (data not shown).
Optimal spacing depended on the orientation of the two half-sites. In the case of the PR fusion protein, optimal spacing was observed at 2–3 bp for inverted repeats and 3 bp for everted repeats. Response elements consisting of direct repeats had no single optimal spacing; the best response was obtained with 4–5-bp spacing or no spacing at all (Fig. 2A). For the ER fusion protein, optimal spacing was at 3–4 bp for direct repeats, 1–2 bp for inverted repeats, and 3 bp for everted repeats (Fig. 2B). It should be noted that there were significant variations in the basal, i.e. ligand-independent, activity of PR and ER fusion proteins, depending on the response element tested. Most notably, increasing the spacing of direct repeats from 3 to 4 bp led to a 1.9-fold higher basal activity of VP64-C7-PR and a 3.7-fold increase in the case of VP64-C7-ER. High basal activity is extremely undesirable for an inducible promoter system, where tight control over the expression levels of a particular gene of interest is often required, especially if the gene product is toxic. Thus, in choosing appropriate response elements, particular attention must be paid not only to hormone inducibility but also to its basal activity in the presence of the regulatory protein. The response element consisting of direct repeats with a spacing of three nucleotides was considered to be a good choice for use in hormone-inducible artificial promoters, since it was compatible with both PR and ER fusion proteins. Significantly, basal activity of these promoters in the presence of either PR or ER fusion proteins was among the lowest of all response elements tested. Furthermore, good hormone-induced activation of transcription was observed with both VP64-C7-PR (3.9-fold) and VP64-C7-ER (9.5-fold).

**Generation of Novel DNA Binding Domains**—While the use of the C7 DNA binding domain was well suited for the preliminary studies described above, it may not be the best choice for incorporation into an inducible transcriptional regulator. The C7 protein is a variant of the mouse transcription factor Zif268 (37), with increased affinity but unchanged specificity for a Sp1-type site (4). We reasoned that the use of alternate DNA binding domains would minimize potential pleiotropic effects of the chimeric regulators. Previously, we described a strategy for the rapid assembly of zinc finger proteins from a family of predefined zinc finger domains specific for each of the 16 5’-GNN-3’ DNA triplets (6, 8). Three-finger proteins binding any desired 5’-(GNN)_3-3’ sequence can be rapidly prepared by grafting the amino acid residues involved in base-specific DNA recognition into the framework of the consensus three-finger protein Sp1C (33). To date, well over 100 three-finger proteins have been produced in our laboratory. Two of these, B3 and N1, were chosen for use in the construction of inducible transcriptional regulators (Fig. 3A). The B3 and N1 proteins are designed to bind the sequences 5’-GGA CCC CCC-GAC-3’ or 5’-GGG GGG GAA-3’, respectively. To verify their DNA binding specificity, these proteins were purified as MBP fusions and tested by ELISA analysis using an arbitrary selection of oligonucleotides containing 5’-(GNN)_3-3’ sequences (Fig. 3B). Significantly, both proteins recognized their target sequence and showed no cross-reactivity to any of the other 5’-(GNN)_3-3’ sequences tested. However, as judged by ELISA, binding of N1 was much weaker than binding of B3. Therefore, affinities were determined by electrophoretic mobility shift analysis. The B3 protein bound its target sequence with a \(K_D\) value of 15 nM, similar to the \(K_D\) values we previously reported for other three-finger proteins (8). In contrast, the N1 protein displayed an affinity for its target that was dramatically lower, and we estimated its \(K_D\) value to be in the range of 5–10 \(\mu\)M. The fact that the two proteins had very different affinities for their respective target sequences allowed us to investigate the influence of affinity on the functionality of the inducible expression systems.

**RU486- and 4-OHT-inducible Systems for the Control of Gene Expression**—To allow for a comparative analysis, a series of RU486- or 4-OHT-inducible transcriptional regulators were constructed containing either the B3 or the N1 DNA binding domain (Fig. 4A). The role of placement of the activation domain was investigated, by fusing it either to the N or the C terminus of the protein. Two different activation domains were compared: the herpes simplex virus VP16 transactivation domain (38) and the synthetic VP64 activation domain, which consists of four tandem repeats of VP16’s minimal activation domain (8).

Synthetic promoters were constructed based on the B3 and N1 DNA target sequences and the optimal response element structure defined above (Fig. 4B). The 10xB3-TATA-luc and 10xN1-TATA-luc plasmids each contain five response elements, consisting of direct repeats spaced by three nucleotides, upstream of a TATA box and a firefly luciferase coding region. The response elements are separated from each other by six nucleotides and should provide for the binding of five dimers and maximal promoter activity.

The activity of the various fusion constructs was assessed by transient cotransfection studies with the cognate TATA reporter plasmids in HeLa cells (Figs. 5 and 6; Table I). In general, the ER fusion proteins were more potent transactivators, and 4-OHT-induced luciferase activity was typically 3–7 times higher than RU486-induced luciferase activity mediated by the PR fusion proteins (compare Figs. 5 and 6). However, since the basal (i.e. ligand-independent) activity of ER chimeras was often somewhat higher, their hormone-induced -fold

![Figure 2](image-url)
stimulation was not generally better. Hormone-dependent gene activation in excess of 2 orders of magnitude was commonly observed with both PR and ER fusion proteins, values that are significantly better than those previously reported for ER and PR fusion proteins (26, 27, 39).

The placement of the activation domain had a significant influence on the activity of the chimeric regulators. However, optimal placement was dependent on the nature of the activation domain. Whereas the VP16 domain yielded the more potent activators when placed at the C terminus, the VP64 domain was more active at the N terminus (Figs. 5 and 6; compare A and B with C and D). Accordingly, direct comparisons demonstrated that an N-terminal VP64 was more potent than a N-terminal VP16 domain, and a C-terminal VP16 was

**FIG. 3.** Generation of designed zinc finger proteins with novel DNA binding specificity. A, amino acid sequence of the three-finger proteins B3 and N1. DNA recognition helix positions –2 to 6, shown in boldface type, were grafted into the framework of the three-finger protein Sp1C. The location of the antiparallel β sheets and the α helices, structural hallmarks of zinc finger domains, are as indicated. DNA binding specificity of each finger is shown on the left. F1 to F3, fingers 1–3. B, ELISA analysis of DNA binding specificity. Zinc finger proteins were expressed in E. coli as MBP fusions and purified. Specificity of binding was analyzed by measuring binding to immobilized biotinylated hairpin oligonucleotides containing the indicated 5′- (GNN)₃-3′ sequences. Black bars, B3; gray bars, N1. Binding of the N1 protein was much weaker, and development time was about 10 times longer than for the B3 protein. The maximal signals were normalized to 1. The KD value for binding to the specific target sequence was measured by electrophoretic mobility shift assay and is labeled above the corresponding bars.

**FIG. 4.** RU486- and 4-OHT-inducible systems for the control of gene expression. A, domain structure of zinc finger protein-steroid hormone receptor fusion proteins. AD, activation domain (VP16 or VP64); ZF, zinc finger DNA binding domains (B3 or N1). B, structure of the 10xB3-TATA-luc- and 10xN1-TATA-luc-inducible promoter constructs. Each contains 10 binding sites, organized into five response elements consisting of direct repeats spaced by three nucleotides. Response elements are spaced by six nucleotides.

**FIG. 5.** Analysis of PR fusion proteins. HeLa cells were cotransfected with expression vectors encoding the indicated PR fusion proteins and 10xB3-TATA-luc (A and C) or 10xN1-TATA-luc (B and D). 24 h after transfection, cells were either left untreated (–), or 10 nM RU486 was added (+). Luciferase activity in total cell extracts was measured 48 h after transfection. Each bar represents the mean ± S.D. of duplicate measurements; vertical axes show the relative light units. RU486-induced stimulation of luciferase expression is indicated. C, control plasmid pcDNA3 that does not express a fusion protein.
The nature and placement of the activation domain was also found to have an influence on the basal activity of the chimeric regulators. In particular, a relatively high basal activity was observed in the case of regulators with N-terminal VP64 domains (Fig. 5C and Expt. 1, Table I).

The nature of the DNA binding domain had a major influence on the extent of ligand dependence of the chimeras. Use of the N1 protein as the DNA binding domain led to more tightly regulated fusion constructs with significantly better fold stimulation of promoter activities (Figs. 5 and 6; compare A and C with B and D) than the use of B3. This difference is most likely due to the dramatic affinity differences between N1 and B3. In particular, the N1-ER-VP64 regulator had no significant basal activity and was capable of mediating a 464-1319-fold activation of the 10xN1-TATA minimal promoter upon 4-OHT induction (Fig. 6D; Table I).

**TABLE I**  Summary of RU486- and 4-OHT-induced gene expression

| PR LBD | ER LBD |
|--------|--------|
| Expt. 1 | Expt. 2 | Expt. 1 | Expt. 2 |
| VP16-B3-LBD | 34 | 36 | 37 | 26 |
| VP64-B3-LBD | 37 | 24 | 26 | 27 |
| B3-LBD-VP16 | 115 | 116 | 47 | 58 |
| B3-LBD-VP64 | 110 | 85 | 62 | 99 |
| VP16-N1-LBD | 188 | 159 | 101 | 39 |
| VP64-N1-LBD | 206 | 390 | 49 | 58 |
| N1-LBD-VP16 | 282 | 203 | 24 | 30 |
| N1-LBD-VP64 | 151 | 129 | 1319 | 464 |

FIG. 6. Analysis of ER fusion proteins. HeLa cells were cotransfected with expression vectors encoding the indicated ER fusion proteins and with 10xB3-TATA-luc (A and C) or 10xN1-TATA-luc (B and D). 24 h after transfection, cells were left untreated (–), or either 10 nM RU486 or 100 nM 4-OHT was added (†). Luciferase activity in total cell extracts was measured 48 h after transfection. Each bar represents the mean value ± S.D. of duplicate measurements; vertical axes show the relative light units. 4-OHT-induced stimulation of luciferase expression is indicated. C, control plasmid pcDNA3 that does not express a fusion protein.

FIG. 7. Independent regulation of two reporter genes. HeLa cells were cotransfected with expression vectors encoding B3-PR-VP16 and N1-ER-VP64 fusion proteins and with the 10xB3-TATA-luc and 10xN1-TATA-β-gal reporter plasmids. 24 h after transfection, cells were left untreated, or either 10 nM RU486 or 100 nM 4-OHT was added. 48 h after transfection, luciferase activity (black bars) and β-galactosidase activity (gray bars) was measured in total cell extracts. Each bar represents the mean value ± S.D. of duplicate measurements. RLU, relative light units.

more potent than a C-terminal VP64 domain (data not shown). The nature and placement of the activation domain was also found to have an influence on the basal activity of the chimeric regulators. In particular, a relatively high basal activity was observed in the case of regulators with N-terminal VP64 domain (Fig. 5C and Expt. 1, Table I).

The nature of the DNA binding domain had a major influence on the extent of ligand dependence of the chimeras. Use of the N1 protein as the DNA binding domain led to more tightly regulated fusion constructs with significantly better fold stimulation of promoter activities (Figs. 5 and 6; compare A and C with B and D) than the use of B3. This difference is most likely due to the dramatic affinity differences between N1 and B3. In particular, the N1-ER-VP64 regulator had no significant basal activity and was capable of mediating a 464-1319-fold activation of the 10xN1-TATA minimal promoter upon 4-OHT induction (Fig. 6D; Table I).

**Coordinated Regulation of Multiple Promoters—**Zinc finger technology has made a large repertoire of DNA binding specificities available for use in protein engineering (6, 8, 9). The availability of different steroid hormone receptor-derived regulatory domains (26, 30) and the ability to redirect chimeric regulators to virtually any desired target sequence should make it possible to independently regulate the expression of multiple genes at the same time. To examine this possibility, a reporter plasmid directing the expression of β-galactosidase under the control of the 10xN1-TATA minimal promoter was constructed. The chimeric regulators B3-PR-VP16 and N1-ER-VP64 were then transiently expressed in HeLa cells along with the 10xB3-TATA-luc and 10xN1-TATA-β-galactosidase reporter plasmids. The transfected cells were treated with either RU486 or 4-OHT, and the luciferase and β-galactosidase activities were monitored (Fig. 7). Significantly, RU486 induced the expression of luciferase while having no effect on β-galactosidase reporter gene activity. 4-OHT, on the other hand, did not affect luciferase expression but efficiently activated β-galactosidase expression. These results demonstrate that the two regulator/promoter combinations act independently of one another and that multiple genes can be efficiently and independently regulated by the selective addition of the desired hormone.

**Potential of Zinc Finger-LBD Fusions for the Regulation of Natural Promoters—**Ready access to polyactyl zinc finger proteins that bind 18 bp of DNA sequence has provided for the generation of artificial transcription factors capable of imposing dominant regulatory effects on endogenous genes (9). For many applications of this approach it may be desirable that the effect on endogenous gene expression is reversible. We therefore sought to extend our previous studies and generate LBD fusion proteins suitable for ligand-dependent regulation of natural promoters. One major drawback of steroid hormone receptors in this approach is that they bind DNA as dimers. Thus, when the fusion protein C7-ER-VP64 was transiently expressed in HeLa cells, it was unable to regulate a reporter construct carrying a single C7 binding site, while it readily regulated a reporter that had two C7 binding sites and therefore accommodated the binding of a dimer to a defined 18-bp site consisting of a repeat of two identical 9-bp sites (Fig. 9A). The most direct approach toward targeting natural promoters would therefore be through binding to a composite target sequence that is in accord with the symmetry requirements imposed by dimerization. However, we reasoned that this approach would be limited, since promoter-specific homodimer binding sites consisting of two identical 9-bp sequences with
the appropriate spacing and orientation would be too rare to be of general use. While targeting promoters using heterodimeric proteins binding a composite site consisting of two distinct 9-bp sites would be feasible, it is also not desirable, because it would require the delivery of a heterodimer and hence two genes. In addition, most steroid receptor LBDs function as homodimers and are not suitable to direct the formation of specific heterodimers. Thus, we reasoned that the type of chimeras that have proven very powerful in regulating artificial promoters would not be optimal for the inducible regulation of endogenous genes. Ideally, inducible regulation of endogenous genes would be achieved by targeting a long DNA sequence lacking specific symmetry requirements, i.e. a long asymmetric sequence. Such sequences might be targeted if the requirement for dimerization of the steroid hormone receptor-based transcription factors could be removed and monomeric hormone-dependent gene switches could be prepared.

Development of a Monomeric Hormone-dependent Gene Switch—We have previously described zinc finger-based artificial transcription factors capable of imposing regulation on the expression of endogenous target genes in a very specific manner (9). These transcription factors contain six zinc finger domains and recognize the 18-bp sequence 5'-GGG GCC GGA GCC GCA GTG-3' in the 5'-untranslated region of the proto-oncogene c-erbB-2 (8, 40). However, when this chimera was tested on the e2c-TATA luciferase reporter plasmid that carries a single E2C binding site upstream of a TATA box, it was constitutively active. Further, the extent of activity was similar to the E2C-VP64 fusion without an ER LBD, and hormone-dependent regulation was also minimal (Fig. 9B). Apparently, the use of a large DNA binding domain recognizing an extended stretch of DNA sequence with high affinity renders the ER LBD fusion protein hormone- and dimerization-independent. To overcome this problem, we produced two types of ER-based chimeric regulators, designed to be capable of regulating gene expression through a single binding site in a hormone-dependent manner. In the first strategy, a heterodimeric regulator was generated consisting of the engineered zinc finger protein E2C fused to an ER LBD as well as an ER LBD fused to a VP64 activation domain (Fig. 8A). When this heterodimeric regulator was expressed in HeLa cells, it had no significant activity on the e2c-TATA luciferase reporter plasmid in the absence of 4-OHT. The addition of hormone led to a 3–5-fold stimulation of luciferase expression, indicating the formation of functional heterodimers (Fig. 9, B and C). However, hormone-induced reporter gene activation was significantly lower than that induced by an E2C-VP64 fusion protein, a result due presumably in part to the formation of E2C-ER and ER-VP64 homodimers. Homodimers were inactive, since neither E2C-ER nor ER-VP64 alone induced luciferase expression (data not shown). In the second strategy, fusion proteins were generated by linking the dimerization partners E2C-ER and ER-VP64 with a flexible polypeptide linker, thereby creating single-chain ligand binding domains. Two linker peptides of 18 or 30 amino acids in length were used to create the proteins E2C-scER/18-VP64 and E2C-scER/30-VP64 (Fig. 8A). Combination of two ER LBDs into one single-chain fusion construct might allow for efficient hormone-induced intramolecular dimerization and therefore yield more efficient activators. Indeed, when E2C-scER/18-VP64 and E2C-scER/30-VP64 were transiently expressed in HeLa cells, they efficiently activated the e2c-TATA-luciferase reporter in a largely hormone-dependent manner (Fig. 9C). 4-OHT-inducible transcriptional regulation was also observed with a luciferase reporter construct under the control of a natural erbB-2 promoter fragment, encompassing nucleotides –758 to –1 with respect to the ATG initiation codon (Fig. 9D). Significantly, the constructs with the longer linkers were reproducibly somewhat more ligand-dependent than the short linker variants (10.3- and 4.9-fold 4-OHT-induction, respectively). In summary, these results show that dimeric regulators requiring response elements similar to those of natural steroid hormone receptors were successfully converted into monomeric ligand-dependent transcription factors capable of modulating gene expression by binding to a long asymmetric DNA sequence.

Monomeric Gene Switch Based on EcR and RXR LBDs—To further explore the concept of ligand-dependent monomeric transcriptional regulators by fusion of two LBDs, the utility of

---

**Chemically Regulated Zinc Finger Transcription Factors**

**Fig. 8.** Domain structure of ligand-dependent, monomeric transcriptional regulators. A, structure of various ER fusion proteins. B, structure of single-chain RXR/EcR fusion proteins. E2C, six finger protein mediating DNA binding; VP64, transcriptional activation domain.
other nuclear hormone receptors was tested. Initially, fusion proteins based on the PR LBD, similar to the single-chain ER constructs, were prepared. However, both 18- and 30-aa linker constructs responded with low induction upon the addition of RU486 (approximately 2-fold stimulation; data not shown). The reason for this discrepancy is presently unclear. Thus, the utility of the LBDs of the Drosophila EcR was investigated. In Drosophila, this receptor functions as a heterodimer between EcR and the product of the ultraspireacle gene (41). However, it has been shown that EcR also efficiently heterodimerizes with ultraspireacle’s vertebrate homologue RXR in response to the ecdysone agonists muristerone A or ponasterone A (PonA) (24, 41, 42). The EcR and RXR LBDs were therefore used to prepare a monomeric gene switch analogous to the scER luciferase reporter (B and C) or the erbB-2 promoter (-758 to -1) luciferase reporter (D). 4-OHT induction and measurement of luciferase activity were carried out as described in A. No induction was observed when cells were treated with RU486 or ponasterone A (not shown).

**DISCUSSION**

Gene function is often probed by methods that either increase the level of the gene product or decrease it. Such perturbations can be imposed at the genomic, transcriptional, or posttranscriptional levels. With the development of zinc finger technology, in particular the development of polydactyl zinc finger proteins with the potential for genome-specific targeting, new approaches toward the directed modulation of gene function have become available. The versatility of this approach could be significantly enhanced if the function of zinc finger-based transcription factors could be controlled in a direct fashion with a small organic molecule. Of the strategies used in nature to control the action of transcription factors with small molecules, the strategy displayed by the steroid hormone receptor family of transcription factors appears to be the most versatile of strategies by which to accomplish this goal. Protein engineering within this family of receptors is further facilitated by their modular structure wherein DNA binding and ligand binding domains are readily defined. Inducible expression systems based on the ER and PR LBDs fused to the Gal4 DBD have been described previously (26, 27, 39). However, while the Gal4 DBD is commonly used in transcriptional studies for the characterization of activation and repression domains (43, 44), it may not be an optimal choice for use in an inducible expression system. In a gene therapy setting, both immunogenicity and specificity or unique targeting may be of concern. In addition, the Gal4 DBD is known to contain both a dimerization function (45) and a nuclear translocation signal (46), both of which have the potential to attenuate hormone dependence of an artificial regulator. Accord-
positions 2

express a fusion protein. No induction was observed when cells were treated with RU486 or 4-OHT (not shown).

three-finger protein; arrangement and bind to a single, contiguous 18-bp DNA sequence.

six-finger “single-chain” fusion proteins undergo an intramolecular re-associate target sequences consisting of two 9-bp subsites.

three-finger fusion proteins undergo intermolecular dimerization and bind composite target sequences.

Left

untreated (−), or 5 μM progesterone A was added (+). Luciferase activity in total cell extracts was measured 48 h after transfection. Each bar represents the mean value ± S.D. of duplicate measurements; vertical axes show the relative light units. pcDNA3.1, control plasmid that does not express a fusion protein. No induction was observed when cells were treated with RU486 or 4-OHT (not shown).

FIG. 11. Schematic representation of the two modes of action of zinc finger-steroid hormone receptor fusion proteins. Left, three-finger fusion proteins undergo intermolecular dimerization and bind composite target sequences consisting of two 9-bp subsites. Right, six-finger “single-chain” fusion proteins undergo an intramolecular rearrangement and bind to a single, contiguous 18-bp DNA sequence. 3F, three finger protein; 6F, six finger protein; ED, effector domain.

ingly, the induction ratios of expression systems based on Gal4-ER and -PR fusion proteins are relatively low (well below 100-fold) (26, 27, 39). In contrast, with designed zinc finger proteins lacking dimerization and nuclear localization functions, we observed induction ratios of up to 3 orders of magnitude. In addition, the availability of various zinc finger framework sequences of human origin coupled with the availability of a wide variety of DNA binding specificities (6, 8, 9) should allow for the construction of regulators with minimal immunogenicity and experimenter-defined targeting specificity.

Little is known about nonspecific effects of designed transcriptional regulators used in the various inducible expression systems described to date (19–29). However, in order to be able to correlate an observed phenotype with the function of the induced gene, the absence of concomitant nonspecific regulation of endogenous genes is essential. Recent progress in DNA microarray technologies has made it possible to assess the specificity of gene regulation in great detail (47, 48), and it will be important to compare available systems using this new approach. It is to be expected that engineered zinc finger proteins will prove to be the DBDs of choice, since they can be easily optimized for minimal pleiotropic effect by screening DBDs of different specificities. In addition, while most zinc finger proteins described to date recognize G + C-rich sequences, recent progress from our laboratory will make proteins binding A + T-rich sequences as readily available (data not shown). An inducible expression system acting through an A + T-rich minimal promoter might present advantages because natural promoters are generally rather G + C-rich. Targeting of A + T-rich sequences should further reduce the probability of cross-reactivity of the designed regulator.

It is commonly believed that heterologous proteins fused to a nuclear hormone receptor LBD are rendered inactive in the absence of a ligand. However, while many of our transcriptional regulators did indeed have very low basal activities, we found significant variations between the degree of ligand independency of the different fusion constructs (Figs. 5 and 6). Evidence of ligand independence can also be found in the previous reports of ER- and PR-based transcription factors (26, 27, 39). Steroid hormone receptors are rendered inactive at least in part due to hsp90 binding to the LBD (14). Since both the ER and the PR fusion constructs described herein contain receptor fragments long enough to encompass the entire hsp90 binding domain, this raises questions concerning the mechanism of ligand-independent transcriptional activation. It seems possible that hsp90 interaction with the chimeric proteins is weaker than with the full-length steroid hormone receptors and that the decreased stability of the hsp90-sequestered state leads to a pool of free monomeric chimeric regulators that are transcriptionally active. This scenario may be true particularly in the case of the ER fusion proteins, since it has been shown that a region immediately preceding the LBD may be required for high affinity hsp90 binding, although there may be no direct interaction (14). Regardless of the mechanism, our results show that the extent of the basal activity is dependent on the zinc finger protein used as DBD, the nature of the LBD, the type of activation domain, and the placement of the activation domain within the protein. The most important of these factors appears to be the DBD, since the N1 fusion proteins generally had a lower basal activity, or none at all, and therefore displayed
higher induction ratios than the B3 fusion proteins. An even more extreme observation was made with the E2C-ER-VP64 construct, which displayed a very high constitutive activity (Fig. 8C). The E2C protein binds DNA with a 0.5 nM dissociation constant. Thus, there is a direct correlation between affinity and basal activity. The N1 protein binds its target sequence with a very low affinity, with a $K_d$ value in the micromolar range (5–10 μM; see Fig. 3B), which may be too weak for efficient transcriptional activation as an unliganded monomer. Dimerization, on the other hand, should cause a significant increase in the effective affinity of the bound transcription factor that may explain the efficient ligand-induced transcriptional activation observed for this transcription factor. Thus, for the construction of the ligand-dependent regulators described here that function as dimeric proteins, zinc finger proteins with low affinity seem to be preferable over those with high affinities. High affinity binding of a monomeric DBD may simply act to shift the equilibrium between the hsp90-sequenced state and the DNA-bound state in favor of the DNA-bound state, thereby enhancing the basal activity and reducing the magnitude of ligand responsiveness.

We have shown that our modular system for controlling target gene expression is able to independently control the expression of two genes within the same transfected cell population, as evidenced by RU486-dependent luciferase induction and 4-OHT-induced β-galactosidase expression (Fig. 7). The lack of β-galactosidase induction by RU486 and luciferase induction by 4-OHT convincingly demonstrates the specificity of the chimeric regulators described here. Not only is the exquisite specificity of the DNA binding domains retained, but also the chimeric regulators described here. Here only is the exquisite specificity of the DNA binding domains retained, but also there is no detectable cross-reaction between RU486 and the ER LBD or between 4-OHT and the PR LBD. The ability to reversibly control the expression of multiple genes or alleles of a gene could prove very useful for many basic research applications. In particular, selective and independent expression of one gene but not another (and vice versa) by small nontoxic ligands would allow for a comparative analysis of gene function, both in vitro and in vivo. In addition, the system described here also allows for both genes to be expressed simultaneously, which could be useful to study the functional interactions between proteins. This is an advantage over a recently described system based on engineered tetracycline-controlled transactivators, which allows switching between the expression of two genes but not their simultaneous expression (32). It should be possible to combine the systems described here with the Tet system to control as many as five genes specifically.

Engineered zinc finger proteins recognizing unique 18-bp addresses have proven to be potent gene-specific targeting devices for the production of artificial transcription factors (9). Typically, long asymmetric DNA sequences have been targeted using polydactyl zinc finger proteins consisting of six zinc finger domains. While the ER- and PR-based transcription factors described above function through recognition of an 18-bp target site, homodimerization places specific constraints on the nature of the target site (Fig. 11). To overcome these targeting constraints, we designed monomeric steroid-hormone receptor variants. We rendered transcriptional regulators prepared from six zinc finger domains ligand-dependent by incorporation of two serially connected LBDs (Figs. 8 and 11). Single-chain regulators containing either two ER LBDs or one EcR and one RXR LBD were functional as monomers and efficiently regulated a natural or synthetic promoter in a ligand-inducible manner by binding to a single 18-bp binding site (Fig. 9 and 10). While we only show ligand-dependent reporter gene activation, this strategy should also be suitable for the regulation of endogenous genes and for transcriptional repression by incorporation of a repression domain. The use of these types of fusion proteins for the inducible regulation of endogenous genes would have significant advantages over other strategies, i.e. expression of the regulator from an inducible promoter. We have recently shown that the use of a tetracycline/doxycycline-dependent expression vector (19) to regulate the expression of an artificial transcription factor is a viable strategy to impose inducible regulation on the endogenous erbB-2 gene (9). However, this strategy requires the delivery of two genes: one encoding the zinc finger protein under the control of a regulatable promoter and the other encoding the regulatory protein. Delivery of multiple vectors is a laborious process and a major hurdle, especially in a gene therapy setting. The use of a monomeric, ligand-dependent transcriptional regulator would be a much more simple and elegant approach and would require only one gene to be delivered. Single-chain steroid hormone receptor LBDs also have the potential for fewer pleiotropic effects, since they are unlikely to engage in heterodimerization with endogenous steroid hormone receptor LBDs due to their intramolecular mode of action. It has not escaped our attention that single-chain steroid hormone LBDs may provide considerable advantages toward the engineering of domains responsive to different ligands. Since homodimerization is no longer required for these receptors to be functional, each of the LBDs present in a single-chain steroid hormone receptor can be independently mutated to create ligand binding sites not accessible to the native LBDs that function by intermolecular dimerization of identical LBDs.

In summary, we describe novel fusion proteins constructed from designed zinc finger DNA binding domains and steroid hormone receptor LBDs. We show that the utility of such chimeras is 2-fold. First, fusion of zinc finger proteins to a transcriptional activation domain and a modified LBD derived from either ER or PR yields potent, ligand-dependent, and highly sequence-specific transcriptional regulators. Together with optimized minimal promoters, these regulators provide 4-OHT- or RU486-inducible expression systems with induction ratios of 2–3 orders of magnitude. Second, we also explore the potential use of zinc finger-steroid receptor fusion proteins for the regulation of natural promoters where single experimenter-defined response elements might be targeted. A gene-specific six-finger protein binding an 18-bp target sequence was converted into a ligand-dependent regulator by fusion with either two ER LBDs or one EcR and one RXR LBD. These engineered single-chain steroid hormone receptor LBDs undergo an intramolecular rearrangement, are functional as monomers, and have potential for the inducible regulation of natural promoters associated with endogenous genes. Together, these chemically regulated transcription factors provide powerful new tools that may be applied to the study of gene function and the alteration of phenotypes of cells or organisms.

Acknowledgments—We are very grateful to Sophia Y. Tsai for providing the GL914VPc' and p17x4TATA-luc constructs and to Trevor D. Littlewood for the pBSK+ ER plasmid.

REFERENCES
1. Desjarlais, J. R., and Berg, J. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7345–7349
2. Choo, Y., and Klug, A. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 11163–11167
3. Rebar, E. J., and Pabo, C. O. (1994) Science 263, 671–673
4. Wu, H., Yang, W. P., and Barbacs, C. F., III (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 344–348
5. Greisman, H. A., and Pabo, C. O. (1997) Science 275, 657–661
6. Segal, D. J., Dreier, B., Beerre, R. R., and Barbacs, C. F., III (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 2758–2763
7. Segal, D. J., and Barbacs, C. F., III (2000) Curr. Opin. Chem. Biol. 4, 34–39
8. Beerre, R. R., Segal, D. J., Dreier, B., and Barbacs, C. F., III (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 14628–14633
9. Beerre, R. R., Dreier, B., and Barbacs, C. F., III (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 1495–1500
10. Jankevich, R., and Hunter, T. (1997) EMBO J. 16, 1620–1627
Chemically Regulated Zinc Finger Transcription Factors

11. Darnell, J. E., Jr. (1997) Science 277, 1630–1635
12. Carson-Jurica, M. A., Schrader, W. T., and O’Malley, W. (1990) Endocr. Rev. 11, 201–220
13. Evans, R. M. (1988) Science 240, 889–895
14. Pratt, W. B., and Toft, D. O. (1997) Endocr. Rev. 18, 306–360
15. Beato, M. (1989) Cell 56, 335–344
16. Kiers, M., Picard, D., Yamamoto, K. R., and Bishop, J. M. (1989) Nature 340, 66–68
17. Superti-Furga, G., Bergers, G., Picard, D., and Busslinger, M. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 5134–5138
18. Samuels, M. L., Weber, M. J., Bishop, J. M., and McMahon, M. (1993) Mol. Cell. Biol. 13, 3343–3356
19. Labow, M. A., Baim, S. B., Shenk, T., and Levine, A. J. (1990) Mol. Cell. Biol. 10, 2256–2260
20. Louvion, J. F., Havaux-Copf, B., and Picard, D. (1993) Gene (Amst.) 131, 129–134
21. Rivera, V. M., Clackson, T., Natesan, S., Pollock, R., Amara, J. F., Keenan, T., Magari, S. R., Phillips, T., Courage, N. L., Cerasoli, F., Jr., Holt, D. A., and Gilman, M. (1996) Nat. Med. 2, 1028–1032
22. Littlewood, T. D., Hancock, D. C., Danielian, P. S., Parker, M. G., and Evan, G. I. (1995) Nucleic Acids Res. 23, 1686–1690
23. Vege, E., Allan, G. F., Schrader, W. T., Tsai, M.-J., McDonnell, D. P., and O’Malley, B. W. (1992) Cell 69, 703–713
24. Baron, U., Schnappinger, D., Helbl, V., Gosse, M., Hillen, W., and Bujard, H. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 1013–1018
25. Desjarlais, J. R., and Berg, J. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2256–2260
26. Liu, Q., Segal, D. J., Ghiara, J. B., and Barbas, C. F., III (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 5525–5530
27. Kim, J. S., and Pabo, C. O. (1997) J. Biol. Chem. 272, 29795–29800
28. Aumais, J. P., Lee, H. S., DeGannes, C., Horsford, J., and White, J. H. (1996) J. Biol. Chem. 271, 12568–12577
29. Pavletich, N. P., and Pabo, C. O. (1990) Science 252, 809–817
30. Sadowski, I., Ma, J., Triezenberg, S., and Ptashe, M. (1988) Nature 335, 563–564
31. Whelan, J., and Miller, N. (1996) J. Steroid Biochem. Mol. Biol. 58, 3–12
32. Nakanishi, K. (1992) Steroids 57, 649–657
33. Seipel, K., Georgiev, O., and Schaffner, W. (1992) EMBO J. 11, 4961–4968
34. Margolin, J. F., Friedman, J. R., Meyer, W. K.-H., Vissing, H., Thiesen, H.-J., and Rauscher, F. J., III (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4509–4513
35. Carey, M., Kakidani, H., Leatherwood, J., Mostashari, F., and Ptashe, M. (1989) J. Mol. Biol. 209, 423–432
36. Shalon, D., Smith, S. J., and Brown, P. O. (1996) Genome Res. 6, 639–645
37. Lockhart, D. J., Dung, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kohsayashi, M., Horton, H., and Brown, E. L. (1996) Nature Biotechnol. 14, 1675–1680