MOLECULAR AND CELLULAR BIOLOGY, Sept. 1993, p. 5794–5804
0270-7306/93/095794-11$02.00/0
Copyright © 1993, American Society for Microbiology
Vol. 13, No. 9

# The Orphan Receptors NGFI-B and Steroidogenic Factor 1 Establish Monomer Binding as a Third Paradigm of Nuclear Receptor-DNA Interaction

THOMAS E. WILSON, TIMOTHY J. FAHRNER, AND JEFFREY MILBRANDT*

*Departments of Pathology and Internal Medicine, Division of Laboratory Medicine, Washington University School of Medicine, St. Louis, Missouri 63110*

Received 13 April 1993/Returned for modification 11 June 1993/Accepted 25 June 1993

**We examined in detail the DNA interaction of the nuclear receptors NGFI-B and steroidogenic factor 1 (SF-1) by using a series of gain-of-function domain swaps. NGFI-B bound with high affinity as a monomer to a nearly linear DNA molecule. The prototypic zinc modules interacted with a half-site of the estrogen receptor class, and a distinct protein motif carboxy terminal to the zinc modules (the A box) interacted with two A/T base pairs 5′ to the half-site. SF-1 bound in the same manner as NGFI-B, with an overlapping but distinct sequence requirement 5′ to the half-site. The key features that distinguished the NGFI-B and SF-1 interactions were an amino group in the minor groove of the SF-1 binding sequence and an asparagine in the SF-1 A box. These results define a common mechanism of NGFI-B and SF-1 DNA binding, which may underlie a competitive mechanism of gene regulation in steroidogenic tissues that express these proteins. This monomer-DNA interaction represents a third paradigm of DNA binding by nuclear receptors in addition to direct and inverted dimerization.**

It is well established that the receptors for steroid hormones are intracellular proteins that are themselves transcription factors. Binding of steroid to the conserved carboxy-terminal domains results in a conformational change that allows shedding of heat shock proteins, phosphorylation, translocation of the protein to the nucleus, and binding to specific DNA sites via the conserved central domain of the proteins (1, 13, 18, 24, 31, 42, 45). Proteins with these characteristic, conserved functional domains also mediate the effects of vitamin D, all-*trans*- and 9-*cis*-retinoic acid, thyroid hormone, the insect molting hormone ecdysone, and probably various fatty acid compounds (5, 9, 12, 17, 22, 33, 37, 63). The term nuclear receptor is now generally applied to this superfamily of proteins to broaden the consideration of their functions while recognizing their general receptor-like characteristics. This term may still be inadequate, since molecular cloning techniques have revealed many proteins that share these structural features but whose function is unknown (the orphan receptors), which may serve as transcription factors regulated by other means (27, 41). At least one protein of this class, NGFI-B (also called nur77), is closely regulated at the transcriptional level by the actions of growth factors and membrane depolarization, but once expressed NGFI-B can activate transcription without exogenously added ligand (6, 16, 38, 43, 61, 64, 65).

The first demonstrations of DNA binding by nuclear receptors were made by examining the promoters of known glucocorticoid- and estrogen-regulated genes (19, 20, 44). The identified response elements function as inverted repeats of protein-specific 6-bp elements (termed half-sites) separated by three nonconserved base pairs, each half-site interacting with one molecule of a steroid receptor dimer (23, 55). Many nonsteroid receptors have since been shown to bind inverted half-site repeats with different spacings (11) or direct rather than inverted repeats (26, 35, 40, 53, 60). One half-site of these direct repeat motifs is often bound by the retinoid X receptor (RXR) in a heterodimeric complex (21, 32, 67, 68). A fascinating feature of these sites is that there are only two currently recognized half-site variants, the glucocorticoid receptor (GR; 5′-AGAACA) and estrogen receptor (ER; 5′-AGGTCA) types, with most proteins recognizing the ER type. This led to the hypothesis that differential DNA binding by nuclear receptors is based primarily on half-site orientation and spacing and thus is based on the dimeric nature of these proteins (40, 60).

Functional and structural analyses have made these characteristics of DNA binding increasingly clear at a molecular level. The central conserved domain of nuclear receptors chelates two zinc atoms via eight invariant cysteines in a manner not unlike that of the TFIII-A protein, which led to the use of the term zinc fingers to describe this domain (3, 8). However, crystal and nuclear magnetic resonance structures have revealed that, unlike the TFIII-A class of zinc fingers, the GR and ER DNA binding domains fold into a single tertiary structure dependent on both zinc-chelating modules (14, 34, 49). The first zinc module contains an α-helix encompassing a motif (the P box [59]) that determines half-site specificity by making contacts in the major groove at the distinguishing central base pairs (34). Within the second zinc module there is a surface created in part by a loop (the D box [59]) that can interact with its counterpart on the other molecule of a dimer bound to an inverted repeat, creating a remarkably strong and specific dimerization interface (34). Finally, we recently identified a dimerization motif immediately downstream of the second zinc module (the T box) that might accommodate the head-to-tail orientation of two monomers bound to a direct repeat by interacting with the opposite face of the second molecule of the dimer pair (66).

To explore the function of the orphan receptor NGFI-B, we previously used a yeast genetic selection system and other analyses to define the nucleotide sequence bound by

* Corresponding author.

this protein as the octamer 5'-AAAGGTCA (the NGFI-B response element, or NBRE [64]). We have since demonstrated that NGFI-B participates in the ACTH-mediated transcriptional activation of the gene encoding steroid 21-hydroxylase from an NBRE within this promoter, strengthening NGFI-B's identity as a growth factor-regulated nuclear receptor (65). The NBRE represented an unusual site for binding and transcriptional activation by a nuclear receptor since it contained only one ER half-site, and in addition it required nucleotides that did not appear to constitute a second half-site. Experiments with NGFI-B–RXR-β DNA binding domain chimeras demonstrated that NBRE binding requires an unusual amino acid motif (the A box) immediately downstream of the T box (66).

In this report, we demonstrate that the essential contacts of the NGFI-B–NBRE interaction can be entirely accounted for by a protein monomer. Specifically, the NGFI-B zinc modules contact the ER half-site in the NBRE in the same manner as those of steroid receptors, and amino acids of the A box contact nucleotides adjacent to the half-site. These additional contacts occur in the minor groove of the DNA, on the same face of the DNA double helix as the zinc module–half-site contacts. Finally, we demonstrate that monomer DNA binding is not unique to NGFI-B, because at least one other nuclear receptor, steroidogenic factor 1 (SF-1 [25]), binds in essentially the same manner, with an overlapping but distinct sequence specificity at the 5' adjacent nucleotides. Since SF-1 and NGFI-B binding activities are coexpressed in the adrenal gland, the tissue of NGFI-B-mediated 21-hydroxylase activation (25, 65), it is likely that this mode of DNA binding will underlie a competitive mechanism of gene regulation.

## MATERIALS AND METHODS

**Preparation of bacterial extracts.** To obtain the NGFI-B and SF-1 glutathione *S*-transferase fusion proteins (GST-B and GST-S, respectively), polymerase chain reaction (PCR) products (with the NGFI-B cDNA [38] and total cDNA made from mouse adrenal gland mRNA as templates, respectively) corresponding to the residues shown in Fig. 3A were ligated into the *Bam*HI and *Eco*RI sites of the vector pGEX-1 (Pharmacia, Uppsala, Sweden). The plasmid encoding GST-GB was obtained in a similar fashion by using the plasmid BGB, which encodes full-length NGFI-B with GR zinc modules, as a PCR template. The constructs GST-BS1, GST-BS7, and GST-BS8 were made by making two PCR products corresponding to the appropriate chimera halves, which were joined by a three-part ligation into the pGEX-1 vector. To facilitate three-part ligation, a *Not*I site was introduced into GST-BS8 by silent mutation, and an *Sph*I site was introduced into GST-BS7 by a V-to-L change. In these constructs, the stop codon was provided by the vector, which resulted in the addition of a few foreign amino acids at the carboxy terminus. The constructs GST-BGt, -Bel, -Bt1, -Bt2, -Bt3, -BSt1, -BSt2, -BSt3, and -BS2 to -BS6 were made by single PCR on appropriate DNA templates with reverse primers that encompassed the *Eco*RI site, a stop codon, and the necessary chimeric regions. All plasmids were sequenced through the entire downstream domain.

*Escherichia coli* DH5α was transformed with the plasmid of interest and grown to saturation at 37°C in Luria-Bertani medium containing 100 μg of ampicillin per ml. This culture was diluted 10-fold into fresh Luria-Bertani medium plus ampicillin and incubated at 37°C with vigorous aeration for 1 h. IPTG (isopropyl-β-D-thiogalactopyranoside) was added to a final concentration of 0.1 mM, and the cultures were incubated for an additional 2 h. Cells (3 ml) were harvested by centrifugation and resuspended in 1 ml of ice-cold gel shift buffer (see below). Cells were lysed by two 15-s sonications, and the insoluble material was pelleted by centrifugation in a microcentrifuge. Five microliters of this crude lysate (~9 μg of protein, as determined by the Bradford protein assay with bovine plasma globulin as a standard) was used as a protein source in each gel shift binding reaction. To visualize and quantify the induced protein, 50 μl of the supernatant was boiled in protein sample buffer and electrophoresed on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel (10% polyacrylamide) followed by staining with Coomassie brilliant blue. To estimate the solubility of a given fusion protein, the crude supernatant was electrophoresed on an SDS-PAGE gel (10% polyacrylamide) next to a one-third volume of bacterial culture boiled in protein sample buffer. When indicated, the GST-B fusion protein was partially purified from the crude lysate by binding to glutathione-agarose beads (Sigma, St. Louis, Mo.), with 1 ml of packed resin per 100 ml of bacterial culture. The beads were washed with 30 volumes of gel shift buffer, and the fusion protein was eluted with 3 volumes of 10 mM reduced glutathione (Sigma). The protein solution was then equilibrated with gel shift buffer by a NAP-10 column (Pharmacia).

**In vitro translation.** The plasmid used to express NGFI-B in vitro was created by ligating an *Nco*I-*Sal*I fragment encoding NGFI-B with a *Bam*HI linker inserted at the *Age*I site into the pCITE vector (Novagen, Madison, Wis.). Plasmid fl-BS6 was obtained by a three-part ligation of PCR products extending between (i) the *Bam*HI site and a *Not*I site introduced in the A box by silent mutation and (ii) the *Not*I site and the unique *Aat*II site of NGFI-B. Plasmid fl-BS1 was created by a three-part ligation into fl-BS6 of (i) a PCR product extending from the *Bam*HI site to an *Sph*I site at the NGFI-B–SF-1 junction and (ii) a double-stranded oligonucleotide spanning the SF-1 T box from the *Sph*I site to the added *Not*I site. Generation of proteins from these plasmids was accomplished in vitro by coupled transcription and translation with the TnT T7 reticulocyte lysate system (Promega, Madison, Wis.) and Express-labeled methionine (NEN), according to the manufacturers' directions. Five microliters of a translation mixture was used in each gel shift binding reaction. This volume was also electrophoresed on an SDS-PAGE gel (10% polyacrylamide) which was dried and exposed to a Phosphorimager screen overnight. The relative amount of radioactivity incorporated into each protein was measured by volume integration with the ImageQuant software package (Molecular Dynamics) to ensure that similar amounts of protein were added to all lanes of Fig. 5B.

**Gel shift procedure.** Oligonucleotides were synthesized on an Applied Biosystems model 394 synthesizer with the "trityl off" option with A, C, G, T, and inosine phosphoramidites from Applied Biosystems and 5-methyl cytosine phosphoramidite from Pharmacia. After cleavage and deprotection in ammonium hydroxide at 55°C overnight, the oligonucleotides were purified on a NAP-10 column (Pharmacia) with Tris-EDTA as the buffer. Gel shift probes were prepared by mixing 25 μg of each oligonucleotide strand, heating to 90°C, and annealing at room temperature for 30 min. After electrophoresis on a 20% acrylamide gel cast with 1× Tris-borate-EDTA buffer, the band corresponding to double-stranded oligonucleotide was cut out and purified by passive elution and ethanol precipitation. Two hundred fifty

nanograms of annealed probe (with an extinction coefficient of 40 mg/ml at an optical density of 260 nm) was phosphorylated to a specific activity of $>10^8$ cpm/μg with 75 μCi of [γ-$^{32}$P]ATP (Amersham) and 10 U of T4 polynucleotide kinase (New England Biolabs) and purified on a Nick column (Pharmacia).

Gel shift analyses were performed as follows. The appropriate amount of a given protein fraction was added to 50 μl of gel shift buffer (10 mM Tris [pH 7.5], 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 5% glycerol) that contained (i) 0.5 ng of $^{32}$P-labeled double-stranded oligonucleotide probe, (ii) 3 μg of sonicated herring sperm DNA, (iii) the indicated amounts of unlabeled competitor oligonucleotides, and (iv) protease inhibitors (1 mM phenylmethylsulfonyl fluoride, 5 μg of leupeptin and pepstatin each per ml [Fig. 5D only]). Binding reaction mixtures were incubated for 30 min at room temperature (except Fig. 5D reaction mixtures, which were incubated on ice) and electrophoresed at 150 mV on a 5% acrylamide gel cast with 0.5× Tris-borate-EDTA. Gels were dried and exposed to film or a Phosphorimager screen. All measurements of radioactivity present in given band (see Table 2) were made with the volume integration option in the ImageQuant software package, with a pixel background correction taken from a region with no radioactive signal.

**Methylation interference.** Methylation interference was performed essentially as described in reference 52. The two strands of NBRE-MI (where the top strand is 5′-CTTC ACACTGGAGTTGT**AAAAGGTCA**TGCTCAATTTCTAG CATC) were synthesized as described above except that the "trityl on" option was used, which allowed for purification and detritylation with an oligonucleotide purification cartridge (Applied Biosystems). Two hundred fifty nanograms of each strand was then labeled with [γ-$^{32}$P]ATP as described above, and EDTA was added to a final concentration of 2.5 mM to inactivate the polynucleotide kinase. One microgram of the appropriate unlabeled opposite strand was added and annealed by heating to 90°C followed by incubation at room temperature for 30 min. The labeled, double-stranded probe was then purified, first on a Nick column and then by electrophoresis on a 20% acrylamide gel followed by passive elution and ethanol precipitation. A total of $1.4 \times 10^6$ cpm of each probe was methylated in 100 μl of Tris-EDTA by adding 5 μl of 20% dimethyl sulfate (made with ethanol) and incubating for 10 min on ice. The reaction was stopped by the addition of 100 μl of stop solution (10% β-mercaptoethanol, 0.6 M sodium acetate, 100 μg of sonicated herring sperm DNA per ml), and the DNA was harvested by ethanol precipitation. A total of $2.5 \times 10^5$ cpm of each methylated probe was then incubated with 20 μl (3.7 μg) of purified GST-B in 200 μl of gel shift buffer for 30 min, followed by electrophoresis on a 5% acrylamide gel cast with 0.5× Tris-borate-EDTA. The bands corresponding to free and GST-B-bound probe were visualized by autoradiography of the wet gel and were excised. The DNA was electroeluted onto NA45 paper (Schleicher & Schuell), eluted with 1 M NaCl at 55°C, and ethanol precipitated with 500 ng of sonicated herring sperm DNA as the carrier. The DNA was cleaved by resuspension in 90 μl of Tris-EDTA, with piperidine added to a final concentration of 10%, and incubation for 30 min at 90°C. The piperidine was removed by three cycles of lyophilization, with 50 μl of $H_2O$ added between cycles, and the DNA was resuspended in formamide. Ten thousand counts per minute of each cleaved DNA was then electrophoresed on a 17% sequencing gel, which was wrapped in plastic wrap and exposed to film.

**Cell transfection and luciferase assays.** The plasmid used to express NGFI-B in tissue culture cells was JDM723, which had a *Bam*HI linker inserted at the *Age*I site. The plasmids used to express fl-BS1 and fl-BS6 were constructed by ligating (i) the *Bst*EII-*Bam*HI fragment of JDM723 encoding the amino terminus of NGFI-B and (ii) the *Bam*HI-*Hin*dIII fragments of the pCITE fl-BS1 and fl-BS6 vectors, encoding the DNA binding domain and carboxy terminus of NGFI-B, into the *Bst*EII and *Hin*dIII sites of JDM723 in a three-part ligation. The luciferase reporters AAA-luc and TCA-luc were made by ligating unphosphorylated double-stranded oligonucleotides containing two copies of the 9-base AAA or TCA sequences (in tandem, separated by five nucleotides) into the *Bam*HI site of Pro-36–luc, just upstream of the minimal prolactin promoter (64). These were screened by PCR to verify that the orientation of the oligonucleotides was the same. CV-1 cells were transfected and luciferase activity was measured exactly as previously described (64). The data in Fig. 5C represent the average of at least two different experiments performed with two different DNA preparations.

## RESULTS

**The NBRE is limited to 9 bp.** We previously suggested that NGFI-B binds to DNA as a monomer largely on the basis of the observation that the NBRE contains only one ER half-site, which implied that only one protein molecule binds by interactions between (i) the zinc modules and the ER half-site and (ii) the A box and the nucleotides 5′ to the half-site (66). Also, mixing differently sized forms of NGFI-B did not reveal dimerization; however, this assay yields a positive result only when dimerization occurs, and the negative result with NGFI-B was not conclusive. However, it is also possible that the role of the A box is to change the specificity of the zinc modules such that they interact as a dimer with atypical, perhaps overlapping, half-sites that include the 5′ A/T base pairs of the NBRE.

To examine the contacts made between NGFI-B and DNA in more detail, we first performed methylation interference with a bacterially expressed protein that contained the NGFI-B DNA binding domain (including the zinc modules through the A box) fused to GST-B. This protein bound specifically and with high affinity to NBRE-containing oligonucleotides (see below). Significant interference with GST-B binding was observed only when methylation occurred on the nucleotides of the ER half-site or the A/T base pairs 5′ to the half-site, with only minor interference observed when nucleotides one position beyond these were methylated (Fig. 1). Thus, the NBRE is completely limited to 9 bp, an observation consistent with previous analyses of NBRE mutants tested with native NGFI-B (64). Furthermore, because methylation at A nucleotides occurs on the minor groove of DNA (52), these interference results gave a preliminary suggestion that the A box contacts the 5′ adjacent nucleotides in the minor groove.

**The Zn module-half-site interaction.** To establish the link between the NGFI-B zinc modules and the prototypic ER half-site in the NBRE, we sought to switch the specificity at this half-site to that of GR. Binding of GST-B to the NBRE oligonucleotide AAA was abolished by a mutation that altered the ER half-site to a GR half-site (probe AAA-G; Fig. 2 and Table 1). However, when the zinc modules in GST-B were replaced with those of GR (GST-GB), the protein now bound to AAA-G and to the glucocorticoid response element (GRE) itself but not to the NBRE. The GST-GB–AAA-G interaction was uniquely dependent on the NGFI-B A box in

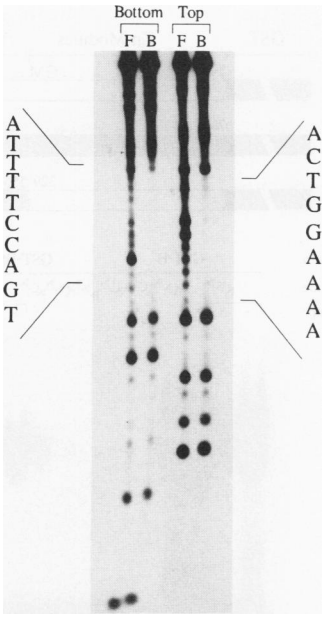


FIG. 1. Methylation interference of the NGFI-B–NBRE interaction. Methylation interference was performed with both strands (top and bottom) of the NBRE-containing oligonucleotide NBRE-MI (see Materials and Methods) by using bacterially expressed, purified GST-B (Fig. 2A). F indicates the cleavage pattern of free oligonucleotide, and B indicates the pattern of protein-bound oligonucleotide. The span of nucleotides where interference was observed is indicated for both strands.

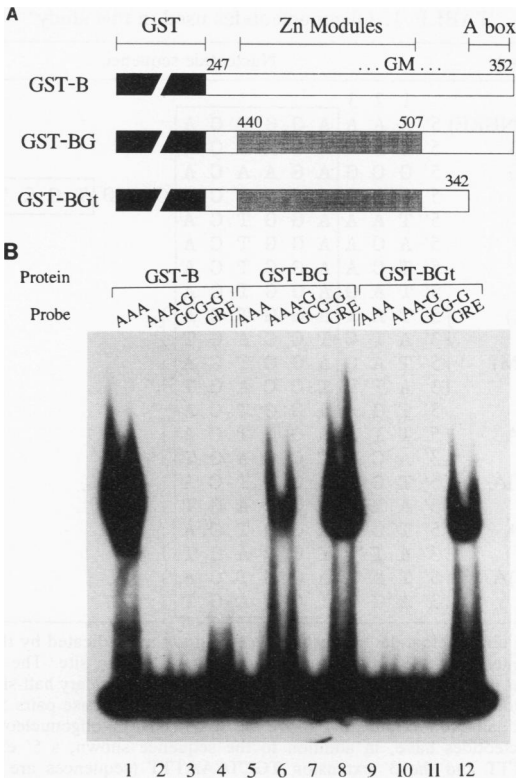


FIG. 2. The NGFI-B zinc modules interact with the ER half-site in the NBRE. (A) A schematic of the GST fusion proteins used in panel B. The GST portion is shown in black (with a break to indicate a discontinuity in scale), the NGFI-B portions are in white, and the GR portions are in gray. Numbers above the bars correspond to the NGFI-B (38) and GR (62) amino acids at the endpoints of the relevant regions. (B) A gel shift assay was performed with crude bacterial lysates containing the proteins diagrammed in panel A and oligonucleotide probes shown in Table 1. The autoradiogram was somewhat overexposed to demonstrate the absence of protein-DNA complexes for the GST-BG–GCG-G and GST-BGt–AAA-G pairs.

GST-GB, since its deletion (protein GST-GBt) eliminated binding to probe AAA-G but not to the GRE. It was also dependent on the A/T base pairs 5′ to the half-site, since no protein bound when these were mutated to G/C base pairs (probe GCG-G). Thus, it is possible to switch the specificity of the NBRE to a GR half-site in a manner that maintains the unique structural requirements for NGFI-B–NBRE binding. In a different sense, GR can be made to bind a single half-site if an appropriate A box-5′ nucleotide combination is provided. We conclude by inference that the NGFI-B zinc modules interact with the ER half-site in the NBRE in a fashion similar to that of steroid receptor zinc modules.

**The A box-DNA interaction.** The above results, in combination with our previous analysis of NGFI-B–RXR chimeras (66), strongly suggest that the A box forms interactions with the A/T base pairs 5′ to the ER half-site in the NBRE. To demonstrate this directly, we sought to switch the specificity at the 5′ nucleotides via compensatory mutations in the A box amino acids. An examination of the literature revealed that the orphan nuclear receptor SF-1 binds to the NBRE, as well as other apparent half-site monomers, found in the promoter of the 21-hydroxylase gene (25, 47). Other reports have shown that SF-1 (termed the embryonal long terminal repeat-binding protein, or ELP, in these studies), as well as its *Drosophila melanogaster* homolog FTZ-F1, binds to an optimal DNA sequence, 5′-TCAAGGTCA, which also contains a single ER half-site (28, 56, 57). We hypothesized that SF-1 binds to these sites as a monomer in the same manner as NGFI-B.

The SF-1 DNA binding domain was expressed in bacteria as a GST fusion (GST-S). The initial endpoints (Fig. 3A and Table 2), which include more residues than are present in GST-B, were chosen because the homology between SF-1 and FTZ-F1 extends for nine amino acids beyond the A box (Table 3). The GST-S–DNA complexes were much weaker than was observed for GST-B, due in part to the fact that bacterially expressed GST-S was not completely soluble. Nonetheless, it is clear that GST-S bound to the NBRE oligonucleotide with a low but measurable affinity but bound significantly better when the nucleotides 5′ to the half-site were changed from 5′-AAA to 5′-TCA (oligonucleotide TCA, Fig. 3 and Tables 1 and 2). In marked contrast, GST-B did not bind at all to TCA (Fig. 3B). This difference in binding was due primarily to the second of these nucleotides, A versus C, since GST-B bound equally well to probes AAA and TAA (with a slight preference for TAA) but bound not at all when only the second position was changed (probe ACA). Thus, while SF-1 is somewhat more tolerant to changes, SF-1 and NGFI-B display markedly different specificities for the nucleotides 5′ to an ER half-site. Replacement of the SF-1 zinc modules in GST-S with those of NGFI-B (protein GST-BS1) did not affect its SF-1 sequence specificity (Fig. 3B). This indicates that the domain downstream of the zinc modules provides the residues necessary for site discrimination, consistent with an interaction between the SF-1 zinc modules and ER half-site in a manner parallel to NGFI-B. We could not test the converse construct (SF-1 zinc modules

TABLE 1. Oligonucleotides used in this study[a]

| Name | Nucleotide sequence | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | | | | | | | | | | | |
| AAA (NBRE) | 5′ | A | A | A | A | G | G | T | C | A | | | | | |
| AAA-G | 5′ | A | A | A | A | G | A | A | C | A | | | | | |
| GCG-G | 5′ | G | C | G | A | G | A | A | C | A | | | | | |
| GRE | 5′ | G | C | G | A | G | A | A | C | A | C | A | G | T | G T T C T |
| TAA | 5′ | T | A | A | A | G | G | T | C | A | | | | | |
| ACA | 5′ | A | C | A | A | G | G | T | C | A | | | | | |
| TCA | 5′ | T | C | A | A | G | G | T | C | A | | | | | |
| TAG | 5′ | T | A | G | A | G | G | T | C | A | | | | | |
| TA(I/C) | 5′ | T | A | I | A | G | G | T | C | A | | | | | |
| | 3′ | A | T | C | T | C | C | A | G | T | | | | | |
| TA(G/M) | 5′ | T | A | G | A | G | G | T | C | A | | | | | |
| | 3′ | A | T | M | T | C | C | A | G | T | | | | | |
| TGA | 5′ | T | G | A | A | G | G | T | C | A | | | | | |
| T(I/C)A | 5′ | T | I | A | A | G | G | T | C | A | | | | | |
| | 3′ | A | C | T | T | C | C | A | G | T | | | | | |
| T(G/M)A | 5′ | T | G | A | A | G | G | T | C | A | | | | | |
| | 3′ | A | M | T | T | C | C | A | G | T | | | | | |
| T(C/I)A | 5′ | T | C | A | A | G | G | T | C | A | | | | | |
| | 3′ | A | I | T | T | C | C | A | G | T | | | | | |
| T(M/G)A | 5′ | T | M | A | A | G | G | T | C | A | | | | | |
| | 3′ | A | G | T | T | C | C | A | G | T | | | | | |

[a] The oligonucleotide probes used in this study are indicated by the name and sequence of the top strand of the relevant binding site. The primary half-sites in the oligonucleotides are boxed, as is the secondary half-site in the GRE probe. The position numbers used to describe the base pairs 5′ to the primary half-site are indicated above the AAA (NBRE) oligonucleotide. All oligonucleotides have, in addition to the sequence shown, a 5′ extension GAGTTTT and the 3′ extension TGCTCAATTT (sequences are the top strand). While all probes were completely double stranded, the bottom strand of an oligonucleotide pair is shown only to indicate the nature of the modified base substitutions, where I represents deoxyinosine and M represents 5-methyl-deoxycytosine.

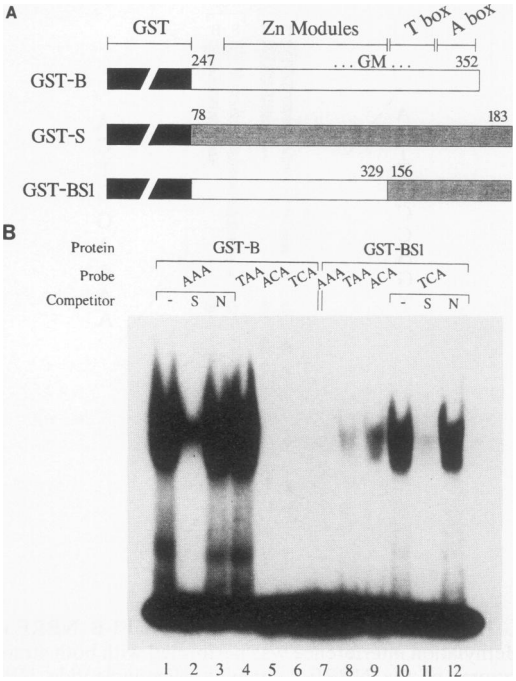


FIG. 3. DNA binding by NGFI-B and SF-1 GST fusion proteins. (A) A schematic of the GST fusion proteins used in panel B and Table 2. The GST portion is shown in black (with a break to indicate a discontinuity in scale), the NGFI-B portions are in white, and the SF-1 portions are in gray. Numbers above the bars correspond to the NGFI-B (38) and SF-1 (56) amino acids at the endpoints of the relevant regions. (B) A gel shift assay was performed with crude bacterial lysates containing the proteins GST-B and GST-BS1 and oligonucleotide probes shown in Table 1. Where indicated, 200 ng of a specific (S) or nonspecific (N) competitor oligonucleotide was added (lane 2, AAA; lane 11, TCA; lanes 3 and 12, GRE). –, no competitor added. Lanes 1 to 4 are somewhat overexposed to bring out the less prominent GST-BS1 complexes. A faint band in lane 7 is not visible but was detectable with a Phosphorimager (Table 2).

with the NGFI-B downstream domain) because it was completely insoluble when expressed in bacteria as a GST fusion.

The proteins shown in Table 2 were used to demonstrate that the minimal downstream domain needed for specific DNA binding is similar for NGFI-B and SF-1. Both GST-B and GST-BS1 could be terminated immediately after the A box with no change in overall DNA binding affinity or specificity (GST-Bt1 and GST-BSt1), indicating that the extended domain of conservation between SF-1 and FTZ-F1 does not affect binding by GST-S or GST-B. Furthermore, both proteins were rendered ineffective at DNA binding when the A box was deleted (GST-Bt3 and GST-BSt3). Interestingly, when a truncation was made within the A box of SF-1 (GST-BSt2), the protein still bound to DNA, albeit with somewhat reduced affinity, in contrast to GST-Bt2, which did not bind. A comparison of GST-BSt1 with GST-BS2, which both contain the NGFI-B zinc modules and the SF-1 T box region, verifies that specificity for differing 5′ adjacent nucleotides lies entirely within the A box, since the change in specificity between these constructs was complete. Further comparison of the constructs GST-BSt1 with GST-BS4 and GST-BS2 with GST-BS5 further limits specificity to the amino-terminal portion of the A box, specifically the first three amino acids (note construct GST-BS3). Thus, mutation of one or two nucleotides in the target DNA can be compensated for by a change in only three A box amino acids, which indicates that contact is made between these amino acids and nucleotides. The fact that no partial change in specificity was observed (for example, equal binding to all probes) further indicates that a small number of contacts confer the differential specificity of NGFI-B and SF-1 and that these are made entirely by the proximal A box.

The panel of NGFI-B–SF-1 chimeras also indicates an important function for the amino acids between the zinc modules and the A box (the T box motif) in monomer DNA binding. In contrast to proteins containing an SF-1 T box (see above), proteins with an NGFI-B T box were only capable of binding in the context of an NGFI-B A box (compare GST-B and GST-BS6, Table 2) even though both proteins were soluble. Viewed in another way, proper function of an SF-1 A box required the presence of an SF-1 T box. To examine this further, we made mutations in the SF-1 T box contained within the parent clone GST-BS1. A comparison of the constructs GST-BS1, GST-BS7, and GST-BS8 implicated Ala-159 as a key factor in the permissive nature of the SF-1 T box. However, changing only this amino acid did not impart measurable SF-1 binding activity to GST-BS6 (not shown). It is important to note that these T box changes are not primarily affecting the SF-1 sequence specificity of GST-BS1, since GST-BS2 contained a complete SF-1 T box but bound as NGFI-B. While the precise role of the T box region is not made clear by these data, it presumably reflects an important secondary structure of the region (see Discussion).

**Identification of minor groove contacts.** We next turned to

TABLE 2. DNA binding by NGFI-B–SF-1 chimeras

| Name | Amino acid sequence[a] | | Extent of binding to probe[b]: | | | |
|---|---|---|---|---|---|---|
| | T box | A box | AAA | TAA | ACA | TCA |
| GST-B | ...GM VKEVVRTDSLKG | RRGRLPS KPeflvtd | 87 | 100 (100) | —[c] | — |
| GST-S | ...GM RLEAVRADRMRG | GRNKFGP MYKTDRALKeflvtd | 15 | 25 | 64 | 100 (30) |
| GST-BS1 | ...GM RLEAVRADRMRG | GRNKFGP MYKTDRALKeflvtd | 6 | 17 | 30 | 100 (100) |
| GST-Be1 | ...GM VKEVVRTDSLKG | RRGRLPS KPKQPPDAS | 91 | 100 (94) | — | — |
| GST-Bt1 | ...GM VKEVVRTDSLKG | RRGRLPS K | 76 | 100 (87) | — | — |
| GST-Bt2 | ...GM VKEVVRTDSLKG | RRGR | — | — | — | — |
| GST-Bt3 | ...GM VKEVVRTDSLKG | | — | — | — | — |
| GST-BSt1 | ...GM RLEAVRADRMRG | GRNKFGP M | 6 | 24 | 41 | 100 (105) |
| GST-BSt2 | ...GM RLEAVRADRMRG | GRNK | 6 | 24 | 45 | 100 (54) |
| GST-BSt3 | ...GM RLEAVRADRMRG | | — | — | — | — |
| GST-BS2 | ...GM RLEAVRADRMRG | RRGRLPS KP | 69 | 100 (70) | — | — |
| GST-BS3 | ...GM RLEAVRADRMRG | RRGKLPS KP | 69 | 100 (69) | — | — |
| GST-BS4 | ...GM RLEAVRADRMRG | RRGKFGP M | 57 | 100 (32) | — | — |
| GST-BS5 | ...GM RLEAVRADRMRG | GRNKLPS K | 7 | 23 | 36 | 100 (69) |
| GST-BS6 | ...GM VKEVVRTDSLKG | GRNKFGP M | — | — | — | — |
| GST-BS7 | ...GM lKEAVRADRMRG | GRNKFGP MYKTDRALKeflvtd | 13 | 31 | 52 | 100 (47) |
| GST-BS8 | ...GM VKEVVRADRMRG | GRNKFGP MYKTDRALKeflvtd | — | — | — | — |

[a] All NGFI-B–SF-1 chimeric proteins were expressed in bacteria as GST fusions similar to Fig. 3A. Uppercase letters (with no underline) indicate NGFI-B residues and uppercase underlined letters indicate SF-1 residues. Lowercase letters indicate residues that correspond to neither NGFI-B nor SF-1. All proteins terminate immediately after the last amino acid shown. The GM amino acid pair that marks the end of the zinc modules is used to indicate their identity in each protein. The small spaces in the sequences (before and after the T and A boxes) were added to facilitate comparison and do not signify a gap in the sequences.

[b] To reveal the different binding patterns, the amount of radioactivity in each complex is expressed as a percentage of the highest value for each given protein. To indicate the relative affinities of the different proteins, the amount of radioactivity in either the TAA or the TCA complexes was normalized to the GST-B–TAA or GST-BS1–TCA values, respectively (shown in parentheses).

[c] —, no retarded complex observed for protein-DNA combination given.

a more detailed analysis of the distinguishing features of the NGFI-B and SF-1 DNA sites, specifically the importance of the minor groove, by using the nucleotide inosine (I). Inosine, which forms Watson-Crick base pairs with C nucleotides, is identical in structure to G except that it lacks the exocyclic amine on position C-2 of the purine ring in the minor groove. Thus, an I/C base pair is identical to G/C in the major groove and A/T in the minor groove (36, 52). Substitution of the A/T base pair at position 3 of the NBRE with G/C completely abolished binding of GST-B, consistent with previous results (probe TAG, Fig. 4A [66]). However, a probe with I/C at position 3 supported GST-B binding at a level comparable to that of the NBRE itself [probe TA(I/C)]. Since no distortion of the DNA double helix was detected as a bend in the bound or unbound NBRE (data not shown), we conclude that the minor groove surface distinguishes an A/T and G/C base pair at position 3 of the NBRE. To confirm this, we used the modified pyrimidine 5-methyl cytosine (M), which pairs with G but has a methyl group at position 5 of the pyrimidine ring in the major groove, similar to T. As shown in Fig. 4A, a G/M base pair at position 3 also abolished GST-B binding [probe TA(G/M)], even though a key functional group of the A/T major groove was present.

At position 2 of the NBRE, substitution of the A/T base pair with G/C significantly decreased, but did not abolish, GST-B binding, also consistent with previous results (probe TGA, Fig. 4B [64]). Again, a substitution of I/C at position 2 allowed nearly complete GST-B binding [probe T(I/C)A], while G/M behaved similarly to G/C [probe T(G/M)A], indicating that the minor groove surface is also critical at position 2. Interestingly, GST-BS1 bound with a higher affinity to probes T(G/C)A and T(G/M)A than to either the NBRE or probe T(I/C)A, indicating that it too interacts with the minor groove, but with a preference for the G/C base pair. To verify this, we examined oligonucleotides with a transversion of the G/C base pair at position 2. Substitutions of the C/G base pair at position 2 of the probe TCA with C/I eliminated GST-BS1 binding [probe T(C/I)A], while probe T(M/G)A was bound as well as TCA. In contrast, substitution of C/I at position 2 conferred a slight but measurable degree of GST-B binding to probe TCA. Thus, NGFI-B and SF-1 respond in precisely opposite manners to changes at position 2 in the minor groove, with SF-1 favoring the presence and NGFI-B favoring the absence of the purine C-2 amino group.

**Verification of structure-function relationships in full-length NGFI-B.** To this point, nearly all of our studies addressing the structure-function relationships of monomer DNA binding have been performed with bacterial fusion proteins. It is necessary to verify our observations with full-length proteins. First, we translated in vitro both wild-type NGFI-B and mutant forms bearing mutations in the downstream domain (Fig. 5A). A focused replacement of the NGFI-B A box with that of SF-1 abolished binding to the NBRE but did not confer binding to TCA (Fig. 5B, fl-BS6). SF-1 binding was conferred only when the T box was changed in combination with the A box (fl-BS1). These results precisely mirror those obtained with bacterial fusion proteins. To verify that this same specificity is observed in cells, full-length NGFI-B mutants were tested in a transcriptional activation assay with promoters bearing two copies of either the NBRE or TCA sequences (AAA-luc and TCA-luc, respectively). As shown in Fig. 5C, the pattern of activity of these chimeric proteins closely mimics the pattern of DNA binding seen in vitro. Most notably, the SF-1 downstream domain conferred a greater promiscuity to NGFI-B so that activation was observed from both sites but with a preference for TCA.

## DISCUSSION

**A model of NGFI-B monomer binding.** NGFI-B binds to and activates transcription from a DNA response element that is atypical among the nuclear receptors characterized to

TABLE 3. Downstream domains of known and putative monomer binding receptors

| Nuclear receptor | Amino acid sequence[a] | |
|---|---|---|
| | T box | A box |
| AA consensus[b] | ...GM VKE VR SL G | RRGRL S K K |
| NGFI-B | ...GM VKEVVRTDSLKG | RRGRLPS KPKQPPDAS |
| NURR1 | ...GM ..... | ..... ...S.Q.P. |
| CEB-1 | ...GM ...I..HG..S. | .....S. .T.LARSED |
| TCA consensus[c] | ...GM +LEAVRADRMRG | GRNKFGP MYKRDRA K |
| SF-1 | ...GM RLEAVRADRMRG | GRNKFGP MYKRDRALK |
| FTZ-F1 | ...GM K..... | ..... .....R. |
| LRH[d] | ...GM K..... | ..... ..... |
| hERR1[d] | ...GM LK.G..L..V.. | ..Q.YKR RPEV.PLPF |
| hERR2[d] | ...GM LK.G..L..V.. | ..Q.YKR RLDSENSPY |
| DHR[d] | ...GM K...I.E..T.. | ..STYQC S.TLPNSML |
| Rev-ErbA-α[e] | ...GM SRDAVRFGRIPK | REKQRML AEMQSAMNL |
| E75[e] | ...GM .....V.. | ...A.I. .A..QSTQN |

[a] +, positively charged amino acids; ., an amino acid common to the respective reference sequence (i.e., NGFI-B, SF-1, or Rev-ErbA-α). A space indicates no consensus. References for the sequences can be found in the text.

[b] The AA consensus sequence was derived from the NGFI-B, NURR1, and CEB-1 sequences.

[c] The TCA consensus sequence was derived from the SF-1 and FTZ-F1 sequences.

[d] This sequence was obtained by comparing the SF-1 downstream domain with the GenBank data base (comparison performed at the National Center for Biotechnology Information with the BLAST network service). The binding preference of this receptor has not been firmly established.

[e] Rev-ErbA-α and its *Drosophila* homolog E75 have been implicated as monomer-binding receptors but lack strong sequence identity with either NGFI-B or SF-1.

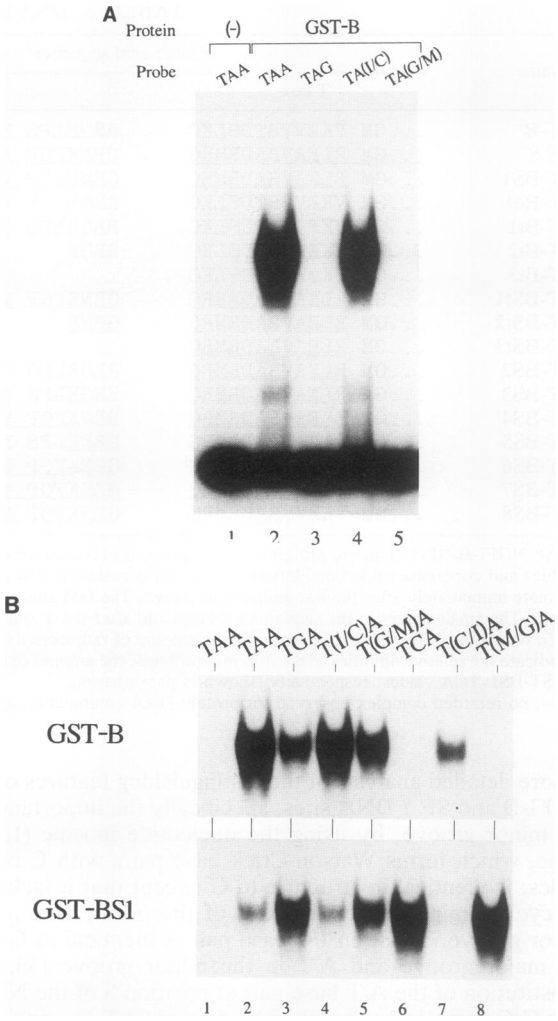


FIG. 4. NGFI-B and SF-1 contact the minor groove of DNA. (A) A gel shift assay was performed with a crude bacterial lysate containing GST-B and oligonucleotides with modified base substitutions at position 3 (Table 1). Lane 1 is probe TAA with no (−) added protein. (B) Gel shift assays were performed with crude bacterial lysates containing GST-B and GST-BS1 and oligonucleotides with modified base substitutions at position 2 (Table 1). Only the protein-DNA complexes are shown to facilitate comparison of the binding patterns of GST-B versus GST-BS1. Note that GST-B and GST-BS1 were incubated in separate sets of binding reactions and run on different gels. Lane 1 is probe TAA with no added protein.

date. This element contains only a single ER half-site and requires at least two A nucleotides 5′ to this half-site for high-affinity binding. We have used compensatory and gain-of-function mutations in the response element and DNA binding domain of NGFI-B to establish the mechanism of this binding event. Chimeras with the DNA binding domains of NGFI-B and GR demonstrated that specificity at the half-site of a target probe can be altered by swapping the zinc modules. This indicates that the NGFI-B zinc modules interact with the ER half-site in the NBRE in the manner demonstrated for the steroid receptors, which fixes the spatial relationship of the protein to the DNA. Further chimeras with the SF-1 protein demonstrated that the proximal residues of the A box make contacts with the minor groove at the 5′ adjacent nucleotides. Thus, amino acids from only one NGFI-B molecule can account for all of the necessary nucleotide contacts revealed by methylation interference of NBRE binding.

The model of the NGFI-B–NBRE complex that emerges from these data is presented in Fig. 6A. In this model, a single DNA binding domain monomer is docked at B form DNA with the known orientation of the GR DNA binding domain and half-site (34). The precise contacts made between the NGFI-B zinc modules and the half-site are not known but are very likely the same as those used by ER, since the P box residues that define this contact are common between NGFI-B and ER (59). With this docking, it is apparent that the domain downstream of the zinc modules is in position to contact the minor groove of the nucleotides 5′ to the half-site. The differential response of NGFI-B and SF-1 to probes containing inosine at position 2 indicates that the exocyclic amine in the G/C minor groove is necessary for high-affinity SF-1 binding, whereas it interferes with NGFI-B binding. The key difference in the side chains of the proximal A box of NGFI-B and SF-1, given that glycine has no side chain, is the substitution of the asparagine amido group of SF-1 for the arginine amino groups of NGFI-B (Fig. 6B). It is likely that the SF-1 asparagine accepts a hydrogen bond from the G amine, an interaction which is not possible in NGFI-B given the lack of hydrogen bond acceptors in the proximal A box. Rather, the arginine(s) in NGFI-B could act as donors to the purine N-3 or the pyrimidine C-2 carbonyl, which are exposed in the minor groove at A/T base pairs but

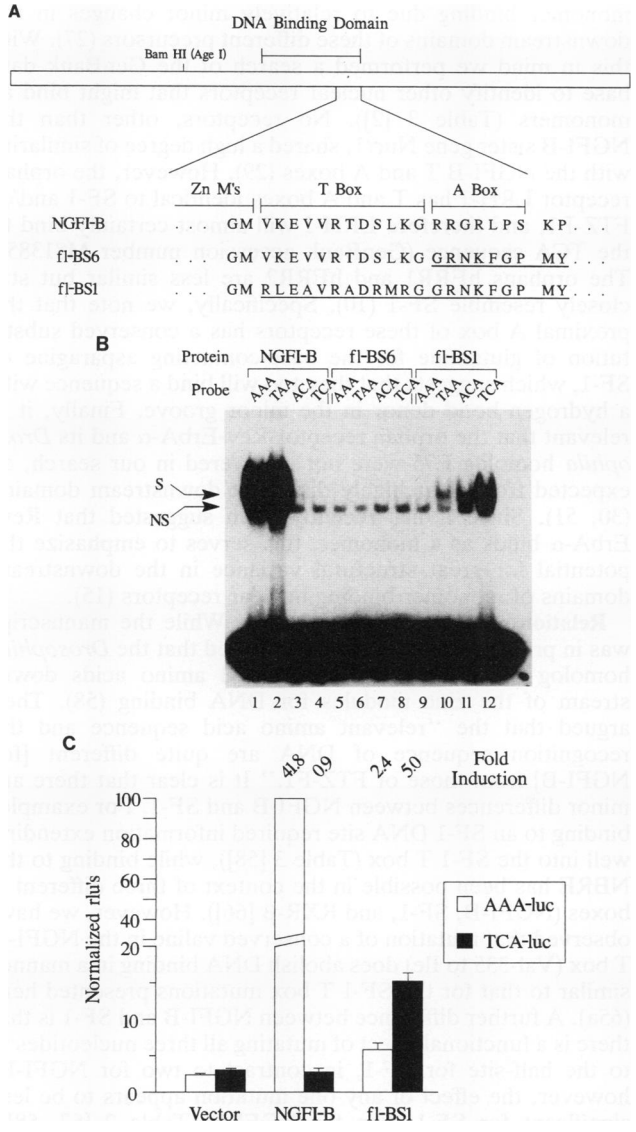


FIG. 5. The A box structure-function relationship applies to full-length NGFI-B. (A) A schematic showing the full-length NGFI-B variants tested in panels B and C. These constructs contained a *Bam*HI linker inserted at the *Age*I site of NGFI-B, which does not affect the activity of the encoded protein (43). The sequences of the T and A boxes are shown, with no underline indicating an NGFI-B residue and an underline indicating an SF-1 residue. (B) A gel shift assay was performed with the proteins shown in panel A (translated in vitro) and oligonucleotide probes shown in Table 1. NS indicates a nonspecific complex that arises from the reticulocyte lysate used to translate the proteins. S indicates the specific complex corresponding to the NGFI-B protein variants. (C) CV-1 cells were transfected with plasmids that express the proteins shown in panel A and luciferase reporter plasmids bearing two copies of the AAA or TCA sequences (Table 1). To allow for comparison between experiments, the relative luciferase units (rlu's) obtained for each combination were normalized to the values of the NGFI-B–AAA pair. Note the discontinuity in scale. Fold induction is expressed relative to the appropriate vector-reporter combination, such that a fold induction of one signifies no increase over background.

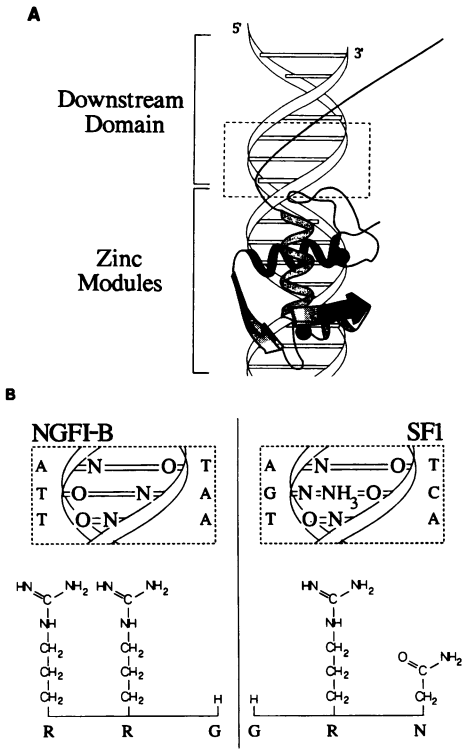


FIG. 6. A model of monomer-DNA binding. (A) Top view of a nuclear receptor DNA binding domain monomer-DNA complex, patterned after Fig. 3A of reference 34, docked at a half-site in the major groove. The zinc modules may be thought of as a prediction of the tertiary structure of the NGFI-B zinc modules or as the GR zinc modules of the GST-BG protein used in Fig. 2. The T and A boxes of the downstream domain are shown in an extended conformation, since no prediction can be made regarding their secondary structure, reaching over and beyond the minor groove of the 5′ adjacent base pairs. The dashed rectangle denotes the region detailed in panel B. (B) The minor groove of the 5′ adjacent base pairs of the optimal NGFI-B and SF-1 binding sites is shown, with the sequence of each strand indicated to the side. Within each minor groove are the functional groups available for hydrogen bonding, corresponding to the S-1, S-1′ and S-2 sites described in reference 50. These are the purine N-3, pyrimidine position 2 carbonyl and G position 2 amine. Also shown are the amino acid side chains of the proximal A boxes of NGFI-B and SF-1. Note the correspondence of the amine hydrogen bond donor in the SF-1 minor groove with the asparagine hydrogen bond acceptor in the SF-1 A box.

potentially are blocked by the exocyclic amine in G/C base pairs.

What is less apparent in our model is the precise secondary structure in the downstream domain. Some folding is clearly necessary, since the A box reaches far beyond the three adjacent nucleotide pairs in the extended conformation shown in Fig. 6A. Our results suggest that this structure will involve the T box region, since mutations within this motif can abolish monomer binding but do not affect sequence specificity. It is of interest to note that the T box of RXR is thought to function by making interpeptide contacts with a second protein monomer (66). These dimerization contacts may be mimicked in NGFI-B by intrapeptide contacts between the T and A boxes, which would be facilitated by a turn at the glycine just upstream of the A box. Such contacts would serve to stabilize the region as well as position the proximal A box near the minor groove. A similar situation

has been observed in an artificial variant of the λ Cro protein, in which it was possible to create a stable monomer binding protein simply by ensuring that the dimerization interface of the molecule formed its essential contacts in an intramolecular fashion (39). It must be noted though that the phenotype observed for our T box mutants is also consistent with the existence of T box-DNA contacts that are common between NGFI-B and SF-1, either nonspecific or with the common position 3 A/T base pair.

In total, it is apparent that contacts between the A box amino acids and the minor groove serve to stabilize the overall interaction of an NGFI-B monomer with DNA. These contacts provide the required binding energy, in addition to that of the zinc module–half-site interaction, that is more typically provided in nuclear receptors by dimerization. It must be emphasized though that the observation of NGFI-B monomer DNA binding in vitro does not alone indicate that NGFI-B functions as a monomer in vivo. This link is established by the demonstrations that (i) NGFI-B can activate transcription from a contrived promoter containing only a single NBRE (64) and (ii) the 21-hydroxylase promoter contains a functional NBRE that shares all of the structural features indicative of monomer binding (65). These results notwithstanding, we consider it likely that NGFI-B can interact with other transcription factors (perhaps nuclear receptors such as SF-1 or NGFI-B itself) bound to nearby promoter elements (43). Also, the results presented here cannot rule out the possibility that a second NGFI-B monomer is part of the protein-DNA complex observed by gel shift, interacting with DNA nonspecifically or not at all. Importantly, though, such higher-order interactions would likely be qualitatively different from the close, spatially rigid and DNA-driven interactions that are needed by dimerizing proteins for DNA binding per se.

This work establishes that there are at least three functionally distinct mechanisms by which nuclear receptors bind to DNA. First, the prototypic steroid receptors bind as dimers to inverted half-site repeats by virtue of D box-D box interactions. Second, the class of nuclear receptors typified by RXR binds to direct (and sometimes everted [54]) half-site repeats by virtue of interactions involving the T box, with many members of this class binding as heterodimers with RXR. Finally, we have demonstrated that NGFI-B and SF-1 represent a novel class of nuclear receptors that bind as monomers by virtue of A box interactions with a 5′ extended half-site. It is curious that the thyroid hormone receptor is at present the only clear example of a nuclear receptor that can bind by more than one of these paradigms (11, 40, 59, 60). The three mechanisms are not mutually exclusive within a given protein, since they are all underscored by the common interaction of the zinc modules with a half-site and since we have created two different bifunctional proteins by fusing the appropriate monomer- and dimer-promoting motifs (Fig. 2 [66]). It is thus an intriguing possibility that NGFI-B and other nuclear receptors might bind to DNA as both monomers and dimers, given the appropriate DNA sites, and that regulatory factors might affect the function of the protein by modifying its preference for these sites.

Structural comparisons reveal that NGFI-B bears the highest degree of amino acid identity in the zinc modules with RXR, a similarity that is apparent functionally in NGFI-B as a latent ability to support dimer binding when the appropriate T box is added (66). In contrast, NGFI-B and SF-1 share little conservation in the zinc modules beyond that of all nuclear receptors, despite their common mode of DNA binding. These observations are most consistent with the primary establishment of distinct dimerizing receptor families followed by the parallel, independent appearance of monomer binding due to relatively minor changes in the downstream domains of these different precursors (27). With this in mind we performed a search of the GenBank data base to identify other nuclear receptors that might bind as monomers (Table 3 [2]). No receptors, other than the NGFI-B sister gene Nurr1, shared a high degree of similarity with the NGFI-B T and A boxes (29). However, the orphan receptor LRH-1 has T and A boxes identical to SF-1 and/or FTZ-F1, and therefore LRH-1 will almost certainly bind to the TCA sequence (GenBank accession number M81385). The orphans hERR1 and hERR2 are less similar but still closely resemble SF-1 (10). Specifically, we note that the proximal A box of these receptors has a conserved substitution of glutamine for the base-contacting asparagine of SF-1, which suggests that they too will bind a sequence with a hydrogen bond donor in the minor groove. Finally, it is relevant that the orphan receptor Rev-ErbA-α and its *Drosophila* homolog E75 were not uncovered in our search, as expected from their highly divergent downstream domains (30, 51). Since it has recently been suggested that Rev-ErbA-α binds as a monomer, this serves to emphasize the potential for great structural variance in the downstream domains of monomer binding nuclear receptors (15).

**Relationship of NGFI-B and SF-1.** While the manuscript was in preparation, Ueda et al. reported that the *Drosophila* homolog of SF-1, FTZ-F1, required amino acids downstream of the zinc modules for DNA binding (58). They argued that the "relevant amino acid sequence and the recognition sequence of DNA are quite different [for NGFI-B] from those of FTZ-F1." It is clear that there are minor differences between NGFI-B and SF-1. For example, binding to an SF-1 DNA site required information extending well into the SF-1 T box (Table 3 [58]), while binding to the NBRE has been possible in the context of three different T boxes (NGFI-B, SF-1, and RXR-β [66]). However, we have observed that mutation of a conserved valine in the NGFI-B T box (Val-335 to Ile) does abolish DNA binding in a manner similar to that for the SF-1 T box mutations presented here (65a). A further difference between NGFI-B and SF-1 is that there is a functional effect of mutating all three nucleotides 5′ to the half-site for SF-1, in contrast to two for NGFI-B; however, the effect of any one mutation appears to be less significant for SF-1 than for NGFI-B (Table 2 [57, 58]). Nonetheless, the nucleotide sequences bound by NGFI-B and SF-1 are extremely similar, since a differential specificity of binding can be obtained by changing a single functional group of 1 bp. Furthermore, the A box amino acids that make the essential DNA contacts localize to precisely the same position relative to the NGFI-B and SF-1 zinc modules. We therefore argue that the mechanism of DNA binding by NGFI-B and SF-1 is fundamentally the same since they share the defining structural features of the DNA and protein.

It is remarkable that the two nuclear receptors characterized to date that bind to DNA as monomers converge on the activity of common promoters. SF-1 was originally identified as a nuclear factor expressed in adrenocortical cells that bound to several single-half-site elements in the promoter of the 21-hydroxylase gene (25). One of these, the −65 element, has recently been demonstrated to be a functional NBRE that may mediate ACTH induction of this promoter (65). It is also clear that the promoters for other steroidogenic enzymes in the adrenal cortex and other tissues have single-half-site elements, some corresponding to NBREs (4, 7, 46,

47). At least three, and probably more, nuclear proteins bind to these elements, one of which may be related to COUP-TF (48). The more promiscuous nature of SF-1 binding is borne out at these sites, because NGFI-B binds to only a subset of the elements bound by SF-1, in a manner predictable by the 5′ adjacent nucleotides (references 25 and 65 and unpublished observations). It follows that there will exist a competition for binding to these sites between proteins that bind more promiscuously, like the constitutively expressed SF-1 protein, and those with a more restricted binding pattern, like the growth factor-inducible NGFI-B. The final transcriptional outcome of this competition would be determined by the nucleotides 5′ to the half-site and by the combination and perhaps interaction of the nuclear receptors that are expressed and bound.

## ACKNOWLEDGMENTS

We thank Keith Parker, Mitch Lazar, and Vincent Laudet for helpful discussions during the preparation of the manuscript.

This work was supported by a grant from Ronald McDonald Children's Charities, Public Health Service grant NS01018 from the National Institute of Neurological and Communicative Disorders and Stroke, and grant P01 CA49712 from the National Cancer Institute. J.M. is an Established Investigator of the American Heart Association, and T.E.W. is an awardee of a Merck Sharpe & Dohme Academic Development Program fellowship.

## REFERENCES

1. **Allan, G. F., X. Leng, S. Y. Tsai, N. L. Weigel, D. P. Edwards, M. J. Tsai, and B. W. O'Malley.** 1992. Hormone and antihormone induce distinct conformational changes which are central to steroid receptor activation. J. Biol. Chem. **267**:19513–19520.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.
3. **Berg, J. M.** 1990. Zinc finger domains: hypotheses and current knowledge. Annu. Rev. Biophys. Biophys. Chem. **19**:405–421.
4. **Bogerd, A. M., A. Franklin, D. A. Rice, B. P. Schimmer, and K. L. Parker.** 1990. Identification and characterization of two upstream elements that regulate adrenocortical expression of steroid 11B-hydroxylase. Mol. Endocrinol. **4**:845–850.
5. **Dreyer, C., G. Krey, H. Keller, F. Givel, G. Helftenbein, and W. Wahli.** 1992. Control of the peroxisomal beta-oxidation pathway by a novel family of nuclear hormone receptors. Cell **68**:879–887.
6. **Fahrner, T. J., S. L. Carroll, and J. Milbrandt.** 1990. The NGFI-B protein, an inducible member of the thyroid/steroid receptor family, is rapidly modified posttranslationally. Mol. Cell. Biol. **10**:6454–6459.
7. **Fitzpatrick, S. L., and J. S. Richards.** 1993. Cis-acting elements of the rat aromatase promoter required for cyclic adenosine 3′,5′-monophosphate induction in ovarian granulosa cells and constitutive expression in R2C Leydig cells. Mol. Endocrinol. **7**:341–354.
8. **Freedman, L. P., B. F. Luisi, Z. R. Korszun, R. Basavappa, P. B. Sigler, and K. R. Yamamoto.** 1988. The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. Nature (London) **334**:543–546.
9. **Giguere, V., E. S. Ong, P. Segui, and R. M. Evans.** 1987. Identification of a receptor for the morphogen retinoic acid. Nature (London) **330**:624–629.
10. **Giguere, V., N. Yang, P. Segui, and R. M. Evans.** 1988. Identification of a new class of steroid hormone receptors. Nature (London) **331**:91–94.
11. **Glass, C. K., J. M. Holloway, O. V. Devary, and M. G. Rosenfeld.** 1988. The thyroid hormone receptor binds with opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. Cell **54**:313–323.
12. **Gottlicher, M., E. Widmark, Q. Li, and J. A. Gustafsson.** 1992. Fatty acids activate a chimera of the clofibric acid-activated receptor and the glucocorticoid receptor. Proc. Natl. Acad. Sci. USA **89**:4653–4657.
13. **Gronemeyer, H., B. Turcotte, C. Quirin-Stricker, M. T. Bocquel, M. E. Meyer, Z. Krozowski, J. M. Jeltsch, T. Lerouge, J. M. Garnier, and P. Chambon.** 1988. The chicken progesterone receptor: sequence, expression and functional analysis. EMBO J. **6**:3985–3994.
14. **Hard, T., E. Kellenbach, R. Boelens, B. A. Maler, K. Dahlman, L. P. Freedman, J. Carlstedt-Duke, K. R. Yamamoto, J. Gustafsen, and R. Kaptein.** 1990. Solution structure of the glucocorticoid receptor DNA-binding domain. Science **249**:157–160.
15. **Harding, H. P., and M. A. Lazar.** 1993. The orphan receptor Rev-ErbAα activates transcription via a novel response element. Mol. Cell. Biol. **13**:3113–3121.
16. **Hazel, T. G., D. Nathans, and L. F. Lau.** 1988. A gene inducible by serum growth factors encodes a member of the steroid and thyroid hormone receptor superfamily. Proc. Natl. Acad. Sci. USA **85**:8444–8448.
17. **Heyman, R. A., D. J. Mangelsdorf, J. A. Dyck, R. B. Stein, G. Eichele, R. M. Evans, and C. Thaller.** 1992. 9-cis retinoic acid is a high affinity ligand for the retinoid X receptor. Cell **68**:397–406.
18. **Hollenberg, S. M., V. Giguere, P. Segui, and R. M. Evans.** 1987. Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor. Cell **49**:39–46.
19. **Jost, J. P., M. Seldran, and M. Geiser.** 1984. Preferential binding of estrogen-receptor complex to a region containing the estrogen-dependent hypomethylation site preceding the chicken vitellogenin II gene. Proc. Natl. Acad. Sci. USA **81**:429–433.
20. **Karin, M., A. Haslinger, H. Holtgreve, R. I. Richards, P. Krauter, H. M. Westphal, and M. Beato.** 1984. Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-IIA gene. Nature (London) **308**:513–519.
21. **Kliewer, S. A., K. Umesono, D. J. Mangelsdorf, and R. M. Evans.** 1992. Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. Nature (London) **355**:446–449.
22. **Koelle, M. R., W. S. Talbot, W. A. Segraves, M. T. Bender, P. Cherbas, and D. S. Hogness.** 1991. The Drosophila EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. Cell **67**:59–77.
23. **Kumar, V., and P. Chambon.** 1988. The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. Cell **55**:145–156.
24. **Kumar, V., S. Green, G. Stack, M. Berry, J. Jin, and P. Chambon.** 1987. Functional domains of the human estrogen receptor. Cell **51**:941–951.
25. **Lala, D. S., D. A. Rice, and K. L. Parker.** 1992. Steroidogenic factor I, a key regulator of steroidogenic enzyme expression, is the mouse homolog of fushi tarazu-factor I. Mol. Endocrinol. **6**:1249–1258.
26. **Latchman, D. S.** 1990. Eukaryotic transcription factors. Biochem. J. **270**:281–289.
27. **Laudet, V., C. Hanni, J. Coll, F. Catzeflis, and D. Stehlin.** 1992. Evolution of the nuclear receptor gene superfamily. EMBO J. **11**:1003–1013.
28. **Lavorgna, G., H. Ueda, J. Clos, and C. Wu.** 1991. FTZ-F1, a steroid hormone receptor-like protein implicated in the activation of fushi tarazu. Science **252**:848–851.
29. **Law, S. W., O. M. Conneely, F. J. DeMayo, and B. W. O'Malley.** 1992. Identification of a new brain-specific transcription factor, NURR1. Mol. Endocrinol. **6**:2129–2135.
30. **Lazar, M. A., R. A. Hodin, D. S. Darling, and W. W. Chin.** 1989. A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-*erbA*α transcriptional unit. Mol. Cell. Biol. **9**:1128–1136.
31. **Lees, J. A., S. E. Fawell, and M. G. Parker.** 1989. Identification of two transactivation domains in the mouse oestrogen receptor. Nucleic Acids Res. **17**:5477–5488.
32. **Leid, M., P. Kastner, R. Lyons, H. Nakshatri, M. Saunders, T.**

Zacharewski, J. Y. Chen, A. Staub, J. M. Garnier, and S. Mader. 1992. Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently. Cell 68:377–395.
33. Levin, A. A., L. J. Sturzenbecker, S. Kazmer, T. Bosakowski, C. Huselton, G. Allenby, J. Speck, C. Kratzeisen, M. Rosenberger, and A. Lovey. 1992. 9-cis retinoic acid stereoisomer binds and activates the nuclear receptor RXR alpha. Nature (London) 355:359–361.
34. Luisi, B. F., W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto, and P. B. Sigler. 1991. Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature (London) 352:497–505.
35. Mangelsdorf, D. J., K. Umesono, S. A. Kliewer, U. Borgmeyer, E. S. Ong, and R. M. Evans. 1991. A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. Cell 66:555–561.
36. Martin, F. H., M. M. Castro, F. Aboul-ela, and I. Tinoco, Jr. 1985. Base pairing involving deoxyinosine: implications for probe design. Nucleic Acids Res. 13:8927–8938.
37. McDonnell, D. P., D. J. Mangelsdorf, J. W. Pike, M. R. Haussler, and B. W. O'Malley. 1987. Molecular cloning of complementary DNA encoding the avian receptor for vitamin D. Science 235:1214–1217.
38. Milbrandt, J. 1988. Nerve growth factor induces a gene homologous to the glucocorticoid receptor gene. Neuron 1:183–188.
39. Mossing, M. C., and R. T. Sauer. 1990. Stable, monomeric variants of Lambda Cro obtained by insertion of a designed B-hairpin sequence. Science 250:1712–1715.
40. Naar, A. M., J. M. Boutin, S. M. Lipkin, V. C. Yu, J. M. Holloway, C. K. Glass, and M. G. Rosenfeld. 1991. The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. Cell 65: 1267–1279.
41. O'Malley, B. W., and B. W. Conneely. 1992. Orphan receptors: in search of a unifying hypothesis for activation. Mol. Endocrinol. 6:1359–1361.
42. Otri, E., J. E. Bodwell, and A. Munck. 1992. Phosphorylation of steroid hormone receptors. Endocrinol. Rev. 13:105–128.
43. Paulsen, R. E., C. A. Weaver, T. J. Fahrner, and J. Milbrandt. 1992. Domains regulating transcriptional activity of the inducible orphan receptor NGFI-B. J. Biol. Chem. 267:16491–16496.
44. Pfahl, M. 1982. Specific binding of the glucocorticoid-receptor complex to the mouse mammary tumor proviral promoter region. Cell 31:475–482.
45. Pratt, W. B., L. C. Scherrer, K. A. Hutchison, and F. C. Dalman. 1992. A model of glucocorticoid receptor unfolding and stabilization by a heat shock protein complex. J. Steroid Biochem. Mol. Biol. 41:223–229.
46. Rice, D. A., S. Kirkman, L. D. Aitken, A. R. Mouw, B. P. Schimmer, and K. L. Parker. 1990. Analysis of the promoter region of the gene encoding mouse cholesterol side-chain cleavage enzyme. J. Biol. Chem. 265:11713–11720.
47. Rice, D. A., M. S. Kronenberg, A. R. Mouw, L. D. Aitken, A. Franklin, B. P. Schimmer, and K. L. Parker. 1990. Multiple regulatory elements determine adrenocortical expression of steroid 21-hydroxylase. J. Biol. Chem. 265:8052–8058.
48. Rice, D. A., A. R. Mouw, A. Bogerd, and K. L. Parker. 1991. A shared promoter element regulates the expression of three steroidogenic enzymes. Mol. Endocrinol. 5:1552–1561.
49. Schwabe, J. W. R., D. Neuhaus, and D. Rhodes. 1990. Solution structure of the DNA-binding domain of the oestrogen receptor. Nature (London) 348:458–461.
50. Seeman, N. C., J. M. Rosenberg, and A. Rich. 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. USA 73:804–808.
51. Segraves, W. A., and D. S. Hogness. 1990. The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila encodes two new members of the steroid receptor superfamily. Genes Dev. 4:204–219.
52. Starr, B. D., and D. K. Hawley. 1991. TFIID binds in the minor groove of the TATA box. Cell 67:1231–1240.
53. Sucov, H. M., K. K. Murakami, and R. M. Evans. 1990. Characterization of an autoregulated response element in the mouse retinoic acid receptor type β gene. Proc. Natl. Acad. Sci. USA 87:5392–5396.
54. Tini, M., G. Otulakowski, M. L. Breitman, L. C. Tsui, and V. Giguere. 1993. An everted repeat mediates retinoic acid induction of the gamma F-crystallin gene: evidence of a direct role for retinoids in lens development. Genes Dev. 7:295–307.
55. Tsai, S. Y., J. Carlsted-Duke, N. L. Weigel, K. Dahlman, J. Gustafson, M. Tsai, and B. W. O'Malley. 1988. Molecular interactions of steroid hormone receptor with its enhancer element: evidence for receptor dimer formation. Cell 55:361–369.
56. Tsukiyama, T., H. Ueda, S. Hirose, and O. Niwa. 1992. Embryonal long terminal repeat-binding protein is a murine homolog of FTZ-F1, a member of the steroid receptor superfamily. Mol. Cell. Biol. 12:1286–1291.
57. Ueda, H., and S. Hirose. 1991. Defining the sequence recognized with BmFTZ-F1, a sequence specific DNA binding factor in the silkworm, Bombyx mori, as revealed by direct sequencing of bound oligonucleotides and gel mobility shift competition analysis. Nucleic Acids Res. 19:3689–3693.
58. Ueda, H., G.-C. Sun, T. Murata, and S. Hirose. 1992. A novel DNA-binding motif abuts the zinc finger domain of insect nuclear hormone receptor FTZ-F1 and mouse embryonal long terminal repeat-binding protein. Mol. Cell. Biol. 12:5667–5672.
59. Umesono, K., and R. M. Evans. 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57:1139–1146.
60. Umesono, K., K. K. Murakami, C. C. Thompson, and R. M. Evans. 1991. Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell 65:1255–1266.
61. Watson, M. A., and J. Milbrandt. 1989. The NGFI-B gene, a transcriptionally inducible member of the steroid receptor gene superfamily: genomic structure and expression in rat brain after seizure induction. Mol. Cell. Biol. 9:4213–4219.
62. Weinberger, C., S. M. Hollenberg, E. S. Ong, J. M. Harmon, S. T. Brower, J. Cidlowski, E. B. Thompson, M. G. Rosenfeld, and R. M. Evans. 1985. Identification of human glucocorticoid receptor complementary DNA clones by epitope selection. Science 228:740–742.
63. Weinberger, C., C. C. Thompson, E. S. Ong, R. Lebo, D. J. Gruol, and R. M. Evans. 1986. The c-erb-A gene encodes a thyroid hormone receptor. Nature (London) 324:641–646.
64. Wilson, T. E., T. J. Fahrner, M. Johnston, and J. Milbrandt. 1991. Identification of the DNA binding site for NGFI-B by genetic selection in yeast. Science 252:1297–1300.
65. Wilson, T. E., A. R. Mouw, C. A. Weaver, J. Milbrandt, and K. L. Parker. 1993. The orphan nuclear receptor NGFI-B regulates expression of the gene encoding steroid 21-hydroxylase. Mol. Cell. Biol. 13:861–868.
65a. Wilson, T. E., K. A. Padgett, M. Johnston, and J. Milbrandt. A genetic method for defining DNA-binding domains: application to the nuclear receptor NGFI-B. Proc. Natl. Acad. Sci. USA, in press.
66. Wilson, T. E., R. E. Paulsen, K. A. Padgett, and J. Milbrandt. 1992. Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. Science 256:107–110.
67. Yu, V. C., C. Delsert, B. Andersen, J. M. Holloway, O. V. Devary, A. M. Naar, S. Y. Kim, J. M. Boutin, C. K. Glass, and M. G. Rosenfeld. 1991. RXR beta: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. Cell 67:1251–1266.
68. Zhang, X. K., B. Hoffmann, P. B. Tran, G. Graupner, and M. Pfahl. 1992. Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. Nature (London) 355:441–446.