Analysis of Estrogen Response Element Binding by Genetically Selected Steroid Receptor DNA Binding Domain Mutants Exhibiting Altered Specificity and Enhanced Affinity*

(Rceived for publication, April 28, 1999, and in revised form, June 2, 1999)

Sudsanguan Chusacultanachai‡§, Kevin A. Glenn†¶, Adrian O. Rodriguez†, Erik K. Read**+, Jeffrey F. Gardner‡, Benita S. Katzenellenbogen†¶, and David J. Shapiro‡§§

From the Departments of ‡Biochemistry, ¶¶Molecular and Integrative Physiology, and **Microbiology, University of Illinois, Urbana, Illinois 61801

To analyze the role of amino acids in the steroid receptor DNA binding domain (DBD) recognition helix in binding of the receptor to the estrogen response element (ERE), we adapted the powerful P22 challenge phage selection system for use with a vertebrate protein. We used the progesterone receptor DNA binding domain and selected for mutants that gained the ability to bind to the ERE. We used a mutagenesis protocol based on degenerate oligonucleotides to create a large and diverse pool of mutants in which 10 nonconsensus amino acids in the DNA recognition helix of the progesterone receptor DNA binding domain were randomly mutated. After a single cycle of modified P22 challenge phage selection, 37 mutant proteins were identified, all of which lost the ability to bind to the progesterone response element. In gel mobility shift assays, approximately 70% of the genetically selected mutants bound to the consensus ERE with a 4-fold higher affinity than the naturally occurring estrogen receptor DBD. In the P-box region of the DNA recognition helix, the selected mutants contained the amino acids found in the wild-type estrogen receptor DBD, as well as other amino acid combinations seen in naturally occurring steroid/nuclear receptors that bind the aGGTCA half-site. We also obtained high affinity DBDs with Trp585 as the first amino acid of the P-box, although this is not found in the known steroid/nuclear receptors. In the linker region between the two zinc fingers, G597R was by far the most common mutation. In transient transfections in mammalian cells using promoter interference assays, the mutants displayed enhanced affinity for the ERE. When linked to an activation domain, the transduced mutants activated transcription from ERE-containing reporter genes.

We conclude that the P-box amino acids can display considerable variation and that the little studied linker amino acids play an important role in determining affinity for the ERE. This work also demonstrates that the P22 challenge phage genetic selection system, modified for use with a mammalian protein, provides a novel, single cycle selection for steroid/nuclear receptor DBDs with altered specificity and greatly enhanced affinity for their response elements.

Steroid/nuclear receptors (1–4) and many transcription factors belong to protein superfamilies whose members bind to related, but distinct, DNA sequences. Individual proteins within the superfamily must bind to their DNA response elements with high specificity and affinity. The steroid/nuclear receptors bind to a specific DNA sequence, termed a hormone response element (HRE). In general, HREs are composed of two core sequences 5′-AGNNCA-3′ that are separated by a spacer region of 0–6 nucleotides and are arranged as either a direct repeat or an inverted or everted palindrome.

Recognition of HREs by steroid/nuclear receptors is mediated through a DNA binding domain (DBD) of 65–70 amino acids. The core DBD is highly conserved (3). Structural analyses of several DBDs (5–8) showed that they usually contain two independent zinc finger motifs connected by a short flexible amino acid linker, with an amphipathic α-helix near the C terminus of each finger. The first helix in the N-terminal zinc finger, called the DNA recognition helix, is important for specific DNA binding. Upon interaction of the DBD with the HRE, amino acid side chains in the recognition helix make sequence-specific contacts with nucleotides exposed in the major groove of the DNA. A dimerization surface, called the D-box, found in the second helix allows the DBD to recognize the two HRE half-sites as a dimer.

Mutational analyses (9, 10) and structural comparisons suggested that the ability of the estrogen receptor (ER), the glucocorticoid receptor (GR), and the progesterone receptor (PR) (5, 6) to discriminate between their respective HREs is at least partially due to three amino acids in the DNA recognition helix of the DBD, called the P-box (11). However, further analysis demonstrated that not all of the side chains of the defined P-box triplet contact the bases of their DNA target. Many contacts involve nucleotides common to both the estrogen re-

* This work was supported by National Institutes of Health (NIH) Grant HD-16720 and USAMRMC Breast Cancer Research Program Grant 17-HD-1-17241 (to D. J. S.), NIH Grant CA 60514 (to B. S. K.), NIH Grant GM 28717 (to J. F. G.), and an NCI, NIH, Minority Predoctoral Fellowship (to A. O. R.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: National Center for Genetic Engineering and Biotechnology, 5392 Gypsum Metropolitan Tower, Bangkok, 10400 Thailand.
‡ Present address: Dept. of Internal Medicine, the University of Iowa Hospitals, Iowa City, IA.
** Present address: Wadsworth Anaerobe Laboratory (151J), West Los Angeles VAMC (Bldg. 304), 11301 Wilshire Blvd., Los Angeles, CA 90073.
† To whom correspondence should be addressed: Dept. of Biochemistry, 413 RAL, 600 S. Mathews, Urbana, IL 61801. Tel.: 217-333-1788; Fax: 217-244-5858; E-mail: djspahip@uiuc.edu.

†† Present address: Technology, 413 RAL, 600 S. Mathews, Urbana, IL 61801.

The abbreviations used are: HRE, hormone response element; DBD, DNA binding domain; ER, estrogen receptor; ERE, estrogen response element; GR, glucocorticoid receptor; PR, progesterone receptor; PREGRE, progesterone response element/glucocorticoid response element; cERE, consensus estrogen response element; CMV, cytomegalovirus; CAT, chloramphenicol acetyltransferase; WT, wild type.
Amino Acids Used by Genetically Selected DBDs to Bind ER

sponse element (ERE) and the progesterone response element/ glucocorticoid response element (PRE/GR) and involve contacts with the side chains of amino acids conserved in the ER-DBDs and GR-DBDs. Of course, several other factors play a role in DNA binding, including steric hindrance, expulsion of water molecules or ions (12, 13), and alterations of the DNA conformation (13, 14) upon DNA-protein interaction.

The reduced ability to activate transcription of a mutant ER in which the P-box amino acid triplet has been changed to alter binding specificity from the ERE to the PRE/GR (9) suggests that additional amino acids may play a role in determining affinity for the HRE. We therefore employed discrimination between the ERE and the PRE/GR as a system for identifying additional amino acids important in binding of a DBD to an HRE.

In the natural process of protein evolution and selection, proteins containing random mutations that confer an advantage on the cell are selected from the large numbers of neutral or deleterious mutations that occur over time. To simulate the process of natural selection in shifting DNA binding specificity from the ERE to the PRE/GR, we needed both a system for producing large numbers of mutants with random amino acid changes and a powerful selection for the relatively rare mutant DBDs exhibiting the desired ERE binding properties. We developed a rapid and simple procedure for saturation mutagenesis of a short region of a protein using degenerate oligonucleotides and Phi DNA polymerase. To select the mutants from this large mutant pool that had gained the ability to recognize the ERE, we adapted the powerful P22 challenge phage (15–17) system for use with a vertebrate protein. In the P22 system, substantial numbers of mutants are screened in a single selection cycle using a life-death selection. In this work, we show that the P22 challenge phage selection system can be used to select for mutants exhibiting a substantial change in DNA binding specificity. The P22 challenge phage selection system provides a new tool for engineering steroid/nuclear receptor DBDs with a desired DNA binding specificity and affinity.

To facilitate identification of amino acids important in discriminating between the ERE and the PRE/GR, it seemed critical to identify the amino acid changes that accompany a shift in DNA binding specificity from the PRE/GR to the ERE. Because the PR-DBD binds to the ERE with a higher specificity than the GR-DBD (18), we employed the PR-DBD in these studies.

We selected and identified mutant PR-DBDs containing amino acid sequences exhibiting high affinity binding to the ERE. We find that the first and third P-box amino acids are the most critical residues for DNA binding specificity, and that mutation of amino acids in the linker region can lead to DBDs whose affinity for the ERE is severalfold higher than that exhibited by the wild-type ER-DBD.

**EXPERIMENTAL PROCEDURES**

**Strains—**Salmonella typhimurium LT2:MS 1582 carrying P22 c2+ mnt+ prophage MS1868, MS1893 (15) and phage P22 mnt+:Kan9 arabinose, 75 µg/ml Timentin (Smith-Kline Beecham, Philadelphia, PA), and incubated overnight at 37 °C. Plasmids from lysogens grown on selective medium were transformed into E. coli DH5α to prepare DNA for sequencing.

**Identification of Specificity Switch Enhanced Affinity Mutants Using Challenge Phage Selection—**S. typhimurium MS1888 was transformed by electroporation with the pool of mutated PR-DBD DNA, plated on LB plates containing 0.2% glucose and 75 µg/ml Timentin, and incubated overnight at 37 °C. For each challenge phage assay, ~5,000 colonies were pooled and grown in LB liquid media containing 0.2% glucose and 75 µg/ml Timentin to an A600 of ~0.6. Bacteria were pelleted and resuspended in LB medium containing 1% arabinose and 75 µg/ml Timentin and an A600 of ~0.2. 100 µl of cells were mixed with the P22-ERE phage lysate at a multiplicity of infection of ~25 and incubated at room temperature for 30 min. The infected cells were plated on LB agar containing 1% arabinose, 75 µg/ml Timentin, and 50 µg/ml kanamycin and incubated overnight at 37 °C. Plasmids from lysogens grown on selective medium were purified and sequenced.

**Protein Expression and Purification—**The T7 expression plasmid pET21PR-DBD, which produces FLAG-PR-DBD, was constructed by cloning the 276-base pair NheI–EcoRI fragment from pBAD PR mutants into the NheI and EcoRI sites of plasmid pET21b (+) and ligated into pBAD ER-DBD, and purified from transformed E. coli DH5α. 20–24 h after transfection, the cells were subjected to a 3-min shock in 20% glycerol, fed with fresh medium, and protein expression was induced by addition of arabinose, 75 µg/ml Timentin, and 50 µg/ml kanamycin; the cultures were incubated for 4 h at 37 °C, and purified FLAG-ER-DBD was purified using a 1 ml column of protein A-Sepharose (Pharmacia Biotech, Piscataway, NJ). The proteins were then dialyzed against 50 mM KCl, 0.5 mM Tris-HCl (pH 7.9), and 50 µg/ml aprotinin. The protein samples were then analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and Coomassie blue staining.

**Gel Mobility Shift Assays—**Gel mobility shift assays were performed essentially as we have described (21). The reactions were carried out in 20 µl in reaction buffer containing 50 mM KCl, 15 mM Tris-HCl (pH 7.9), 4 mM dithiothreitol, 0.2 mM EDTA, 25 ng of poly[(dI-dC)], and 10% glycerol. Free probe and protein-DNA complexes were quantitated using a PhosphorImager (Molecular Dynamics, Inc., Sunnyvale, CA).

**Promoter Interference Assays—**HepG2 cells were transfected with a total of 8 µg of DNA including 100 ng of the CMV-(ERE)-CAT promoter reporter construct (23). The reporter plasmids were pCMV-FLAG-DBD expression plasmid, 400 ng of pCMV-luciferase (23), and carrier DNA (pTZ18U). 20–24 h after transfection, the cells were subjected to a 3-min shock in 20% glycerol, fed with fresh medium, and analyzed by Coomassie blue staining.

**Plasmid and Phage Constructions—**Plasmid pBAD ER-DBD containing the ER-DBD gene under the control of the arabinose promoter (Pant), was inserted into a double-stranded oligonucleotide containing the consensus ERE (5′-AGGTCTACcagTGACCT-3′) into the Smal site of pFY190, which carries a ~500-bp DNA fragment of phage P22 immI DNA cloned into the EcoRI–HindIII sites of pBR322 (22). Plasmid pFY190 containing the ER-DBD was transformed by electroporation into S. typhimurium ms1883 (15), and the cells infected by P22 mnt+:Kan9 arabinose, 75 µg/ml Timentin, and recombinant phages were selected as a large clear plaque and purified twice on a lawn of MS1862. High titer phage lysates were prepared and purified from MS1883 (15). The presence of the ER-DBD in the P22 phage was confirmed by DNA sequencing.

Plasmid pBAD ER-DBD containing the ER-DBD gene under the control of the arabinose promoter (19) was constructed from pCMV-ER by three successive cycles of polymerase chain reaction amplification. The three cycles were performed by a Shone-Delgado reaction using a unique NheI site at the 5′-end, and an EcoRI site, stop codon, and HindIII site at the 3′-end. The final product was digested with SmaI and HindIII and cloned into pBAD18 (19) digested with NheI and ligated into pBAD ER-DBD with the same enzymes.

**Mutationase—**Saturated random mutagenesis of the DBDs was carried out by a modified mutagenesis protocol. To simulate the process of natural selection in shifting DNA binding specificity from the ERE to the PRE/GR, we needed both a system for producing large numbers of mutants with random amino acid changes and a powerful selection for the relatively rare mutant DBDs exhibiting the desired ERE binding properties. We developed a rapid and simple procedure for saturation mutagenesis of a short region of a protein using degenerate oligonucleotides and Phi DNA polymerase. To select the mutants from this large mutant pool that had gained the ability to recognize the ERE, we adapted the powerful P22 challenge phage (15–17) system for use with a vertebrate protein. In the P22 system, substantial numbers of mutants are screened in a single selection cycle using a life-death selection. In this work, we show that the P22 challenge phage selection system can be used to select for mutants exhibiting a substantial change in DNA binding specificity. The P22 challenge phage selection system provides a new tool for engineering steroid/nuclear receptor DBDs with a desired DNA binding specificity and affinity.

To facilitate identification of amino acids important in discriminating between the ERE and the PRE/GR, it seemed critical to identify the amino acid changes that accompany a shift in DNA binding specificity from the PRE/GR to the ERE. Because the PR-DBD binds to the ERE with a higher specificity than the GR-DBD (18), we employed the PR-DBD in these studies.

We selected and identified mutant PR-DBDs containing amino acid sequences exhibiting high affinity binding to the ERE. We find that the first and third P-box amino acids are the most critical residues for DNA binding specificity, and that mutation of amino acids in the linker region can lead to DBDs whose affinity for the ERE is severalfold higher than that exhibited by the wild-type ER-DBD.

**EXPERIMENTAL PROCEDURES**

**Strains—**Salmonella typhimurium LT2:MS 1582 carrying P22 c2+ mnt+ prophage MS1868, MS1893 (15) and phage P22 mnt+:Kan9 arabinose, 75 µg/ml Timentin, and incubated overnight at 37 °C in order to induce protein expression. FLAG-ER-DBD, FLAG-PR-DBD, and the selected FLAG PR-DBD mutants were purified to near homogeneity by immunoaffinity chromatography using the M2 anti-FLAG monoclonal antibody and elution with FLAG peptide (22).

**Gel Mobility Shift Assays—**Gel mobility shift assays were performed essentially as we have described (21). The reactions were carried out in 20 µl in reaction buffer containing 50 mM KCl, 15 mM Tris-HCl (pH 7.9), 4 mM dithiothreitol, 0.2 mM EDTA, 25 ng of poly[(dI-dC)], and 10% glycerol. Free probe and protein-DNA complexes were quantitated using a PhosphorImager (Molecular Dynamics, Inc., Sunnyvale, CA).

**Promoter Interference Assays—**HepG2 cells were transfected with a total of 8 µg of DNA including 100 ng of the CMV-(ERE)-CAT promoter reporter construct (23). The indicated amounts of CMV-FLAG-DBD expression plasmid, 400 ng of pCMV-luciferase (23), and carrier DNA (pTZ18U) 20–24 h after transfection, the cells were subjected to a 3-min shock in 20% glycerol, fed with fresh medium, and analyzed by Coomassie blue staining.
harvested 40–48 h after the glycerol shock. Cell lysates were prepared and assayed for luciferase activity. CAT activity was determined by our mixed phase assay (24).

Transactivation by DBD-VP16 Chimeras—Transfections were performed as described above with some modifications. HepG2 cells were transfected with a total of 8 μg of DNA including 2 μg of (ERE)-TATA-luciferase reporter plasmid,2 40 ng of pRL-Renilla luciferase internal control plasmid (Promega, Madison, WI), one of the pCMV-DBD-VP16 expression plasmids, and pTZ18U as carrier DNA. Dual luciferase assays were performed according to the manufacturer’s protocol (Promega, Madison, WI).

RESULTS

The ER-DBD Is Highly Toxic in Bacteria—Both wild-type ER-DBD and high affinity ERE binding mutants expressed in E. coli or in the Salmonella typhimurium used in the challenge phage assay are highly toxic to the bacteria. Evidence that the ability of the DBDs to bind to ERE sequences was critical to their toxicity came from our observation that ER-DBD mutants, which had lost the ability to bind to the ERE in vitro, were not toxic in E. coli (data not shown). We concluded that the plasmid loss and cell death that resulted from toxicity of the ERE binding DBDs was due to the presence of nine consen- sus EREs (cEREs) in the E. coli genome (NCBI number U00096). We found that expression from the tightly regulated arabinose promoter (19) minimized toxicity when the inducer was absent. We also replaced the ampicillin in the growth medium with Timentin, a combination of clavulanic acid and ticarcillin, which more effectively blocks growth of bacteria that have lost the expression plasmid and no longer produce β-lactamase.

The Challenge Phage Assay Is a Powerful Assay for Specificity Switch Mutants—Challenge phage are derivatives of bacte riophage P22 that are designed to study protein-nucleic acid interactions in vivo (15). The presence of the imm I region makes bacteriophage P22 especially well suited to genetic selections based on lysis or lysogeny. The imm I region is not present in bacteriophage λ. We use Salmonella, not E. coli, because P22 cannot infect E. coli. Upon infection with a P22 challenge phage, the decision between lysis of the infected Salmonella and lysogeny is controlled by expression of the ant gene, whose product, the antirepressor (Ant), prevents the estab lishment and maintenance of lysogeny. Our challenge phage contain a cERE (aGGTCAcagTGACCt) inserted into the ant promoter at −3 relative to the transcription start site. We selected for mutant PR-DBDs that bound with high affinity to the ERE. If an infected host cell transformed with the PR-DBD mutant pool does not express a mutant PR-DBD that binds to the ERE, the cell is killed by the P22 phage. If an infected cell expresses a mutant PR-DBD that binds with good affinity to the ERE, the cell survives, because binding of the mutant DBD to the ERE blocks ant transcription. In addition, since the challenge phage carry a Kan9 cassette, lysogens can be selected as kanamycin- and ampicillin-resistant colonies (Fig. 1). While this system had found significant application, it had not previously been used with a vertebrate protein.

Since our application of the challenge phage selection required a single step selection of mutant proteins exhibiting a substantial change in DNA sequence specificity, in preliminary studies we tested the effectiveness of the selection system. To determine the background of false positive colonies, 107 S. typhimurium host cells were transformed with different ratios of a control plasmid (pBAD PR-DBD) that produces wild-type PR-DBD, which does not bind to the ERE. The system was spiked with various ratios of the positive plasmid (pBAD ER-

2 G. De Haan, S. Chusacultanachai, and D. J. Shapiro, submitted for publication.
and EGA in the ER-DBD, Gly585 and Val589 were mutated with
frequency (Table II). In the P-box, which is GSV in the PR-DBD
acids and the P-box amino acids were mutated with high fre-
switch mutants showed that all of the nonconserved amino
amino acids (Table II).

100 plates of mutants screened produced lysogens (20–1026
P22-ERE phage. Ten million induced cells were challenged
to create a highly saturated mutant library with a controlled
mutation rate.

Selection of PR-DBD Specificity Switch Mutants and Identifi-
cation of Mutated Amino Acids—In each screen used to iden-
tify the specificity switch mutants ~5,000 independent trans-
formed cells were pooled and plated, DBD expression was
induced with arabinose, and the cells were infected with the
P22-ERE phage. Ten million induced cells were challenged
with the ERE phage and plated on selective medium. 50 out of
100 plates of mutants screened produced lysogens (20–1026
colonies/plate). One lysogen was selected from each positive
plate for further analysis by DNA sequencing. We obtained 37
independent mutants containing an average of five mutated
amino acids (Table II).

Sequence analysis of all 37 genetically selected specificity
switch mutants showed that all of the nonconserved amino
acids and the P-box amino acids were mutated with high fre-
quency (Table II). In the P-box, which is GSV in the PR-DBD
and EGA in the ER-DBD, Gly\textsuperscript{585} and Val\textsuperscript{589} were mutated with
97 and 92% frequency, respectively. Gly\textsuperscript{585} was mutated either
to Trp or to Glu, the amino acid in the ER-DBD. Ser\textsuperscript{586} was
unchanged in 40% of the mutants and was mutated to Gln in the
remaining 60%. Val\textsuperscript{589} was mutated to Ala, Gly, and Ser
with 68, 25, and 7% frequency, respectively. While there is no
single amino acid mutation in the linker region common to
most of the mutants, changes to basic amino acids occurred
with a high frequency, and the mutation G597R was present in
10 of the mutants.

Characterization of the Specificity Switch Mutants—The 37
selected mutant DBDs were subcloned into the FLAG expres-
sion system, expressed as FLAG epitope-tagged proteins, and
purified by immunoaffinity chromatography with anti-FLAG
monoclonal antibody. The affinity of each of the mutant DBDs
for the cERE, for the imperfect pS2 ERE (5′-aGGTCAnnnTG-
GCCCC-3′), and for the PRE/GRE was compared with that of the
ER-DBD and the PR-DBD in protein titrations using quanti-
tative gel mobility shift assays (Fig. 2, A and B). Relative
affinity for the consensus ERE was determined from the con-
centration of protein required to upshift 50% of the probe. In
agreement with our earlier work (21), wild-type ER-DBD
showed little or no detectable binding to the imperfect pS2 ERE
in gel shift assays. In contrast, the genetically selected mutants
displayed high affinity binding to the pS2 ERE (Fig. 2B). While
the starting PR-DBD effectively bound to the PRE/GRE (5′-
AGAACAnnnTCCTGTT-3′) and we carried out only a positive
selection for binding to the cERE (5′-GGTCAnnnTGACCC-3′),
all 37 of the selected mutants completely lost the ability to bind
to the PRE/GRE (Fig. 2C and Table II). This indicates that high
affinity binding to one DNA recognition sequence is incompat-
ible with binding to a different recognition sequence. Mutants
selected using the P22 challenge phage system are therefore
highly specific for binding to the DNA sequences of interest.

In gel shift assays, 17 of the 37 mutants exhibited 10–15-fold
higher affinity for the cERE than wild-type ER-DBD. 14 of the
mutants exhibited 2–9-fold higher affinity binding to the cERE
than was shown by the ER-DBD. Two of the mutants bound to
the cERE with an affinity lower than the ER-DBD, and four
mutants showed no detectable binding to the ERE or to the
PRE/GRE (Table II). Whether these four mutants bind to the
ERE with an affinity below the threshold of detection in our gel
shift assays or are false positives was not examined. Evidence
suggesting that these mutants may bind weakly to the ERE
and are not random false positives comes from the observation
that all four of the nonbinders contained mutations that
changed one of the three amino acids in the P-box of the
PR-DBD to the corresponding amino acid in the ER-DBD P-
box. In contrast, in all 33 of the mutants exhibiting binding to
the cERE in gel shift assays, at least two of the three critical
amino acids in the PR-DBD P-box were mutated.

Mutations in the Linker Amino Acids Enhance Affinity for the
ERE—Surprisingly, high affinity binding to the ERE by the
selected mutants was associated with mutations in the linker
region (amino acids 594–600). In the structures of steroid
hormone receptor DBDs, this region is rather poorly ordered
and forms a flexible linker between the first and second zinc
fingers (5, 6). Consistent with the importance of flexibility in
this region, mutations to Pro were relatively common. The most
striking mutation was G597R, which was present in 10 of the
23 mutants exhibiting >7-fold higher affinity for the ERE than
wild-type ER-DBD. In contrast, none of the six selected mu-
tants whose affinity for the ERE was lower than that of the
wild-type ER-DBD contained the G597R mutation. Since 14 of
the 17 mutants exhibiting a >9-fold increase in binding rela-
tive to the ER-DBD contain at least one linker region mutation
to a basic amino acid, mutations to basic amino acids are
clearly important. While these positively charged residues
probably exhibit electrostatic interactions with the negatively
charged phosphate backbone, they appear to increase affinity
for the DNA without decreasing the specificity of ERE recog-
nition. The importance of ionic interactions is illustrated by
comparing mutants 50 and 56 (Table II), which contain the
same mutations in the P-box amino acids. Mutant 50, with an
affinity for the ERE 12-fold higher than wild-type ER-DBD has
an M595K mutation, while mutant 56 with an affinity for the
ERE 10 times lower than ER-DBD has a Q598D mutation.
Mutations to Asp were rare in the proteins exhibiting high
affinity binding to the cERE and were present in both mutants
exhibiting reduced binding to the cERE.

Mutations to nonpolar amino acids containing aliphatic side
chains were also common in the high affinity binders. Several
amino acids (Cys, Met, Phe, and Trp) present at low abundance
(1–3%) in proteins were also rarely seen in the genetically
selected mutants. Since our mutagenesis was random, their
absence in the selected mutants suggests that their presence
imposes structural or folding constraints on the DBD.

Mutants Exhibiting High Affinity Binding to the Consensus
ERE Also Bind to the Imperfect pS2 ERE with High Affinity—
Imperfect EREs, not the cERE, are found in almost all ERE-
containing genes. Both the full-length ER and the ER-DBD
exhibit reduced affinity for these imperfect EREs (30). We
tested the ability of the mutants to bind to the imperfect ERE,
found in the estrogen-inducible human pS2 gene (31). The
wild-type ER-DBD exhibits extremely weak binding to the pS2
ERE (Ref. 21; Fig. 2C and Table II). The highest affinity mu-
tants bound to the pS2 ERE with >1000-fold higher affinity
than the wild-type ER-DBD (Fig. 2C and Table II). The mutant's ability to bind to the imperfect ERE was particularly striking, since even the highest affinity mutants retained specificity for ERE binding and showed no binding at all to the PRE/GRE.

The Mutants Selected in Bacteria Exhibit High Affinity Interaction with the ERE in Mammalian Cells—Since the mutants were identified by genetic selection in bacteria and assayed for ERE binding in vitro, it was important to evaluate the ability of a few of the mutants to function in mammalian cells. To more directly evaluate the ability of the mutants to bind to the ERE in vivo, we carried out promoter interference assays (23). In these assays, mutants bound to ERs near the initiation site of the CMV promoter compete for binding with basal transcription factors. The amount of transfected expression plasmid required to produce a given level of interference with transcription provides a measure of the interaction of the expressed protein with the ERE. We transfected increasing amounts of DNA encoding mutant DBDs into HepG2, human hepatoma cells, and determined the extent of promoter interference for each DNA. The control PR-DBD did not inhibit transcription. All three tested mutants were clearly more effective in interfering with the activity of the CMV-(cERE)2-CAT promoter than the wild-type ER-DBD (Fig. 3). Mutant 26, with an affinity for the consensus ERE twice that of the wild-type ER-DBD was only slightly more effective than the wild-type ER-DBD. Mutants 5 and 15, with affinities for the ERE 15- and 13-fold higher than wild-type ER-DBD, respectively, required 5–20-fold less transfected DNA to achieve 40% inhibition of promoter activity than the ER-DBD. These data demonstrate that the selected mutants bind to the ERE in intact human cells with far higher affinity than the wild-type ER-DBD.

To analyze the ability of the mutants to activate transcription, we fused the strong VP16 transactivation domain (32) to each of the mutants and to the ER and PR-DBDs, and we expressed the chimeric proteins from the CMV promoter. HepG2 cells were cotransfected with a range of concentrations of each of the chimeric proteins and an ERE-containing luciferase reporter gene. The control PR-DBD-VP16 was unable to activate the reporter gene, while the wild-type ER-DBD-VP16 elicited detectable transactivation only at the highest level of transfected DNA, 25 ng. All three of the mutants exhibited higher levels of transactivation than the WT ER-DBD-VP16. Transactivation by the mutants was related to their affinity for the ERE. Mutant 26 was the least effective, while mutant 5 was slightly more potent than mutant 15 (Fig. 4). Similar results were obtained using a reporter gene containing a single ERE (data not shown).

### Table II

Relative binding affinity of the PR-DBD specificity switch mutants for the consensus ERE, PRE/GRE, and pS2 ERE

| Number | Consensus ERE | Relative binding affinity<sup>a</sup> | pS2 ERE |
|--------|---------------|--------------------------------------|---------|
| 41     | EGCKAFFKRS1QGHND (WT ER) | 100 | 0 | 0 |
| 42     | WEGKASSQAGHND (WT PR) | 100 | 0 | 0 |
| 57     | WEGKASSQAGHND | 1500 | 0 | >1000 |
| 55     | WEGKASSQAGHND | 1500 | 0 | >1000 |
| 15     | EAGSHY | 1500 | 0 | >1000 |
| 7      | WEGKASSQAGHND | 1500 | 0 | >1000 |
| 29     | WEGKASSQAGHND | 1500 | 0 | >1000 |
| 31     | WEGKASSQAGHND | 1500 | 0 | >1000 |
| 50     | WEGKASSQAGHND | 1200 | 0 | >1000 |
| 52     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 25     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 27     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 14     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 43     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 49     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 45     | WEGKASSQAGHND | 1000 | 0 | >1000 |
| 53     | WEGKASSQAGHND | 900 | 0 | >1000 |
| 12     | WEGKASSQAGHND | 900 | 0 | >1000 |
| 48     | WEGKASSQAGHND | 800 | 0 | >50 |
| 59     | WEGKASSQAGHND | 800 | 0 | >50 |
| 47     | WEGKASSQAGHND | 800 | 0 | >50 |
| 22     | WEGKASSQAGHND | 700 | 0 | >50 |
| 13     | WEGKASSQAGHND | 500 | 0 | >50 |
| 51     | WEGKASSQAGHND | 500 | 0 | >10 |
| 9      | WEGKASSQAGHND | 500 | 0 | >10 |
| 37     | WEGKASSQAGHND | 300 | 0 | >10 |
| 54     | WEGKASSQAGHND | 200 | 0 | >10 |
| 17     | WEGKASSQAGHND | 200 | 0 | >10 |
| 26     | WEGKASSQAGHND | 200 | 0 | >10 |
| 40     | WEGKASSQAGHND | 50 | 0 | >10 |
| 56     | WEGKASSQAGHND | 10 | 0 | >10 |
| 16     | WEGKASSQAGHND | 10 | 0 | >10 |
| 18     | WEGKASSQAGHND | 10 | 0 | >10 |
| 19     | WEGKASSQAGHND | 10 | 0 | >10 |

<sup>a</sup> Binding of WT ER-DBD to the consensus ERE and of the WT PR-DBD to the PRE/GRE were set equal to 100. Estimated binding of the WT ER-DBD to the pS2 ERE was set to 1. Binding affinity was calculated from the amount of purified DBD required to shift 50% of the labeled probe in gel mobility shift assays. Relative binding is the ratio of binding by the mutant to binding by WT protein.

<sup>b</sup> The mutated amino acids in the PR-DBD are shown in boldface type.
DISCUSSION

Mutation of Amino Acids in the P-box Is Necessary for Altering HRE Specificity but Is Insufficient for High Affinity Binding to the ERE—Previous studies showed that mutating three amino acids in the P-box (amino acids 585–589) from GSckV found in the PR-DBDs and GR-DBDs to EGckA, which is found in the ER-DBD, was critical to the ability of a DBD to discriminate between the PRE/GRE and the ERE (9, 10). The introduction of amino acid substitutions in the P-box, one amino acid at a time, has also been reported (33–37). We used random mutagenesis to simultaneously mutate the P-box amino acid triplet and the previously unstudied linker region and used a powerful genetic selection to isolate mutants that had gained the ability to bind to the ERE. Our data indicate that there is some flexibility in both the number and nature of the P-box mutations.

Of the 17 selected mutants exhibiting >9-fold higher binding to the ERE than the wild-type ER-DBD, seven retained the Ser at the second position of the P-box (amino acid 586) seen in the PR, and 10 contained the Gly found in the ER. Although the second amino acid in the P-box appears to play a very limited role in discrimination between different HREs, there are rigid requirements for Ser or Gly at this site. All 37 selected mutants contain either Ser or Gly at this position. The first and third amino acids in the P-box (Gly585 and Val589 in the PR) are the most critical residues for HRE recognition. Since none of the selected mutants that bind to the ERE in vitro exhibit changes in only one of these amino acids, we conclude that Gly585 and Val589 must both be mutated for effective ERE binding. While the spectrum of amino acids tolerated at these positions is quite limited, a unique set of amino acids is not required. The amino acids we observed at Val589 of the P-box (Ala, Gly, and Ser) are
The mutants exhibit increased binding to the ERE in human cells. HepG2 cells were transfected with 100 ng of the promoter interference reporter plasmid CMV-ERE-luc and increasing amounts (1, 10, 100, and 500 ng) of the expression plasmid encoding FLAG-DBD5 (○), FLAG-DBD15-VP16 (■), FLAG-DBD26 (▲), FLAG-ER-DBD (▼), and FLAG-PR-DBD (●). CAT activity in the absence of DBD expression plasmid was set equal to 100%, and the percentage of inhibition of CAT activity was determined for each mutant. Each point represents the average of at least two separate transfections.

Fig. 3. The mutants exhibit increased transactivation ability in human cells. HepG2 cells were transfected with 2 µg of 4ERE-TATA-luciferase reporter plasmid and increasing amounts of FLAG-DBD5-VP16 (○), FLAG-DBD15-VP16 (■), FLAG-DBD26-VP16 (▲), FLAG-ER-DBD-VP16 (▼), or FLAG-PR-DBD-VP16 (●). The data for each point represent the average from at least two separate transfections. RLU, relative luciferase units.

Fig. 4. The mutants exhibit enhanced transactivation ability in human cells. HepG2 cells were transfected with 2 µg of 4ERE-TATA-luciferase reporter plasmid and increasing amounts of FLAG-DBD5-VP16 (○), FLAG-DBD15-VP16 (■), FLAG-DBD26-VP16 (▲), FLAG-ER-DBD-VP16 (▼), or FLAG-PR-DBD-VP16 (●). The data for each point represent the average from at least two separate transfections. RLU, relative luciferase units.

Amino Acids Used by Genetically Selected DBDs to Bind EREs

Amino Acids Used by Genetically Selected DBDs to Bind EREs

Amino Acids Used by Genetically Selected DBDs to Bind EREs

Amino Acids Used by Genetically Selected DBDs to Bind EREs

All present in known members of the steroid/nuclear receptor superfamily that recognize the ERE half-site.

In the ER and in most steroid/nuclear receptors that recognize the ERE half-site, Glu is present at the first position in the P-box. G585E was present at this position in about half of the mutants, and G585W, which is not found at this position in any member of the steroid/nuclear receptor superfamily, was present in the other half of the mutants. Of the 17 mutants whose affinity for the ERE was >9-fold higher than that of wild-type ER-DBD, 11 contained G585W, and only six contained G585E.

In a previous study in which this Gly in the GR-DBD was mutated to Trp, this change resulted in promiscuous binding to many response elements including the PRE/GRE and the ERE (33–34). Although there was no genetic selection against PRE/GRE binding in our study, all of the selected mutant proteins containing G585W showed no detectable binding to the PRE/GRE in both protein titration gel mobility shift assays (Fig. 2C and Table II) and in competition gel mobility shift assays performed with a 200-fold excess of unlabeled PRE/GRE (data not shown). This high specificity for ERE binding may result from the presence of multiple amino acid mutations in the recognition helix of our mutants. Mutants with Trp at the first P-box position also bind to the ERE in vivo in promoter interference assays using mutant number 5 (Fig. 3), and effectively activated transcription when linked to the VP16 activation domain (Fig. 4).

While changing the P-box amino acids was essential for altering specificity from binding to the PRE/GRE to the ERE, it was insufficient for high affinity binding to the ERE. Mutant number 26, in which the only changes are to the ERE sequence seen in the ER, exhibited a lower affinity for the ERE than 75% of the selected mutants.

Mutations in the Linker Region of the DBD Result in Strongly Enhanced Binding to the ERE—The linker region between the two zinc fingers in the ER-DBD is identical in almost all species (38). In the crystal structures of steroid receptor DBDs, the linker appears to be flexible, without higher order structure, and is in close proximity to the phosphate backbone of the DNA helix. Mutation of amino acids in this region dramatically increases binding to the ERE. Amino acids with basic side chains, like Lys and Arg, are associated with high affinity mutants, whereas amino acids with acidic side chains are found primarily in the mutants exhibiting reduced or undetectable binding to the ERE (Table II). Despite the prevalence of basic amino acid substitutions in this region, lysine and arginine were not always interchangeable. While G597R was present in 10 of the 22 mutants exhibiting >7-fold higher affinity for the ERE than wild-type ER-DBD, G597K was absent. Mutations to amino acids with aliphatic side chains and to tyrosine with its phenolic hydroxyl group were also commonly found in the high affinity DBDs. These side chains may contact the sugar ring of the DNA backbone (8) and help stabilize the protein-DNA complex.

Imperfect EREs often contain a consensus half-site and a nonconsensus half-site, which differs from the consensus half-site by 1–3 nucleotides. The ER-DBD recognizes these imperfect sequences by low affinity binding using an alternative side chain conformation (39). Most of the selected mutants exhibited a 50–1000-fold higher affinity for the imperfect ERE found in the pS2 gene than the wild-type ER-DBD. We believe this dramatic increase in binding relative to ER-DBD is due to a combination of the higher affinity for the ERE half-site of the selected mutants and to the presence of a robust dimerization interface in the PR-DBD (6). Many of the mutants with a high affinity for the ERE were able to bind to the ERE as a monomer (Fig. 2A) and will effectively occupy the consensus ERE half-site in the pS2 ERE. The formation of a dimerization interface on the DNA facilitates binding of the mutants to the imperfect pS2 half-site. When a mutant dissociates from its low affinity binding site on the imperfect pS2 half site, it remains tethered to the DNA through the strong dimerization interface, and its high local concentration strongly facilitates rebinding to the
imperfect half site. This combination of enhanced affinity for the consensus half-site and dimerization to facilitate rebinding to the low affinity imperfect half-site is probably responsible for efficient binding of the selected mutants to the pS2 ERE. Our observation that several of the mutants have gained the ability to bind to the ERE as monomers and the strong bias of the mutations toward basic amino acids strongly support the view that enhanced binding is a result of direct interaction between the mutated region of the DBD and the DNA. However, it remains possible that some of the mutants exhibit enhanced dimerization. In a study in which the P-box amino acids were mutated, protein-protein interactions appeared to make a major contribution to the ability of T3Rβ-RXα heterodimers to bind to HREs (37).

Production of Specificity-shifted Enhanced Affinity DNA Binding Proteins Using the P22 Challenge Phage System—Production of recombinant proteins targeted to a DNA sequence of interest requires methods for producing large pools of mutants and a powerful selection technique to identify and isolate the mutants of interest.

We found that available mutagenesis methods were unsuitable for saturation mutagenesis of a defined segment of a protein, such as the DNA recognition helix. We therefore developed a simple rapid mutagenesis method using doped oligonucleotides and *Pfu* DNA polymerase. The use of degenerate oligonucleotides allows precise delineation of the amino acids to be mutated and permits retention of amino acids important in protein function. Because doped oligonucleotides are used and the nucleotide ratios can be adjusted, true random mutagenesis is readily obtained. Using *Pfu* DNA polymerase, under the conditions we describe, allows production of large mutant pools without isolation of DNA or ligation, steps that often limit the number of independent sequences in mutant pools.

While the P22 selection system had been used in a number of prokaryotic systems, it had not been applied to a vertebrate protein and had not previously been used to isolate proteins exhibiting far higher affinity for a DNA sequence than the naturally occurring protein that recognizes the site. Instead, most efforts to isolate mutant proteins with defined DNA sequence specificity have focused on the use of selection strategies based on phage display (40–42). Despite its unquestioned utility, the number of false positives generated and the relatively low signal/noise ratio of phage display almost always makes it necessary to perform multiple cycles of selection. In contrast, we show that the P22 challenge phage system can be used to identify one positive cell in a million cells in a single selection cycle (Table 1).

The tailless subfamily of orphan receptors carries Asp at the first position of the P-box (25) and binds to an 5'-AAGTCA-3' half-site that differs from the consensus ERE half-site used in our selections by one nucleotide (5'-AGGCTA-3'). The impressive DNA sequence selectivity of the challenge phage selection system is illustrated by the fact that none of the 37 mutant DBDs we isolated and characterized contained Asp at the first position of the P-box. The high sequence selectivity of the challenge phage system may be related to its use of *in vivo* selection in the presence of the bacterial chromosome. Since the bacterial DNA is present in great excess over the target sequence, it serves as a nonspecific competitor DNA during the *in vivo* selection.

In this work we describe modified conditions for using the bacteriophage P22 challenge phage selection system with a toxic vertebrate protein and demonstrate the feasibility of using this selection system to generate DNA-binding proteins with altered sequence specificity and greatly enhanced affinity for a recognition sequence. This system should prove useful in studying other protein-DNA interactions and for engineering proteins with novel DNA binding specificity. After fusion to activation, repression, or catalytic domains, these engineered DNA binding modules can have a variety of potential regulatory and therapeutic applications.

Acknowledgments—We thank Dr. S. Maloy for helpful advice on the use of the P22 system and Drs. R. Dodson and G. De Haan for helpful comments on the manuscript.

REFERENCES
1. Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Gunther, S., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R. (1995) Cell 83, 835–839
2. Beato, M., Herrlich, P., and Schatz, G. (1995) Cell 83, 851–857
3. Parker, M. G. (1996) Biochem. Soc. Symp. 63, 45–50
4. Freedman, L. P. (1999) Cell 97, 5–8
5. Schwabe, J. W. R., Chapman, L., Finch, T., and Rhodes, D. (1993) Cell 75, 567–578
6. Luisi, B. F., Xu, W. X., Owinsowski, Z., Freedman, L. P., Yamamoto, K. R., Sigler, P. B. (1991) Nature (Lond.) 352, 497–505
7. Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995) Nature 375, 203–211
8. Zhao, Q., Khorasanizadeh, S., Miyoshi, Y., Lazar, M. A., Rastinejad, F. (1998) Mol. Cell 1, 849–861
9. Moller, S., Kumar, V., Verneuil, H. D., Chambon, P. (1989) Nature 343, 271–274
10. Alroy, I., and Freedman, L. P. (1992) Nucleic Acids Res. 20, 1045–1052
11. Umesono, K., and Evans, R. M. (1989) Cell 57, 1139–1146
12. Arbuszke, N. D., and Luisi, B. (1995) Struct. Biol. 2, 341–346
13. Kosztin, D., and Schulten, K. (1997) Biophys. J. 73, 557–570
14. Nardulli, A. M., and Shapiro, D. J. (1992) Mol. Cell. Biol. 12, 2037–2042
15. Maloy, S., and Youderian, P. (1994) Methods Genet. 3, 203–233
16. Lee, E. C., Hales, L. M., Crompt, R. L., and Gardner, J. F. (1992) EMBO J. 11, 305–313
17. Numrych, T. E., and Gardner, J. F. (1995) Semin. Virol. 6, 5–13
18. Petz, L. N., Nardulli, A. M., Kim, J., Horwitz, K. B., Freedman, L. P., and Shapiro, D. J. (1997) J. Steroid Biochem. Mol. Biol. 60, 31–41
19. Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) J. Bacteriol. 177, 4211–4215
20. Parikh, A., and Guengerich, P. F. (1998) BioTechniques 24, 428–431
21. Kunz, M. A., and Shapiro, D. J. (1997) J. Biol. Chem. 272, 27949–27956
22. Chiang, C.-M., and Roeder, R. G. (1993) Peptide Res. 6, 62–64
23. Reeder, J. C., and Katzenellenbogen, B. S. (1992) Mol. Cell. Biol. 12, 4531–4538
24. Nielsen, D. A., Chang, T.-C., and Shapiro, D. J. (1998) Anal. Biochem. 179, 19–23
25. Deters-Wadleigh, S. D., and Fanning, T. G. (1994) Mol. Phylogen. Evol. 3, 209–215
26. Weiler, I. J., Lew, D., and Shapiro, D. J. (1987) Mol. Endocrinol. 1, 355–362
27. Kastner, P., Krust, A., Turcotte, B., Stropp, U., Tora, L., Gronemeyer, H., and Chambon, P. (1990) EMBO J. 9, 1603–1614
28. Gallwitz, R. C., and Joyce, G. F. (1992) PCR Methods Applications 2, 28–33
29. Leung, D. W., Chen, E., and Geedell, D. V. (1989) Technique 1, 11–15
30. Nardulli, A. M., Romine, L. E., Carpo, C., Greene, G., and Rainish, B. (1996) Mol. Endocrinol. 10, 694–704
31. Brown, A. M. C., Jeltsch, J. M., Roberts, M., Chambon, P. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 6334
32. Campbell, M. E. M., Palfreyman, J. W., and Preston, C. M. (1994) J. Biol. Chem. 269, 1–10
33. Zilliacus, J., Wright, A. P. H., Carstede-Duke, J., Nilsson, L., and Gustafsson, J.-A. (1995) Protein 21, 57–67
34. Zilliacus, J., Carstede-Duke, J., Gustafsson, J.-A., and Wright, P. H. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4175–4179
35. Nelson, C. C., Hendy, S. C., and Romanuk, P. J. (1995) J. Biol. Chem. 270, 16891–16897
36. Nelson, C. C., Hendy, S. C., and Romanuk, P. J. (1995) J. Biol. Chem. 270, 16898–16994
37. Nelson, C. C., Hendy, S. C., Farris, J. S., and Romanuk, P. J. (1996) J. Biol. Chem. 271, 19464–19474
38. Todo, T., Adachi, S., and Yamauchi, K. (1996) Mol. Cell. Endocrinol. 119, 37–45
39. Schwabe, J. W. R., Chapman, L., Rhodes, D. (1995) Structure 3, 201–213
40. Rebar, J. E., and Pabo, C. (1994) Science 263, 671–675
41. Jameson, A. C., Kim, S.-H., and Wells, J. A. (1994) Biochemistry 33, 5689–5695
42. Greisman, H. A., and Pabo, C. (1997) Science 275, 657–661