Journal of Cellular Biochemistry 51:140–150 (1993)

On the Mechanism of DNA Binding by Nuclear Hormone Receptors: A Structural and Functional Perspective

Leonard P. Freedman and Ben F. Luisi

Cell Biology and Genetics Program, Sloan-Kettering Institute, New York, New York 10021 (L.P.F.); MRC Virology Unit, Glasgow G11 5JR, United Kingdom (B.F.L.)

Abstract The nuclear hormone receptor DNA-binding domain consists of two zinc finger-like modules whose amino acids are highly conserved among the members of the receptor superfamily. In this review, we describe the various genetic, biochemical, and structural experiments that have been carried out primarily for the DNA-binding domains of the glucocorticoid and estrogen receptors. We describe how the structural and functional information have permitted us to predict properties of the DNA-binding domains of other nuclear receptors. We postulate how receptors discriminate closely related response elements through sequence-specific contacts and distinguish symmetry of target sites through protein-protein interactions. This mechanism explains in part how the receptors regulate diverse sets of genes from a limited repertoire of core response elements. Lastly, we describe the stereochemical basis of nuclear receptor dysfunction in certain clinical disorders. © 1993 Wiley-Liss, Inc.

Key words: steroid receptors, hormone response elements, zinc fingers, protein dimerization, transcription, gene regulation

Nuclear hormone receptors directly transduce signals presented by levels of hormones and other small molecules into effects on gene expression. Upon binding these ligands, the receptors can associate with specific DNA sequences and modulate the transcription of target genes. Unified by functional analogy and a characteristic, punctuated sequence homology, these molecules have been grouped as the "nuclear receptor superfamily," and members have been identified and isolated from species as diverse as mammals and arthropods. Vertebrate receptors have been characterized that specifically bind hormonal forms of vitamins A and D, thyroid hormone, peroxisomal activators, and steroid hormones, such as glucocorticoid, progesterone, estrogen, androgen, and aldosterone [reviewed in Parker, 1991, and references therein]. A number of "orphan receptors" whose ligands and function are unknown have also been discovered. Several putative nuclear receptors have been isolated from *Drosophila* that may play important regulatory roles in embryogenesis, including a receptor for the insect steroid, ecdysone [Seagraves, 1991].

Like many eukaryotic transcription factors, nuclear hormone receptors are composed of domains that correspond to discrete functions. The domains of the receptors encompass the functions of ligand-binding (carboxyl terminal region), DNA-binding (preceeding the ligand-binding domain), nuclear localization (coinciding with DNA- and ligand-binding domains), and transcriptional modulation (localized to more variable regions, including the amino terminus). The amino termini of nuclear receptors vary considerably in length and composition, but the sequences of the DNA and ligand-binding domains have been well conserved [Amero et al., 1992; Laudet et al., 1992]. The sequence homology is greatest in the DNA-binding domain, and indeed, many receptor-encoding genes were cloned by screening cDNA libraries with probes against this region. The ligand- and DNA-binding domains, as well as transcriptional activation regions, all can confer specific function when linked to unrelated, nonreceptor proteins [for example, see Picard et al., 1988; Godowski et al., 1988]. The DNA-binding specificity of the

Received June 22, 1992; accepted June 25, 1992.

Address reprint requests to Dr. Leonard Freedman, Cell Biology and Genetics Program, Sloan-Kettering Institute, 1275 York Avenue, New York, NY 10021.

© 1993 Wiley-Liss, Inc.

receptors has been shown to be encoded entirely by the DNA-binding domain and can be switched by swapping the corresponding segment between different receptors [Green et al., 1988].

The DNA-binding segment of the glucocorticoid receptor (GR) has been shown to be a true domain in that if folds stably and retains function in isolation from the remainder of the receptor [Freedman et al., 1988a]. This has been a great convenience for experimental characterization. Several amino acids in the DNA-binding domain are invariant throughout the family, including eight cysteines that were shown in the case of GR to tetrahedrally coordinate two Zn ions [Freedman et al., 1988a]. Mutagenesis experiments have indicated that specific residues within the Zn module region of GR are critical for DNA binding specificity [Mader et al., 1989; Danielson et al., 1989; Umesono and Evans, 1989], DNA-dependent dimerization [Umesono and Evans, 1989; Dahlman-Wright et al., 1991], and positive control of transcription [Schena et al., 1989]. These results have recently been confirmed and expanded by three-dimensional structural analysis [Hard et al., 1990b; Schwabe et al., 1990; Luisi et al., 1991].

Here, we review and correlate the various genetic and biochemical experiments that have been described primarily for the DNA-binding domains of the glucocorticoid and estrogen receptors with the structural information. We describe how the structure and function have permitted us to predict properties of the DNA-binding domains of the other nuclear receptors. In particular, we postulate how receptors discriminate closely related response elements through sequence specific contacts and distinguish symmetry of target sites through protein-protein interactions. This mechanism explains in part how the receptors regulate diverse sets of genes from a limited repertoire of core response elements. Lastly, we describe the stereochemical basis of nuclear receptor dysfunction in certain clinical disorders.

STRUCTURE OF THE DNA-BINDING DOMAIN

The three-dimensional structures of the DNA-binding domains of the glucocorticoid and estrogen receptors have been established by nuclear resonance spectroscopy [Hard et al., 1990b; Schwabe et al., 1990] and the crystal structure of the glucocorticoid receptor DNA-binding domain (GRdbd) in complex with DNA has been determined [Luisi et al., 1991]. These structures agree well with each other, but there is a difference in secondary structure in a small segment that may be nucleated by DNA binding, as we will describe.

The salient feature of the DNA-binding domain is its Zn coordination sites. In agreement with stoichiometry determinations and spectroscopic evaluation [Freedman et al., 1988a], two Zn ions are each tetrahedrally coordinated by four cysteines to stabilize two peptide loops and cap amino termini of two amphipathic α-helices (Fig. 1). The metal binding sites differ structurally and functionally from those found in the other eukaryotic transcription factors bearing Zn fingers, such as TFIIIA, ADR1, and Xfin, where the metal is coordinated by two histidines and two cysteines [reviewed in Berg, 1990]. It also differs from the Zn coordination site of GAL4 [Marmomstein et al., 1992], where two metal ions share cysteine ligands. We therefore refer to the nuclear receptor loop/helix subdomains as *Zn modules* to distinguish them from the other Zn-bearing structures. Residues coordinating the Zn and supporting the fold of the domain are conserved throughout the family, which strongly suggests that the Zn module structures are also conserved among the receptors.

Although they appear to be structurally similar, the Zn modules of the GRdbd serve different functions. The amino terminal module exposes an α-helix to the major groove and directs contacts with the bases of the target site. The carboxy terminal module forms a dimerization interface, which is mediated principally through contacts made by residues 476–482 (Fig. 1), in a region referred to as the "D box" for its postulated role in dimerization [Umesono and Evans, 1989]. The loops of both modules make phosphate backbone contacts. Although structurally distinct, the Zn fingers, GAL4 Zn center, and the nuclear receptor Zn modules do share the general feature that they all serve to stabilize and orient an α-helix for interaction with the DNA target through major groove contacts [Luisi, 1992]. In the case of the Zn fingers and the GAL4 Zn center, the recognition helix lies within the loop region, while it begins at the carboxy terminal end of the loop in the nuclear receptor modules.

The apparent structural similarity of the two modules gives the protein the appearance of having an approximate structural repeat. Moreover, the modules are encoded by separate exons

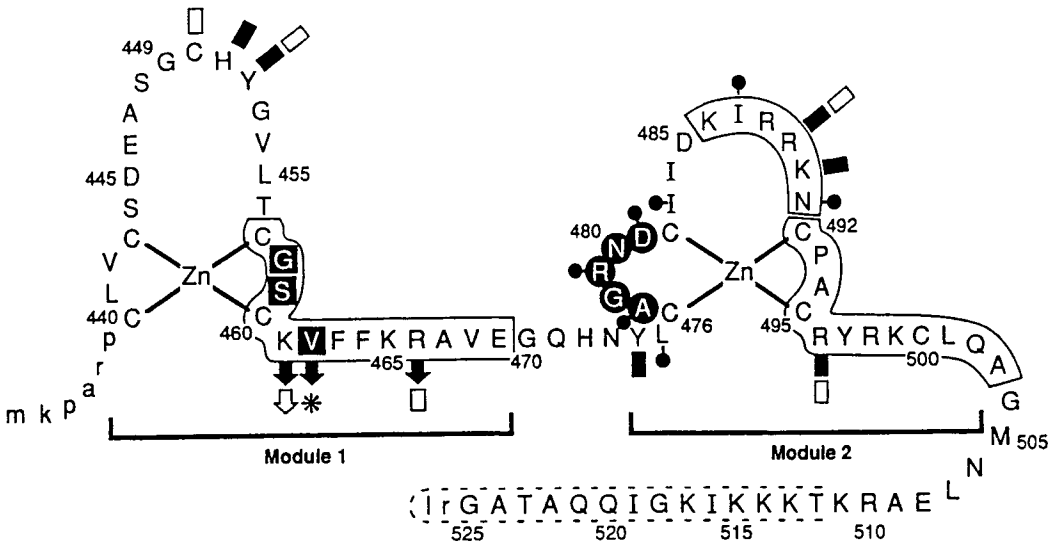

Fig. 1. Zinc coordination scheme of the rat glucocorticoid receptor DNA-binding domain. Numbering scheme is based on the full-length receptor. Indicated residues and regions are based on the crystal structure of the protein bound to a GRE [Luisi et al., 1991]. Residues making specific phosphate backbone contacts are indicated by solid rectangles; those making nonspecific contacts are indicated by open rectangles. Residues making direct contacts with bases are depicted by solid arrows. Amino acids involved in dimer interface interactions are indicated by a solid dot. Three amino acids (residues 458, 459, 462) that confer specificity in mutagenesis experiments (see text) are shown in solid boxes; those demonstrated to confer half-site spacing requirements (residues 477–481) are shown in solid circles. Solid lines enclose α-helical regions. A disordered section at the C-terminus is enclosed by dashed lines. Amino acids as lower-case letters derive from vector sequence of the overexpression plasmid. Reprinted by permission from Luisi et al. [1991].

[Ponglikimongkol et al., 1988], which suggests that they may have arisen by gene duplication and then evolved with different functions. However, the modules are actually topologically nonequivalent, so their relationship may not be so simple. The topology is defined here as the chirality of Zn coordinating residues, which can be *R* or *S*. The amino terminal Zn module, like the Zn fingers, [Berg, 1988] has the *S* configuration. The carboxy terminal module, however, has the mirror image *R* configuration about the metal, and it cannot be changed to that found for the amino terminal module without breaking a Zn-S bond. Despite their structural similarities, the two Zn modules may not have arisen by gene duplication unless one refolded catastrophically about the Zn ion in the very early stages of its evolution.

The two amphipathic helices of GRdbd pack together to form a hydrophobic core which is important for maintaining the globular fold of the domain [Luisi et al., 1991]. The residues in and supporting this core are very strongly conserved in the superfamily. Functionally nonconservative changes here result in loss of function of the glucocorticoid receptor in vivo [Schena et al., 1989] and in the vitamin D and androgen receptors, a substitution here is associated with two clinical disorders (see below).

The NMR structure of the protein in the absence of DNA and the crystal structure of the protein/DNA complex agree quite well. There is a difference, however, in the secondary structure of the carboxyl terminal module. In the crystal structure, there is a short α-helix in the loop region of this module that makes phosphate backbone and dimer interface contacts. This helix is not present in the NMR structures, and it may be nucleated by DNA binding. There is precedent for this effect, for interaction with nucleic acid may nucleate secondary structure in the DNA-binding domains of GAL4 and leucine zipper transcription factors [Marmomstein et al., 1992; Patel et al., 1990]. It is not clear if folding of the short α-helix occurs in the context of the full-length glucocorticoid receptor, which may already exist as a dimer before it binds DNA.

RESPONSE ELEMENTS OF STEROID RECEPTORS

The glucocorticoid response element (GRE) was first identified in DNA-binding studies with the mouse mammary tumor virus long-terminal

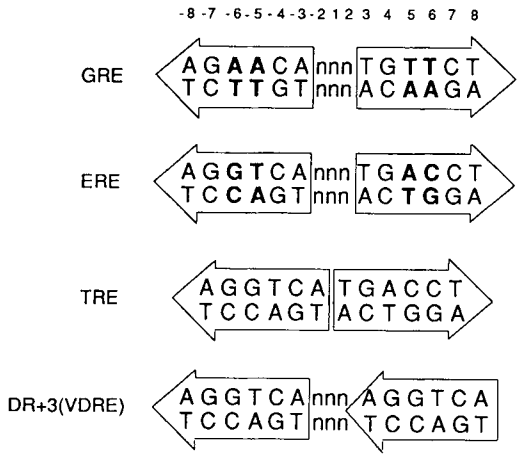

Fig. 2. Idealized hormone response elements are organized as two half-sites. Arrows indicate the directionality of the half sites.

repeat (MMTV LTR) [Payvar et al., 1983]. Glucocorticoid-inducibility was found to be mediated through a 15 base-pair, partially palindromic sequence that consists of two hexameric half-sites separated by three bases. A similar sequence was discovered in the controlling region of the genes for human metallothionein IIA, chicken lysozyme, and rat tyrosine aminotransferase [reviewed in Beato, 1989]. Based on these and other subsequently characterized GREs, a functional consensus was proposed: 5′-GGTACAnnnTGTTCT-3′ [Beato, 1989] where n can be any nucleotide. In nature, the sequences of the half-sites of GREs may vary considerably; nonetheless, the spacing between half-sites is always three bases. A GRE can be constructed that has a perfect palindrome and which will impart glucocorticoid-inducible gene expression in vivo [Strahle et al., 1987]: 5′-AGAACAnnnTGTTCT-3′ (Fig. 2). LaBaer Ph.D. thesis, UCSF [1989] and Hard and collegues [1990a] have shown that the GRdbd association constant for a specific site relative to a nonspecific site differs by only two orders of magnitude ($10^8\,M^{-1}$ vs. $10^6\,M^{-1}$, respectively). This is a much weaker interaction than found for the prokaryotic gene regulatory proteins, and the crystal structure shows that there are fewer protein/DNA contacts. The relatively weak DNA binding is fairly typical of eukaryotic transcription factors.

The hormone response elements of other steroid receptors are also palindromic with a fixed, intervening spacing. These receptors also bind their targets with relatively low affinity. The

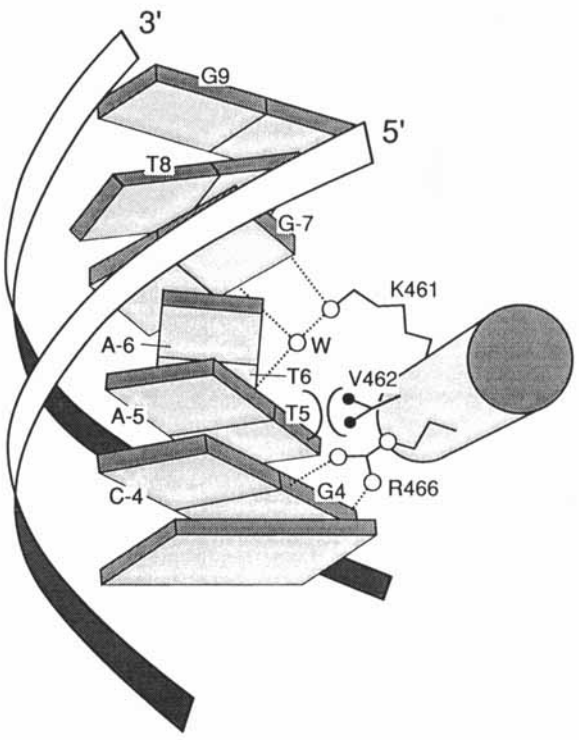

Fig. 3. Direct side-chain contacts by GR amino acids of the first α-helix to bases within the major groove of one GRE half-site. Reprinted by permission from Luisi et al. [1991].

estrogen response element consensus sequence has been identified and has similarity to the GRE: 5′-AGGTCAnnnTGACCT-3′ [Klock et al., 1987].

RECEPTOR DIMERIZATION: GRADING OF TARGETS

The palindromic nature of GREs suggests that the receptor binds its targets as dimers. Furthermore, the center-to-center separation of half-sites in the palindrome is nine bases, which is nearly a helical repeat, suggesting that the two subunits lie on nearly the same face and could contact each other. Indeed, crystallographic analysis [Luisi et al., 1991] and electrophoretic mobility studies [LaBaer Ph.D. thesis, UCSF, 1989; Hard et al., 1990; Dahlman-Wright et al., 1990; Alroy and Freedman, 1992] show this to be the case for the GRdbd. Two monomers of GRdbd bind to the target site and lie on one surface of the DNA, making extensive protein-protein contacts (see Fig. 4a). The subunits expose recognition α-helices into adjacent major grooves. The GRdbd/DNA complex resembles that of prokaryotic proteins having the helix-turn-helix motif. These proteins also bind to palindromic targets

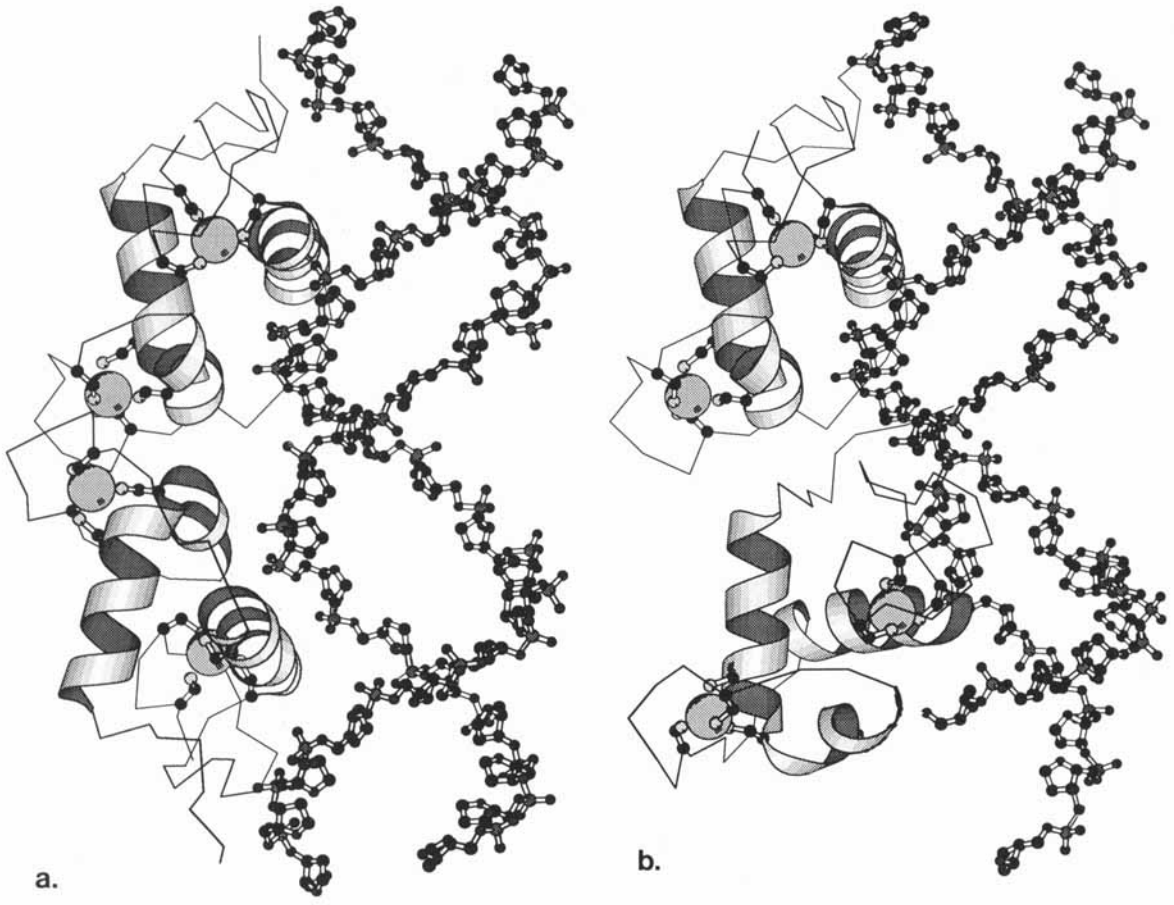

Fig. 4. **a:** Two monomers of the GRdbd bind on one surface of the response element and make an extensive dimer interface. Because the response element is palindromic, the monomers form a symmetrical dimer with a self-complementary interface. The zinc ions are depicted as spheres. Only the phosphate backbone of the DNA is shown for clarity. Helices are represented by ribbons and connecting segments by lines. **b:** Arrangement of DNA-binding domains on an element consisting of a direct repeat of the hexameric recognition core with a separation of three bases (DR+3). In this model, the protein monomers can contact in this head-to-tail arrangement to form an asymmetrical dimer. The figures were prepared with MOLSCRIPT [Kraulis, 1991].

as dimers on one surface of the DNA and interact in the major groove via a recognition α-helix.

The GRdbd is monomeric in solution [Freedman et al., 1988b; Hard et al., 1990], but two molecules bind a GRE cooperatively [LaBaer, Ph.D. thesis, UCSF, 1989; Hard et al., 1990a; Dahlman-Wright et al., 1990; Alroy and Freedman, 1992]. The binding of the first monomer increases the affinity of the second by two orders of magnitude [LaBaer, Ph.D. thesis, UCSF, 1989; Hard et al., 1990a]. Cooperativity results from favorable protein-protein contacts made through an interface which is aligned by DNA binding. If the spacing between half-sites is increased or decreased by a single base, the contacts should be disrupted and, indeed, cooperativity is lost in vitro [Dahlman-Wright et al., 1990; Freedman and Towers, 1991] and transactivation is abolished in vivo [Dahlman-Wright et al., 1990; Nordeen et al., 1990]. As already mentioned, crystallographic analysis shows that the protein-protein contacts are mediated through residues in the carboxy terminal module, including a region referred to as the D box. Freedman and Towers [1991] found that substitution of the D-box residues of the GRdbd with those of the vitamin D_3 receptor (VDR) abolishes cooperativity. The swap substitutes residues that cannot provide all the favorable interactions and would weaken association of subunits. Interestingly, VDR binds nonpalindromic targets, as we will discuss below.

The crystal structure of the GRdbd has been studied in two complexes with DNA [Luisi et al., 1991]. In one, the consensus target was employed; in the other, the spacing between half-sites was increased from three to four bases. In

both cases, the contacts of the dimer interface are the same. One subunit of the dimer is forced out of alignment with the recognition sequence, and it does not make specific base contacts with DNA. We proposed that this interaction represents a non-specific complex [Luisi et al., 1991]. In nature, GREs may not be perfectly palindromic, and deviations from the consensus may be exploited for grading of affinity and activity of the response elements. The structural results show that the receptor could still form a dimer on the surface of the most extreme deviation, but one subunit would be forced to make a non-specific complex. The affinity of such a complex would clearly not be as great as for that with a target conforming to the consensus. Thus, dimerization may serve as a means of modulating target affinities and, possibly, transcriptional responsiveness. The hormone response elements of other steroid receptors are also palindromic and presumably bind to these sites in an analogous fashion to the GR/GRE complex. For instance, the ERdbd is also monomeric in solution and binds its target cooperatively [J. Schwabe, Ph.D. thesis, MRC Cambridge, 1992]. As we describe later, dimerization is also mediated by the ligand-binding domain present in the full-length receptors.

DNA-BINDING SPECIFICITY AND DISCRIMINATION OF RESPONSE ELEMENTS

The glucocorticoid and estrogen receptors discriminate each other's response elements in vivo, and the consensus elements differ at two bases in each half-site: 5′-AG*GT*CAnnnTG*AC*CT-3′ of the ERE vs. 5′-AG*AA*CAnnnTG*TT*CT of the GRE (Fig. 2). Studies using in vivo transactivating assays have identified three amino acids in the DNA-binding domain of each receptor that direct this discrimination [Mader et al., 1989; Danielson et al., 1989; Umesono and Evans, 1989]. These residues are located in the amino terminus of the α-helical region of the first Zn module, in a segment referred to as the "P box" (Fig. 2). Mader et al. [1989] demonstrated that changing three amino acids, *Glu-Gly*-cys-lys-*Ala,* as it occurs in ER, to *Gly-Ser*-cys-lys-*Val,* as it occurs in GR, completely changed specificity so that this mutant ER transactivated strongly from a GRE-driven reporter and not at all from an ERE. Substitution of two of these amino acids in GR to the corresponding ER residues (Gly-Ser → Glu-Gly) partially switches the receptor's specificity from a GRE to an ERE; the third substitution, Val → Ala is required for the full switch [Danielson et al., 1989; Umesono and Evans, 1989]. The binding affinities of these mutants in the context of the DNA-binding domain have been investigated in vitro and correlate with corresponding in vivo affects [Alroy and Freedman, 1992].

The crystal structure of the GRdbd/DNA complexes show that only one residue of the discriminatory triplet (*Gly-Ser*-cys-lys-*Val*) makes a base contact [Luisi et al., 1991]. The Val makes a favorable van der Waals contact with 5-methyl group of the *T* of the GRE half-site TG*T*TCT. The ERE has an A at the corresponding position (Fig. 2). Corroborating the importance of the Val/methyl contact, Alroy and Freedman [1992] found that substitution of Val with Ala, as it occurs in the ER, decreases binding affinity in vitro by a factor of ten. Surprisingly, they found that substitution of the Val with Ala does not increase GRdbd's affinity for an ERE. This residue may not be making an attractive interaction in the ER/ERE complex, but may instead have been selected in evolution for the effective repulsion it presents to a GRE. We propose that only one residue of the ER triplet makes an attractive interaction with the ERE: Glu of the ER recognition helix (of *Glu*-Gly-cys-lys-Ala) accepts a hydrogen bond from the *C* of TGA*C*CT. Indeed, in vitro DNA binding of the GRdbd carrying the Glu at the first position supports the importance of this residue in directing ERE binding specificity [Alroy and Freedman, 1992].

The GR and ER probably share two base contacts in common, and these do not contribute to ERE/GRE discrimination, although they do contribute to recognition of the common features of the hormone response element motif: TGxxCT, where x is a discriminating base. Lys461 and Arg466 in the recognition α-helix of GR contact *G* of A*G*AACT and its complement T*G*TTCT, respectively (Fig. 3). Lys and Arg are also found at the corresponding positions in the recognition helix of ER (and, in fact, in all nuclear receptors reported to date) and probably make corresponding base contacts in the ER/ERE complex: T*G*ACCT and A*G*GTCA (or T*G*xxCT and A*G*xxCA of the general response element).

RESPONSE ELEMENTS FALL INTO CLASSES

Nuclear receptors tend to fall into subgroups which recognize the same response element core, and the amino acid sequence of the P box can be

used to predict core preference. Receptors carrying the GS–V motif at the amino terminus of the recognition α-helix (i.e., glucocorticoid, progesterone, mineralacorticoid, androgen receptors) all recognize a GRE with high affinity. Receptors carrying the EG—A or EG—G motif at the corresponding position (i.e., estrogen, vitamin D_3, thyroid hormone, retinoic acid, ecdysone receptors, and many orphan receptors) all appear to bind to the ERE core (5′-AGGTCA-3′) with high affinity, and probably in all these cases, the Glu of the P boxes accepts a hydrogen bond from one of the two *C*s of the complement (5′-TGA*CC*T-3′) half-site. Within the latter group, specificity may be confered by the spacing and relative orientation of the two half-sites (see below). It is interesting to note that GREs can function virtually as efficiently as response elements for progesterone, androgen, and mineralocorticoid hormones although these hormones elicit distinct physiological effects in organisms. The distinguishing effects of these hormones may arise from tissue-specific expression of the corresponding receptors or the role of auxilary factors in directing specific gene activation.

Wilson et al. [1991] have found that the mammalian orphan receptor NGFI-B recognizes A-T rich sequences flanking the 5′ end of the ERE hexameric core. A peptide segment at the carboxy terminal end of the DNA-binding domain, which Wilson et al. [1992] term the "A box," was shown to direct interactions with these sequences. The crystal structure of GRdbd/DNA complex suggests that the corresponding residues of the orphan receptors could interact with the flanking sequence in the minor groove [Luisi et al., 1991]. This segment does exhibit sequence variation in nuclear receptors, and may impart different DNA-binding properties. For instance, the ERE may be tuned by flanking sequences [Alroy and Freedman, 1992], and it is possible that the A box of this receptor makes analogous minor groove contacts.

DISCRIMINATION OF DNA TARGETS BY PROTEIN-PROTEIN CONTACTS

The subgroup of non-steroid receptors that recognize the ERE core half-site, AGGTCA, may distinguish targets by recognizing the relative orientation and spacing of two such sites. This group includes the thyroid hormone, retinoic acid, and vitamin D_3 receptors (TR, RAR, and VDR, respectively) and certain orphan receptors. In some cases, the response elements are arranged as direct repeats, which suggests that the proteins may bind as asymmetrical dimers (i.e., in a head-to-tail orientation). It is intriguing to note that when the D-box of GRdbd is changed to that of VDR, cooperative binding to the palindromic GRE is abolished [Freedman and Towers, 1991]. This supports the idea that VDR (and the related receptors) may not form symmetrical dimers on the DNA target. Instead, VDR and other receptors may form asymmetrical dimers stabilized by protein-protein interactions, as we shall describe.

TR can activate genes from elements with two ERE half-sites arranged in an inverted repeat with no spacing between half sites (see Fig. 1) [Glass et al., 1988] or as a direct repeat with a spacer of four bases [Umesono et al., 1991; Naar et al., 1991]. For the former element, the center-to-center spacing of the two half-sites is six bases, which is a little more than half a structural repeat of the DNA helix. If two monomers bind to this target, they would lie on opposite faces of the DNA and could not make protein-protein contacts through the DNA-binding domain. However, they could contact through other uncharacterized parts of the receptor, such as the ligand binding domain. It is formally possible that two dimers bind this element, in which a tetramer would cover both faces of the DNA. In the TRE arranged as a direct repeat with a four base spacer, the center-to-center separation of half-sites is 10 bases. Although this distance corresponds to roughly a helical repeat, modeling suggests that the DNA-binding domains would be spaced too far to contact, but would be too close for two pairs of dimers to bind this element side by side.

RAR preferentially induces reporter genes under the control of a target with ERE half-sites arranged as direct repeats with a spacing of five bases [Umesono et al., 1991; Naar et al., 1991]. In this case, the center-to-center spacing would be 11 bases, but two monomers would also be too far to contact through the DNA-binding domain, and two pairs of dimers could not bind as they would clash. Again, it is possible that contacts could be mediated through another part of the full-length receptor.

VDR can bind and activate from ERE half-sites arranged as a direct repeat with a spacing of three base pairs (DR + 3) [Umesono et al., 1991]. Using the GRdbd to model VDR interactions with such a target suggests that two monomers bound at this site would lie on the same surface of the DNA and could make favorable protein-protein interactions. Consistent with

this, the VDR binds the DR + 3 target cooperatively [L.P.F., B.F.L., and T. Towers, in preparation]. Specific asymmetric contacts may be made between residues in the carboxy terminus of the DNA-binding domain with residues in the segment of the amino terminal module of a neighboring receptor (Fig. 4). Wilson et al. [1992] have recently used genetic screens to identify a region in the mouse RXR-β receptor which mediates preference for tandem binding to ERE half-sites. This segment corresponds to the carboxy terminus of the DNA-binding domain, and they have proposed that this segment, which they term the "T box," may contact other regions of tandemly bound receptors. Intriguingly, the modelled VDR contacts correspond in part to residues of the T box.

The preference of the VDR, TR, and RAR to activate genes from artificial controlling elements that are direct repeats of ERE half-sites with a spacing of three, four, and five bases, respectively, has been termed the "3-4-5 rule" [Umesono et al., 1991]. Binding preferences in vivo may not be so simple, because these receptors may form heterodimers with the receptor for 9 cis-retinoic acid, RXR [Yu et al., 1991; Leid et al., 1992; Zhang et al., 1992; Kliewer et al., 1992; Bugge et al., 1992; Marks et al., 1992]. This interaction potentiates receptor DNA-binding and transactivation, presumably by forming a more stable dimer than the homodimeric species. Whether the 3-4-5 rule represents a biologically relevant means to confer specific hormone receptor responsiveness to genuine target genes is not presently clear. However, these observations do suggest that binding site selectivity of this group of receptors may be influenced more by organization of the response elements than by particular nucleotide differences per se, as appears to be true for the discrimination of GRE and EREs. We suggest that spacing and orientation may be influenced by protein-protein contacts in homo- and heterodimers in VDR, TR, RAR, and related receptors. Protein-protein contacts mediating binding of TR and RAR to direct repeats with a spacing of four and five bases, respectively, may be made by regions outside the DNA-binding domains, while critical VDR interfaces appear to take place within and just beyond the DNA binding domain.

DIMERIZATION THROUGH THE LIGAND BINDING DOMAIN

In the absence of DNA, the full-length GR and ER are predominantly homodimers, and bind exclusively as dimers to their targets. Kumar and Chambon [1988] showed that ER dimerization occurs principally through residues within the ligand-binding domain. This dimerization function is stronger than that of the DNA-binding domain. By deletion and site-directed mutagenesis, this region has been sub-localized to ER residues between 500 and 540 [Farwell et al., 1990]. Interestingly, this region, when aligned with other members of the nuclear receptor superfamily, contains a conserved heptad repeat of hydrophobic amino acids that could form a dimerization interface analogous to a coiled-coil interface of leucine zippers [Landschultz et al., 1988]. Conceivably, coiled-coil interactions could take place within this region. Forman et al. [1989] have proposed that this region can direct both homo- and heterodimerization. Indeed, this region appears to mediate heterodimerization between several nuclear receptors and RXR (see above). Recent work also suggests that GR can functionally trimerize with jun-fos and jun-jun complexes, possibly through this domain. This interaction will in turn determine whether transcription of a given gene will be enhanced or represented by the receptor [reviewed in Miner and Yamamoto, 1991]. The ligand-binding domain may also mediate interactions with the heat-shock protein hsp90, which forms a complex with the receptor in the cytoplasm [Pratt et al., 1988]. The regulatory consequences of such baroque interactions of glucocorticoid and, possibly, other nuclear receptors with different classes of transcription factors are clearly profound.

STEREOCHEMISTRY OF NUCLEAR RECEPTOR DYSFUNCTION

A number of mutations have been described in the DNA-binding domain of nuclear receptors which are associated with clinical disorders. Six independent mutations have been mapped to the DNA-binding domains of the vitamin D receptor in patients suffering from hereditary rickets, and several mutations have been reported in the domain of the androgen receptor in patients suffering from testosterone insensitivity syndrome. The crystal structure of the GRdbd/GRE complex provides insight into the sterochemical basis of these disorders.

Arg466, Arg489, and Arg496 of GR (Fig. 1) are conserved at the corresponding positions throughout the superfamily, reflecting the important functional roles they play. These residues are all substituted by Gln in VDR in three

patients with hereditary rickets [Hughes et al., 1989; Sone et al., 1990, 1991], which in all cases could not support the same interactions of the arginines. Arg 466 (residue 47 of VDR) donates two hydrogen bonds to the G of the GRE core hexamer TGTTCT (Fig. 3), and the Arg of the VDR may interact with an equivalent G in VDREs. Gln at this position could not reach the base. Arg489 of GR (residue 70 of VDR) interacts with a DNA phosphate and with Asp445, a residue that is also absolutely conserved at the corresponding position throughout the superfamily. Substitution of the Arg with Gln would allow one but not both of the above interactions to be made. Arg496 (residue 77 of VDR) forms a phosphate backbone contacts that a Gln side chain could not duplicate.

Gly is strongly conserved in the nuclear receptor superfamily at the position corresponding to 449 of GR, and it is substituted by Asp in VDR (position 40) from a patient suffering from the rickets syndrome [Hughes et al., 1989], and as Val in the androgen receptor (position 559) from an individual with partial androgen resistance. The Gly is probably required for two purposes. First, it maintains the peptide backbone in a special confirmation to orient His451 to donate a hydrogen bond to the DNA phosphate and to orient Tyr452 (also Tyr in AR, but Phe in VDR) to pack against the hydrophobic interior of the domain. Second, Gly449 permits the peptide backbone to approach sufficiently close to the phosphate backbone to make a favorable complex. Any other amino acid here would not permit the peptide backbone to follow the same path, due to steric restrictions, and, furthermore, would clash with the phosphate backbone.

A well-conserved hydrophobic core stabilizes the compact, globular fold of the GR DNA-binding domain [Luisi et al., 1991]. This core is formed by the interface between the α-helices of the two modules, and the central elements of this core of GR include Phe463, Phe464 and Tyr 497, which intercalates between the side chains of the two Phe's. In the androgen receptor, Leu 616 occurs at the position corresponding to Tyr 497 and could make analogous hydrophobic contacts with the Phe doublet. Two children suffering from partial androgen resistance carry a Leu-to-Arg substitution at position 616 in the androgen receptor [A. De Bellis and F. French, personal communication]. This substitution may disrupt the hydrophobic core and affect the stability of the DNA-binding domain.

CONCLUSIONS

The crystal structure of the GRdbd/DNA complex has provided stereochemical details of protein-DNA and protein-protein interactions, and the importance of these contacts has been corroborated by mutational analyses. The marriage of structural and functional methods has permitted a better understanding of not only GR DNA binding, but also binding by its highly homologous relatives comprising the nuclear receptor superfamily. The combined methodologies have provided a framework for understanding hereditary disorders associated with defects in the androgen and vitamin D_3 receptors mapping to the DNA-binding domain. The structural and functional data from GR and ER, which allowed us to explain how a few amino acids direct discrimination between closely related GREs and EREs, have been consolidated and extrapolated to describe the DNA binding modes of other nuclear receptors, such as the vitamin D_3, thyroid hormone, and retinoic acid receptors. These latter receptors appear to bind DNA specifically through discrimination of the spacing and orientation of homologous half-sites.

The results described here represent an early stage of an expanding investigation of an intriguing and important superfamily. In our work, only one aspect of nuclear receptor action has been addressed, leaving many questions concerning other functions unanswered. How, for example, are ligands recognized and discriminated? How does ligand binding activate nuclear localization and/or transcriptional activation? How do interactions with other proteins, such as hsp90 and AP-1, affect function, and where is the site of interaction? What converts a nuclear receptor from a transcriptional activator to a repressor? These and other questions are fundamental to a detailed understanding of nuclear receptor action, and will be greatly enhanced by further studies combining structural and functional approaches.

ACKNOWLEDGMENTS

We thank A. De Bellis and F. French for permission to refer to unpublished results on androgen resistance mutations. We also acknowledge P.B. Sigler and K.R. Yamamoto, in whose laboratories much of this work was initiated.

REFERENCES

Alroy I, Freedman LP (1992): DNA binding analysis of glucocorticoid receptor specificity mutants. Nucleic Acids Res 20:1045–1052.

Amero SA, Kretsinger RH, Moncrief ND, Yamamoto KR, Pearson WR (1992): The origin of nuclear receptor proteins: A single precursor distinct from other transcription factors. Mol Endocrinol 6:3–7.

Beato M (1989): Gene regulation by steroid hormones. Cell 56:335–344.

Berg JM (1988): Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc Natl Acad Sci USA 85:99–102.

Berg J (1990): Zinc fingers and other metal-binding domains. J Biol Chem 265:6513–6516.

Bugge TH, Pohl J, Lonnoy O, Stunnenberg H (1992): RXRα, a promiscuous partner of retinoic acid and thyroid hormone receptors. EMBO J 11:1409–1417.

Dahlman-Wright K, Siltata-Roos H, Carstedt-Duke J, Gustafsson J-A (1990): Protein-protein interactions facilitate DNA binding by the glucocorticoid receptor DNA binding domain. J Biol Chem 265:14030–14035.

Dahlman-Wright K, Wright A, Gustafsson J-A, Carstedt-Duke J (1991): Interaction of the glucocorticoid receptor DNA-binding domain with DNA as a dimer is mediated by a short segment of five amino acids. J Biol Chem 266:3107–3112.

Danielson M, Hinck L, Ringold GM (1989): Two amino acids within the knuckle of the first zinc finger specify response element activation by the glucocorticoid receptor. Cell 57:1131–1138.

Farwell SE, Lees JA, White R, Parker MG (1990): Characterization and localization of steroid binding and dimerization activities in the mouse estrogen receptor. Cell 60:953–962.

Forman BM, Yang C-R, Au M, Cassanova J, Ghysdael J, Samuels HH (1989): A domain containing leucine zipper-like motifs mediate novel in vivo interactions between the thyroid hormone and retinoic acid receptors. Mol Endocrinol 3:1610–1626.

Freedman LP, Luisi BF, Korszun ZR, Basavappa R, Sigler PJ, Yamamoto KR (1988a): The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. Nature 334:543–546.

Freedman LP, Towers T (1991): DNA binding properties of the vitamin D3 receptor zinc finger region. Mol Endocrinol 5:1815–1826.

Freedman LP, Yamamoto KR, Luisi BF, Sigler PJ (1988b): More fingers in hand. Cell 54:444.

Glass CK, Holloway JM, Devary OV, Rosenfeld MG (1988): The thyroid hormone receptor binds with opposite transcriptional effects to common sequence motif in thyroid hormone and estrogen response elements. Cell 54:313–323.

Godowski PJ, Picard D, Yamamoto KR (1988): Signal transduction and transcriptional regulation by glucocorticoid receptor-LexA fusion proteins. Science 241:812–816.

Green S, Kumar V, Theulaz I, Whali W, Chambon P (1988): The N-terminal DNA binding zinc finger of the oestrogen and glucocorticoid receptors determines target gene specificity. EMBO J 7:3037–3044.

Hard T, Dahlman K, Carstedt-Duke J, Gustafsson J-A, Rigler R (1990a): Cooperativity and specificity in the interactions between DNA and the glucocortiocid receptor DNA-binding domain. Biochemistry 29:5358–5346.

Hard T, Kellenbach E, Boelens R, Maler BA, Dahlman K, Freedman LP, Carlstedt-Duke J, Yamamoto KR, Gustafsson J-A, Kaptein R (1990b): Solution structure of the glucocorticoid receptor DNA binding domain. Science 249: 157–160.

Hughes MR, Malloy PJ, Kieback DG, Kesterson RA, Pike JW, Feldman D, O'Mally BW (1989): Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. Science 242:1702–1705.

Kliewer SA, Umesono K, Mangelsdorf DJ, Evans RM (1992): Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone, and vitamin D_3 signalling. Nature 355:446–449.

Klock G, Strahle U, Schutz G (1987): Oestrogen and glucocorticoid responsive elements are closely related but distinct. Nature 329:734–736.

Kraulis, PJ (1991): J Appl Crystallogr 24:946–950.

Kumar V, Chambon P (1988): The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. Cell 55:145–156.

Landschulz WH, Johnson PF, McKnight SL (1988): The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science 240:1759–1764.

Laudet V, Hanni C, Coll J, Catzflis F, Stehelin D (1992): Evolution of the nuclear receptor gene superfamily. EMBO J 11:1003–1013.

Leid M, Kastner P, Lyons R, Nakshatri H, Saunders M, Zacharewski T, Chen J-A, Staub A, Garnier J-M, Mader S, Chambon P (1992): Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences. Cell 68:377–395.

Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB (1991): Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. Nature 352:497–505.

Luisi BL (1992): Zinc standard for economy. Nature 356:379–380.

Mader S, Kumar V, deVereneuil H, Chambon P (1989): Three amino acids of the oestrogen receptor are essential to its ability to distinguish an oestrogen from a glucocorticoid-responsive receptor. Nature 338:271–274.

Marks MS, Hallenbeck PL, Nagata T, Segars JH, Appella E, Nikodem VM, Ozato K (1992): H-2RIIBP (RXRβ) heterodimerization provides a mechanism for combinatorial diversity in the regulation of retinoic acid and thyroid hormone responsive genes. EMBO J 11:1419–1435.

Marmomstein R, Carey M, Ptashne M, Harrison SC (1992): DNA recognition by GAL4: Structure of a protein-DNA complex. Nature 356:408–414.

Miner JN, Yamamoto KR (1991): Regulatory cross-talk at composite response elements. Trends Biochem Sci 16:423–426.

Naar AM, Boutin J-M, Lipkin SM, Yu VC, Holloway JM, Glass CK, Rosenfeld MG (1991): The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. Cell 65: 1267–1279.

Nordeen SK, Suh BJ, Kuhnel B, Hutchison CA (1990): Structural determinants of a glucocorticoid receptor recognition element. Mol Endocrinol 4:1866–1873.

Parker MG (1991): "Nuclear Hormone Receptors." London: Academic Press.

Patel L, Abate C, Curran T (1990): Altered protein conformation on DNA binding by Fos and Jun. Nature 347:572–575.

Payvar F, DeFranco D, Firestone GL, Edgar B, Wrange O, Okret S, Gustafsson JA, Yamamoto KR (1983): Sequence-specific binding of the glucocorticoid receptor to MTV-DNA at sites within and upstream of the transcribed region. Cell 35:381–392.

Picard D, Salser SJ, Yamamoto KR (1988): A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. Cell 54:1073–1080.

Ponglikimongkol M, Green S, Chambon P (1988): Genomic organization of the human oestrogen receptor gene. EMBO J 7:3385–3388.

Pratt WB, Jolly DJ, Pratt DV, Hollenberg SM, Gigure V, Cadepond FM, Schweizer-Groyer G, Catelli M-G, Evans RM, Baulieu E-E (1988): A region of in the steroid binding domain determines formation of the non-DNA-binding 9S glucocorticoid receptor complex. J Biol Chem 263:267–273.

Schena M, Freedman LP, Yamamoto KR (1989): Mutations in the glucocorticoid receptor zinc finger region that distinguish interdigitated DNA binding and transcriptional enhancement activities. Genes Dev 3:1590–1601.

Schwabe JWR, Neuhaus D, Rhodes D (1990): Solution structure of the DNA binding domain of the oestrogen receptor. Nature 348:458–461.

Seagraves WA (1991): Something old, some things new: The steroid receptor superfamily in Drosophila. Cell 67:225–228.

Sone T, Kerner SA, Saijo T, Takeda E, Pike JW (1991): Mutations in the DNA binding domain of the vitamin D receptor associated with hereditary resistance of 1, 25-dihydroxyvitamin D3. In Norman AW, Bouillon R, Thomasset M (eds). "Vitamin D: Gene Regulation, Structure-Function Analysis, and Clinical Applications." New York: Walter de Gruyter, pp 84–85.

Sone T, Marx SJ, Liberman UA, Pike JW (1990): A unique point mutation in human vitamin D receptor chromosomal gene confers hereditary resistance of 1, 25-dihydroxyvitamin D3. Mol Endocrinol 4:623–631.

Strahle U, Klock G, Schutz G (1987): A DNA sequence of 15 base pairs is sufficient to mediate both glucocorticoid and progesterone induction of gene expression. Proc Natl Acad Sci USA 84:7871–7875.

Umesono K, Evans RM (1989): Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57: 1139–1146.

Umesono K, Murakami KK, Thompson CC, Evans RM (1991): Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. Cell 65:1255–1266.

Wilson TE, Fahrner TJ, Johnston M, Milbrandt J (1991): Identification of the DNA binding site for NGFI-B by genetic selection in yeast. Science 252:1296–1300.

Wilson TE, Paulsen RE, Padgett KA, Milbrandt J (1992): Participation of non-zinc finger residues in DNA binding by two orphan receptors. Science 256:107–110.

Yu VC, Delsert C, Andersen B, Holloway JM, Devary OV, Naar AM, Kim SY, Boutin J-M, Glass CK, Rosenfeld MG (1991): RXRβ: A coregulator that enhances binding of retinoic acid thyroid hormone, and vitamin D receptors to their cognate response elements. Cell 67:1251–1266.

Zhang X-K, Hoffmann B, Tran PB-V, Graupner G, Pfahl M (1992): Retinoid X receptor is an auxillary protein for thyroid hormone and retinoic acid receptors. Nature 355: 441–446.