Abstract

The functions of the group of proteins known as nuclear receptors will be understood fully only when their working three-dimensional structures are known. These ligand-activated transcription factors belong to the steroid-thyroid-retinoid receptor superfamily, which include the receptors for steroids, thyroid hormone, vitamins A- and D-derived hormones, and certain fatty acids. The majority of family members are homologous proteins for which no ligand has been identified (the orphan receptors). Molecular cloning and structure/function analyses have revealed that the members of the superfamily have a common functional domain structure. This includes a variable N-terminal domain, often important for transactivation of transcription; a well conserved DNA-binding domain, crucial for recognition of specific DNA sequences and protein:protein interactions; and at the C-terminal end, a ligand-binding domain, important for hormone binding, protein:protein interactions, and additional transactivation activity. Although the structure of some independently expressed single domains of a few of these receptors have been solved, no holoreceptor structure or structure of any two domains together is yet available. Thus, the three-dimensional structure of the DNA-binding domains of the glucocorticoid, estrogen, retinoic acid-β, and retinoid X receptors, and of the ligand-binding domains of the thyroid, retinoic acid-γ, retinoid X, estrogen, progesterone, and peroxisome proliferator activated-γ receptors have been solved. The secondary structure of the glucocorticoid receptor N-terminal domain, in particular the tau1 transcription activation region, has also been studied. The structural studies available not only provide a beginning stereochemical knowledge of these receptors, but also a basis for understanding some of the topological details of the interaction of the receptor complexes with coactivators, corepressors, and other components of the transcriptional machinery. In this review, we summarize and discuss the current information on structures of the steroid-thyroid-retinoid receptors. © 1999 Elsevier Science Inc. All rights reserved.

Keywords: Nuclear receptor; Structure; Helix; Ligand-binding domain; DNA-binding domain; Transactivation
some proliferator activated receptors; and the largest category, receptors for which no ligand is known, or orphan receptors. In the amino terminal end of the molecule from the DBD is a widely varying amino acid sequence that in many receptors is important for gene transactivation. Although these proteins are called the nuclear receptor family, some members (glucocorticoid, mineralocorticoid receptors) concentrate in the nucleus only after binding their ligand. The structural information to date comes not just from steroid receptors, but also from other members of the nuclear receptor family. Single-domain structural motifs seem to be shared, and valuable insights are obtained from each; therefore, the data available from all nuclear receptor family members will be discussed here.

Where their function has been demonstrated clearly, it has been shown that all these proteins act by regulating gene transcription. This may occur with or without a ligand bound, depending on the particular receptor/transcription factor and the molecular circumstances. (Of course, the consequences of regulated transcription may be the creation of molecules that post-translationally regulate the expression of other genes. This can be the predominant manner by which some genes are controlled by steroids.) The regulation of transcription by members of the nuclear receptor family occurs by a mixture of mechanisms, involving direct interactions of each receptor with specific DNA sequences [8–10] and/or with a variety of proteins. These include other sequence-specific DNA-binding transcription factors, the basal transcription factors, various coactivators/corepressors, and integrators of transcription [11–13]. A large and growing number of such cofactors have been identified [13–17]. These do not seem to bind DNA directly, but they associate with activated, DNA-bound nuclear receptors and participate strongly in the transcriptional induction and repression that ensues. Very exciting is the suggestion that varying cellular content of these cofactors, together with the influence on cofactor/receptor interactions of particular ligands, may determine the extent and nature of the resulting transcriptional activity. For the first time there is a hint that we shall be able to understand why certain ligands act as antagonists in some cells and agonists in others, or why, in the same cell, various structurally related ligands show agonist or antagonist activity. They may be causing subtle but critical changes in the shape of their receptors, resulting in differing patterns of interactions with coactivators/corepressors/other transcription factors. These are obviously ideas that must be tested and understood at the structural level. Such interactions do not, however, entirely explain the gene-specific action of the nuclear receptors. That specificity depends on mechanisms for attaching the receptor, directly or indirectly, to the correct gene at the correct regulatory region in the DNA.

The classic high specificity DNA-binding sites for the nuclear receptors are called Response Elements (REs). Originally found for the glucocorticoid receptor in the MMTV promoter, these short, patterned DNA sequences consist of an essential sequence of 5–6 nucleotides, a variable length spacer sequence, and then a direct or reversed repeat of the 5–6 essential nucleotides. The various specific sequences and their relationships to one another have a strong determination on the particular nuclear receptor that can bind. This information has been well and often reviewed [2,4,7,8,18]. But how do the various classes of nuclear receptors bind these response elements, with their varying organization, so as to affect gene transcription? Surely the orientation of the receptor towards the transcription initiation site is important. For a glucocorticoid receptor (GR), bound as a head-to-head dimer to its palindromic response element, this may not be a problem because the same GR aspect should present itself regardless of whether the element is 3′ or 5′-wards from the transcription start point. Even so, the relation of the receptor to the transcription factor IID (TFIID) and RNA polymerase complex would differ greatly. GREs can in fact function as enhancers [19], though not as effectively, in general, as those 5′-ward. For a heterodimeric thyroid:retinoid-X-receptor (RXR) bound head-to-tail to a direct repeat element, the problem is more complex because the receptor surface presented to the transcription complex will differ depending on whether the response element lies 3′ or 5′ to the start point, and also depending on the orientation of the direct repeat-towards for away from start site.

Further complicating matters is the fact that nuclear receptors may be interactive with other site-specific transcription factors. In the proliferin gene, the GR has only a hexameric half-site as a potential DNA-binding point and depends on interaction with the transcription factor cJun bound to its nearby site to affect transcription. In addition to the interaction of the receptors with proteins directly involved in transcription, the participation of chromatin structure in gene regulation by nuclear receptors is becoming increasingly evident. In fact, the proteins of the nuclear receptor family, like other transcription factors, probably act on natural genes embedded in chromatin in concert with a complex set of regulatory and structural proteins [20,21]. Certain factors/sites often predominate in the control of specific genes. To understand the complex interactions on genes in chromatin, it is necessary initially to have a clear picture of the structure of each transcription factor and its interaction with its dominant site.

1. Structural features of specific domains of the nuclear receptor family

1.1. The N-terminal domain

Relatively little information is available on the structure of the N-terminal domain of any member of this family. Some studies have explored the secondary structure of this region from the GR; none are available showing its three-dimensional structure. The N-terminal part of the GR con-
tains a powerful transactivation domain, activation function 1 (AF1/taul/enh2), that molecular genetic evidence indicates plays an important role in gene regulation [22,23]. Circular dichroism and nuclear magnetic resonance (NMR) studies have shown that the AF1/taul region, which is rich in acidic amino acid residues, is largely unstructured in water solutions under a variety of pH conditions [24]. However, in the presence of the strong α-helix stabilizing agent trifluoroethanol (TFE), as many as three segments at the C-terminal end of the AF1/tau1 region exhibit α-helical characteristics. The first potential helix region of the AF1/tau1 domain in the human GR shows some similarity to the amino-terminal region of the retinoic acid receptor β2 (RAR β2) [25] responsible for transactivation activity. Mutational studies in the GR tau1 transactivation domain have suggested that the ability to activate a reporter gene in vivo correlates with the ability to form an α-helical conformation in vitro; viz. introduction of helix-breaking proline residues in the potential helix-forming regions of the tau1 greatly disrupts the transactivation capacity, suggesting that α-helical structure is important for the receptor’s activity [24,26]. Other studies have shown that hydrophobic residues of the AF1/tau1 play a critical role during transcriptional activation [24]. There are also reports that transactivation domains of certain other proteins, such as VP16, NF-kB and c-Myc, are largely unstructured in aqueous solution at neutral pH and that they can adopt an α-helical conformation in the presence of TFE [27–29]. Thus, the GR AF1/tau1 transactivation domain probably has a propensity for forming structure under specific conditions. However, it should be noted that although TFE can induce α-helical structure in peptides, which have a propensity for helix formation [30–32], it may also force helicity in peptides, which do not naturally adopt this conformation. Therefore, the importance of the induction of α-helical structure in the AF1/tau1 region by TFE is a topic of considerable interest. It is unclear whether and under what conditions analogous conformational changes occur in vivo.

It is not clear how AF1/tau1, or for that matter any part of the N-terminal region, interacts with other proteins to induce transcription. Hypothetically, the transactivation domain is unstructured until it reaches the putative target factors, then by an induced fit mechanism, interaction of the transactivation domain with its target factor leads to a stabilized structure [33]. It has been reported that the AF1/tau1 domain makes functional, physical interactions with the basal transcriptional machinery, specifically with the Ada and TFIID complexes via the Ada2 and TATA-binding protein subunits, respectively [34,35]. Mutants of the AF1/tau1 core affected each of these interactions similarly. Interactions with other adaptors, coactivators, or mediators are also important [15,36–39]. Thus, data short of actual structural proof supports the idea that conditional folding of the activation domain is an important requirement for its interaction with target factors and subsequent roles in gene regulation [29,34,36,38,39].

1.2. The DNA-binding domain

The DBD has the most conserved amino acid sequence among the members of the steroid-thyroid-retinoid receptor superfamily, and consists of two zinc-finger motifs, each containing four highly conserved cysteine molecules coordinating binding of a zinc atom. This results in the formation of a tertiary structure containing helices that interact specifically with DNA sequences that are organized appropriately in response elements [40–42]. The peptide loops that can be diagrammed around each zinc finger are superficially similar to those of the zinc fingers described for the TFIIIA from Xenopus [43,44] and are inferred to exist in a wide variety of nucleic acid binding proteins [45,46]. The sequence similarity of the human GR DBD to the TFIIIA was recognized at once, and consequently each zinc with its associated amino acids is still referred to as a zinc finger and diagrammed as such in two dimensions (as here, in Fig. 1). In the nuclear receptors, however, they do not form classic three-dimensional zinc finger structures as in the Xenopus protein and other true zinc finger proteins [47]. The two zinc-fingers of TFIIIA act as independent, conformationally stable structural units, in the presence or absence of DNA, and each contributing independently to DNA binding. The zinc-fingers of the nuclear receptors fold together as part of a larger, unified globular domain [48–50], the parts of which do not bind DNA equivalently and whose structure may be somewhat modified upon DNA binding.

The DBDs of several nuclear hormone receptors have been expressed as recombinant peptides and well characterized structurally. With the use of NMR spectroscopy, the structures in solution of the DBDs of the estrogen, glucocorticoid, retinoic acid-β, and retinoid-X receptors have been determined [48,51–53]. The crystal structures of the DBDs of the glucocorticoid and estrogen receptors have been solved in complex with DNA [40,54]. The central feature of the secondary structure elements within the GR DBD is found in three helical regions. Helices I and III are oriented
the major groove [56]. The dimer interface, a loop of five amino acids in the protein helix (Fig. 1) and certain bases in the DNA helix provides critical contacts between three structures show that helix I fits into the major groove of and showed more symmetry of DNA interactions. The crys-
two complexes, which gave fundamentally the same results global structural similarities and differences between the contacts. However, the detail was sufficient to examine the was inadequate to satisfactorily resolve all side chain-base between hexamers of the GRE, although the resolution seen also obtained with the natural three base-pair separation other forms a looser, non-specific interaction. Crystals were -helices, whereas helix II is somewhat distorted [48] (Fig. 2). The DBD may be considered as two interdependent subdomains, each consisting of a zinc ion, tetrahedrally coordinated to the sulphurs of four cysteine residues, and an amphipathic helix. The two subdomains differ both structurally and functionally. The helix of the first subdomain is mainly involved in site-specific recognition based on its interaction with certain bases in the cognate response element hexamer. Also within this helix are the amino acids that are responsible for site-specific discrimination of binding. These 3–4 amino acids have been termed the P box [40,53]. A loop formed in the second subdomain provides the DBD homodimerization interface and the helical region, less specific DNA interactions.

The best crystal structure of the GR DBD in complex with DNA has been provided using a GRE sequence modified to have a four base-pair spacer sequence between the two GRE half sites, instead of the proper three [40]. The structure shows a receptor DBD dimer on the palindromic GRE, which is consistent with the known cooperative binding kinetics of the DBD to the GRE. As a consequence of the unnatural four base-pair spacer, however, the two DBDs in the dimer do not bind the GRE equivalently. One half of the dimer appears to bind specifically to the DNA, while the other forms a looser, non-specific interaction. Crystals were also obtained with the natural three base-pair separation between hexamers of the GRE, although the resolution seen was inadequate to satisfactorily resolve all side chain-base contacts. However, the detail was sufficient to examine the global structural similarities and differences between the two complexes, which gave fundamentally the same results and showed more symmetry of DNA interactions. The crystal structures show that helix I fits into the major groove of the DNA helix and provides critical contacts between three amino acids in the protein helix (Fig. 1) and certain bases in the major groove [40]. The dimer interface, a loop of five amino acids that is also called the D box (Fig. 1), lies between the first two cysteines of the second zinc finger.

The agreement between the structures of the DBD:GRE crystals and the DBD alone in solution is generally quite good [40,49]. The DBD in solution, determined by NMR, is a monomer, and there has been some discussion regarding the conformation and flexibility of parts of the second zinc-finger, which is well resolved in the crystal, but less well in solution. This region is also poorly defined in the solution structure of the estrogen receptor (ER) DBD [51]. Perhaps the conformation of this fragment is stabilized upon formation of the dimeric DBD-DNA complex [40,48]. We await data that tests this hypothesis. NMR relaxation measurements and molecular dynamics simulations of the GR DBD revealed a uniform and limited mobility along the backbone [55,56]. Concerted motions in and between the subdomains could facilitate structural rearrangements that contribute to the cooperativity of DNA binding [55,56].

The two subdomains in the DBD suggest a direct structural repeat of a zinc-coordinating center and amphipathic helix. By assigning priorities to the members of zinc-coordinating residues, however, the metal binding sites can be defined as having a chirality [40]. With this scheme, the module around the first zinc has the S configuration, whereas the second has R configuration. Consequently, the zinc-binding sites are topologically distinct and can not be superimposed. The hydrophobic and hydrophilic surfaces of the helices are also reversed in the two modules [40].

The crystal structure of the ER DBD:ERE (with native three-base pair spacing) complex has been determined at 2.4Å resolution compared to the glucocorticoid receptor DBD:GRE complex at 2.9Å, and hence is able to give a better picture of the stereochemistry of protein:protein and protein:DNA interfaces. The structure of the ER DBD:ERE complex is very similar to that of the GR DBD:GRE. The ER DBD structure also consists of two amphipathic α-helices packed at right angles. An extensive hydrophobic core is formed between these two helices. As a result of folding of the residues at the N-terminal to the helix I, two loops are formed. The hydrophobic residues present in these loops pack with the hydrophobic core of two helices mentioned above. In the structure of the ER DBD, there is an intricate network of hydrogen bonds formed among ordered water molecules, amino acid side chains, and nucleotide bases. This contrasts with the apparent scarcity of ordered water molecules in the GR DBD structure, but this difference may be attributed to the higher resolution of ER DBD structure.

The conserved spacing of three bases between the ERE and GRE hexameric half-sites separates their centers by nine base pairs, which is nearly one turn of the DNA helix. Consequently, the two DBD monomers lie on the same surface of the DNA duplex. The relative orientation of the two DBDs in the homodimeric complex is determined not only by the RE sites, but also by monomer interactions critical for recognition of spacing and orientation of the hexameric half-sites [40]. The (DBD)2:DNA complex of the
glucocorticoid and estrogen receptors shows that each DBD exposes an α-helix to the bases in the major groove of the DNA, and the recognition surfaces of these complexes are supported in the major groove. Each monomer contacts the sugar phosphate backbone on either side of the major groove it faces.

The three-dimensional structure of the recombinant human RAR-β DBD has been determined by NMR, and it shows an overall structure similar to that observed in the DBD of the rat GR, having two α-helices crossing each other at a perpendicular. The two zinc atoms in the first and second fingers have S and R chirality, respectively, which is similar to that seen in the structure of the GR. The C-terminal α-helix is two residues shorter than that of the GR DBD and is followed by a well defined region of extended backbone structure [52].

The three-dimensional structure in solution of another member of the superfamily, the RXR-α, as determined by NMR, also reveals similar folding of the two zinc fingers to form two perpendicularly oriented helices [53]. However, the most interesting finding in the structure of the RXR DBD is an additional helix immediately after the second zinc finger. This additional helix presumably facilitates both protein:DNA and protein:protein interactions required for high affinity binding of the RXR DBD to its cognate RE. The additional helix also provides insight into the structural features that underlie receptor binding to direct repeat HREs. The amino acids in this third helix are conserved in the isoforms of RXR found in different species; so the third helix may be a general feature. There are suggestions that the third helix in the RXR DBD functions not only in RXR homodimerization, but may also function in the well known interactions of RXR with other members of the nuclear hormone receptor superfamily. From modeling binding of the RXR DBD to a direct repeat, it is speculated that RXR binds each response element half-site individually in a manner similar to that of the GR, with helix I lying across the major groove. In this orientation, the third helix projects toward the minor groove of DNA. The binding of two DBDs of RXR to a direct repeat thus brings the third helix of one subunit into close proximity to the second zinc finger of the second monomer, suggesting that homodimerization on the direct repeat sequence may involve a novel dimer interface. This would involve the amino acid residues in the second metal binding loop in one subunit and the third helix of the other. This is different from the glucocorticoid receptor DBD, which dimerizes on its consensus inverted repeat target sequence through symmetrical contacts that involve residues in the D-loop of the second zinc finger of each DBD [40]. Perhaps the differences in DBD structure between RXR and GR/ER, therefore, explain how they bind to response elements differing in organization, i.e. direct repeat versus palindrome. The crystals of the heterodimer formed by RXR and the thyroid receptor DBDs on a TRE composed of a direct tandem of half-sites separated by four base pairs (TREdr4) show that the bound DBDs engage the major grooves of successive repeats with the same geometry, placing them on the same face of the TREdr4 target [57].

The GR coding sequence from *Onchorhynchus mykiss*, the rainbow trout, have been obtained. The deduced sequence shows the typical two zinc fingers, but differs from the available mammalian GR sequences by having an additional nine amino acids between the two zinc fingers [58]. The fish GR functions to stimulate transcription from a consensus mammalian GRE [58], but how it folds to accommodate to the binding site remains to be seen. However, comparison between the sequences of the rainbow trout ER and human ER shows that the homology in the DNA and hormone-binding regions is highly conserved, although it is not conserved in a homologous manner all along the receptor [59].

### 1.3. The ligand-binding domain

The LBD not only serves the function of providing a class-specific ligand binding site, but also as a homodimerization domain. It also interacts with other proteins in transactivating genes and still others in regulating receptor activation. A common overall layout has been identified in the several LBD structures determined. The three-dimensional structure of the recombinant LBDs of the RAR-γ, TR, and ER with their ligands bound have been solved by x-ray diffraction, as has that of the unliganded RXR-α [60–63]. All are composed of a series of 11–12 α-helices (H1–H12) closely folded in a similar manner. Because no universal numbering convention has been agreed upon, there are slight differences in the specific helix number given for topologically analogous sites when LBDs from different receptors are compared. Still, inspection of LBDs and their helices, regardless of the numbering system, shows conservation of helical arrangement, with only slight differences. Although the data so far do not do so completely, eventually the structure of each LBD must explain its multiple functions.

The unliganded LBD of the RXR-α was the first of these structures solved. The data revealed a structure in which about 65% of the total volume is occupied by 11 α-helices organized in a three layer, antiparallel sandwich [61]. When the three-dimensional structures of the ligand-bound LBDs of TR, RAR-γ, and ER were determined, they showed similar overall folding [60,62,63]. Indeed, the secondary structural features are superimposable, except for the C-terminal helix [64]. This structural information is summarized in Table 1.

Although the secondary structure elements in the unliganded RXR-α and RAR-γ with ligand bound are similar, the latter is more compact overall. The ligand-bound LBDs of the TR and ER also reveal a more compact structure compared to unliganded RXR-α. Presumably, these differences are due to conformational changes following ligand binding, so that ligand binding may function to stabilize the
The way the LBD participates in natural, monomer-dimer relationships of these receptors is not spelled out by the data so far. The unliganded RXR crystallized as a dimer, with the dimer interface mainly formed by helices H10 and, to a lesser extent, H9 and the loop between H7 and H8. Each monomer of the RXR LBD contributes 11% of its solvent accessible surface to the dimerization interface [61]. This is typical of specific protein:protein interactions. However, such a dimer interface is not seen in the ligand-bound structures of the RAR-γ and TR [62,63], whereas the ligand-bound ER LBD formed homodimers that were reminiscent of the unliganded RXR homodimer. The dimer contacts between the ER LBD monomers are made primarily through H11, which intertwine to form a rigid backbone, but also involve H8 from one monomer and parts of H9 and H10 from other [60].

The C-terminal part of the LBDs of the RAR, the TR, and the ER has been shown to have a ligand-inducible activation function, termed AF2 [66,67]. For this function to manifest itself, ligand-dependent interaction with a number of putative transcriptional coactivators must occur [68]. In the unliganded RXR-α LBD structure, the region shown by molecular genetics to be essential for AF2 function adopts a helical structure, which corresponds to the C-terminal helix H11. This helix is often known as the AF2 activation helix (H12 in Fig. 3). It is seen to extend from the core of LBD, pointing away from the dimer axis at an angle of about 45°. It was reasoned that this could be optimal for facilitating interactions with putative transcriptional intermediary factors [69]. In the liganded RAR, TR, and ER LBDs, the equivalent AF2 activation helix corresponds to the C-terminal amphipathic helix H12, and it is positioned as a tight part of the globular LBD that is different in position from its counterpart in the unliganded RXR [60–63]. Both deletion and mutation studies have shown that AF2 activation helix is essential for ligand induced transcriptional activation. It appears that the activation of AF2 upon ligand binding corresponds to major conformational changes, which create the proper surface required for efficient interaction with transcriptional factors, the putative mediators of the AF2 function [70,71].

During the meeting on Nuclear Receptor Gene Family held at Incline Village, Nevada, March 28–April 3, 1998, the crystal structure of the LBD of hormone bound progesterone receptor was discussed [72]. The overall structure

---

### Table 1

Structural comparisons from the crystal structures of the LBDs of the nuclear hormone receptors

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Ligand</th>
<th>No. of α-helices</th>
<th>No. of β-sheets</th>
<th>Layered arrangement*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RXRγ61</td>
<td>—</td>
<td>11</td>
<td>2</td>
<td>1,2,3/4,5,8,9,N-term of 11/6,7,10</td>
</tr>
<tr>
<td>RAR-γ62</td>
<td>All-trans retinoic acid</td>
<td>11</td>
<td>2</td>
<td>1,2,3/4,5,8,9,10,11</td>
</tr>
<tr>
<td>TR63</td>
<td>dimit, T3 and IpBr2</td>
<td>12</td>
<td>4</td>
<td>1,2,3,4/5-6,9,10/7,8,11,12</td>
</tr>
<tr>
<td>ER60</td>
<td>E2, RAL</td>
<td>12</td>
<td>2</td>
<td>1,2,3,4/5-6,9,10/7,8,11</td>
</tr>
</tbody>
</table>

* T3 = 3,5,3'-triiodothyronine; IpBr2 = 3,5-dibromo-3'-isopropylthyronine; dimit = 3,5-dimethyl-3'-isopropylthyronine; E2 = 17β-estradiol; RAL = raloxifene; H = helix as assigned by authors. The arrangement of helices in all receptors is very similar, despite the differences suggested by the variations in numbering, e.g. helices 4–10 in the TR-T3 correspond to helices 3–9 in the RAR-RA and helix 11 in the TR-T3 corresponds to helices 10–11 in the RAR-RA.
reveals that both the pattern of folding and the corresponding regions involved in ligand contact are similar to homologous nuclear receptor LBDs. The hormone-bound LBD is dimeric in the crystalline state. During the meeting, the crystal structure of human peroxisome proliferator activated receptor delta LBD was also presented. The overall structure reveals a large interior pocket framed by a common fold found in nuclear receptors [73]. The amphipathic C-terminal activation helix, even in the absence of ligand, constitutes an active conformation, which has been seen in the structures of the ligand-bound RAR-α and TR. That may explain ligand-independent activation of this receptor. (The progesterone and peroxisome proliferator activated receptor-γ LBD structures also have been published very recently [74–76].)

In sum, the existing evidence shows that the structure of the LBD of nuclear receptors consists of a conserved arrangement of helices, sandwiching and participating in a ligand-binding pocket (Fig. 3). How the binding of ligand affects the structure is not certain, but it probably causes intramolecular shifts that result in repositioning of the C-terminal helix essential for AF2, so that LBD interaction with cofactor proteins is changed. Agonists would provoke a structural shift favoring heterodimerization with cofactors, (some) antagonists would not.

2. Summary and conclusion

Studies of recombinant proteins corresponding to individual domains of members of the steroid hormone receptor family give an initial picture of the overall structure of these proteins. They appear to share basic features. At the amino terminal end there is a domain of varying length and sequence with relatively little structure. In many cases, this region contains a powerful activation function. The relatively limited data on this region suggests that it is rather unstructured. Possibly, specific structure in this domain is achieved by ligand and/or DNA binding elsewhere in the receptor, or even more likely, by direct interactions with other proteins involved in the transcription transactivation function. Recent data suggests that the binding of holoestrogen receptor to differing EREs results in differing global ER structure [77]. The DNA-binding domain lies to the C-terminal side of the above. The DBD has an overall globular structure, subdivided into two subdomains. The more amino-terminal subdomain contains a zinc-stabilized helix that is responsible for identification and initial binding to specific response element DNA sequences. The more C-terminal subdomain binds DNA in a less specific way and contributes to the affinity of binding. Portions of the second subdomain are also involved in receptor homodimerization. The DBD also is involved in heterodimerization with certain non-receptor transcription factors. Nuclear localization signals, whose structure has not been defined, follow as one moves C-terminal through the receptor molecule. These are succeeded by the large, globular ligand-binding domain comprised of 11–12 helices in a structure that supplies the ligand-specific binding site. A strong transcription transactivation function, provided in the ligand-binding domain, requires the correct positioning of the C-terminal helix of the domain. This positioning is determined by the ligand binding. The LBD also is a strong contributor to homodimerization. In some steroid receptors at least, when no ligand has activated by binding, the LBD binds a set of heat shock and other proteins in a pre-activation complex.

This overall view of structure now opens the way to ask more advanced questions. What is the structure of the holoreceptor with and without DNA binding? What are the effects of each domain on the others in structure and function? Mutational analysis of the ER, coupled with studies of transcriptional activities, suggest that the LBD affects transactivation functions in the amino terminal region [78]. How does this receptor interact structurally with its many possible binding partners, including co-activators, corepressors, other transcription factors, heat shock proteins, and other proteins? The ultimate question is: How does this complex of proteins interact with chromatin and the machinery of transcription in a mega complex to regulate the expression of genes?

2.1. Structural information needed to solve receptor complexity

Although it is valuable to understand the structures of the different functional domains separately, eventually one wants to know how they are organized in the holoreceptor.
There is speculation that the structure of specific domains may not be the same in the full length receptor under in vivo conditions as that seen in studies of isolated individual domains. What structural changes take place when a hormone binds to its respective receptor? Data indicates that ligand interacts with its cognate receptor to induce conformational changes in the tertiary protein structure, enabling productive interactions between the hormone-bound receptor and its target response element DNA sequences [79,80]. Many interpret recent results on ligand-specific receptor/ cofactor interactions to mean that specific steroids subtly alter receptor shape, determining what surfaces present themselves for cofactor binding. We speculate that binding to a specific HRE also leads to some conformational changes, particularly in the N-terminal domain, which in turn may lead to the receptor entering transcriptionally active form. The interaction with cofactors also may bring some structural changes in the receptor. How do the holoreceptors interact with their specific HREs and does binding to HRE cause conformational changes in the receptor structure? Are there some factors that impose structure in the N-terminal domain under in vivo conditions?

Acknowledgement

We are thankful to Feng Zhou, Ph.D. for his help in creating the computer graphics figures.

References


