Allosteric Modulators of Steroid Hormone Receptors: Structural Dynamics and Gene Regulation

Raj Kumar and Iain J. McEwan
Department of Basic Sciences (R.K.), The Commonwealth Medical College, Scranton, Pennsylvania 18510; and School of Medical Sciences (I.J.M.), University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, Scotland, United Kingdom

Steroid hormones are synthesized from cholesterol primarily in the adrenal gland and the gonads and play vital roles in normal physiology, the control of development, differentiation, metabolic homeostasis, and reproduction. The actions of these small lipophilic molecules are mediated by intracellular receptor proteins. It is just over 25 yr since the first cDNA for steroid receptors were cloned, a development that led to the birth of a superfamily of ligand-activated transcription factors: the nuclear receptors. The receptor proteins share structurally and functionally related ligand binding and DNA-binding domains but possess distinct N-terminal domains and hinge regions that are intrinsically disordered. Since the original cloning experiments, considerable progress has been made in our understanding of the structure, mechanisms of action, and biology of this important class of ligand-activated transcription factors. In recent years, there has been interest in the structural plasticity and function of the N-terminal domain of steroid hormone receptors and in the allosteric regulation of protein folding and function in response to hormone, DNA response element architecture, and coregulatory protein binding partners. The N-terminal domain can exist as an ensemble of conformers, having more or less structure, which prime this region of the receptor to rapidly respond to changes in the intracellular environment through hormone binding and posttranslation modifications. In this review, we address the question of receptor structure and function dynamics with particular emphasis on the structurally flexible N-terminal domain, intra- and interdomain communications, and the allosteric regulation of receptor action.

I. Introduction
Steroid hormone receptors (SHRs) are ligand-regulated transcription factors that control a diverse array of physiological processes. The importance of SHRs in the regulation of cellular and developmental processes is well established (1–6), yet the mechanism by which transcription of target genes is controlled by steroid hormones and SHRs is not completely understood. In recent years, studies from many laboratories have revealed details of the mechanisms by which SHRs modulate transcription regulation and activate cell signaling pathways (for recent reviews, see Refs. 7–11). This has resulted in the identification of molecules with diverse activities that regulate receptor function and the elucidation of the role of SHRs in both normal physiological processes and pathological conditions, including cancer, car-

Abbreviations: AF, Activation function; AR, androgen receptor; CTE, C-terminal extension; DBD, DNA binding domain; ER, estrogen receptor; ERE, ER element; ERR, estrogen-related receptor; GR, glucocorticoid receptor; GRE, glucocorticoid response element; HDX, hydrogen deuterium exchange; HRE, hormone response element; ID, intrinsically disordered; LBD, ligand binding domain; MR, mineralocorticoid receptor; NLS, nuclear localization signal; NMR, nuclear magnetic resonance; NTD, N-terminal domain; PPAR, peroxisome proliferator-activated receptor; PR, progesterone receptor; RXR, retinoid X receptor; SRC-1, steroid receptor coactivator-1; SRM, SHR modulator; TBP, TATA box binding protein; TFE, trifluoroethanol; TFII, transcription factor II, VDR, vitamin D3 receptor.
diovascular disease, and several other disorders (12–19). However, despite remarkable progress, the mechanisms by which SHRs function in a tissue-specific manner remains unclear. To more fully understand how SHRs function as transcription factors, it is therefore necessary to know how these receptor proteins are activated by hormone and bind to specific DNA sequences in the regulatory regions of target genes to either up- or down-regulate the rate of transcriptional initiation by RNA Pol II. Identification of the specific factors and molecular mechanisms that contribute to the normal actions and aberrant pathological effects of SHRs is critical for the development of novel and potent therapeutic agents. It is worth noting that steroid hormones also mediate rapid actions through both classical SHR and membrane-associated receptors: these mechanisms have been discussed in a number of excellent reviews (20–23) and will not be specifically discussed here.

SHRs are part of a large superfamily of nuclear receptors that share structural and functional properties (24, 25). The cloning of the cDNA for different SHRs has greatly facilitated the molecular understanding of their actions as well as receptor protein structure and function (26–33). In animals, the main classes of steroid hormones are progestins (C22 steroids), corticosteroids (C21), androgens (C19), and estrogens (C18), whereas ecdysone (C27), the moulting hormone, is found in insects. Related molecules that also act as ligands for members of the nuclear receptor superfamily include bile acids and the sterol, vitamin D3. On the basis of amino acid sequence homology and DNA binding specificity, the classical SHRs are found in subfamily III, which includes estrogen receptors (ER) α and β (NR3A1 and A2 respectively); the estrogen-related receptors (ERR) α, β, and γ (NR3B1-3); the glucocorticoid receptor (GR; NR3C1); the mineralocorticoid receptor (MR; NR3C2); the progesterone receptor (PR; NR3C3); and the receptor for androgens (AR; NR3C4) (25). The receptors for vitamin D3 (VDR; NR1I1) and ecdysone (NR1H1) are found in subfamily I and act as heterodimers with retinoid X receptors (RXR) α, β, or γ (NR2B1-3) or the insect homolog ultraspiracle, respectively.

It is well established that transcriptional activation of genes involves the regulated assembly of multiprotein complexes on enhancers and promoters (34–36). However, despite an exponentially increasing number of reported coregulatory protein binding partners for SHRs and the mapping of receptor interacting domains within these proteins, the mechanisms by which SHRs function remains an area of important investigation. Furthermore, observations from both basic biochemical analysis and recent cell-based studies led to the inescapable conclusion that in cells, SHRs behave very dynamically, showing rapid and reversible interactions with partner proteins, together with chromatin and DNA (Fig. 1) (37–41). The coupling of the functional and structural dynamics of SHRs with differences in the local concentrations of potential coregulatory proteins is likely to result in receptors bringing differing sets of binding partners together in response to agonist or antagonist ligands, such that an agonist in one cell type can be an antagonist in another cell type (42–45). Binding of agonistic or antagonistic ligands leads to different allosteric changes in SHRs, making them competent to exert positive or negative effects on the expression of target genes by different mechanisms. The result is SHR-mediated modulation of specific responses in different tissues depending on the physiological and genetic context of the cell (46–48).

Although some of these complex behaviors have been explained by identifying cell-specific binding partners for SHRs, it is as yet unclear whether other dynamic considerations must be taken into account while defining the underlying mechanisms, especially given the ubiquitous nature of a large number of known coregulatory proteins. We propose that the often disparate effects of selective ligand modulators of SHR function, the many and varied nonconsensus DNA response elements for SHRs, and the effects of proteins that bind reversibly to various sites of SHRs can be explained by the intra- and interdomain coupling and structural dynamics within the receptor protein. Although conformational stability of most proteins determines their biological function, it has been proposed that the structural plasticity of the SHR proteins can explain much about their observed behaviors (49–51). The good and bad effects of steroids on physiology, pathophysiology, and in therapeutic use are well documented. We contend that understanding the differential effects of ligands, response element sequences, and various protein binding partners on the dynamic structure of SHRs will provide essential knowledge for those seeking to design specific steroid (and nonsteroidal) selective SHR modulators (SRMs), which minimize adverse side effects.

A functionally competent structure can be achieved by the SHR protein through different regions of the receptor rapidly and reversibly adopting various conformations, controlled by ligand and DNA binding. This in turn may help to create protein surfaces that are readily available for selective binding to coregulatory proteins, resulting in SHR-mediated transcriptional regulation of target genes (Fig. 1). How and under what conditions such surfaces are created to include or exclude coregulatory proteins in a promoter- and cell-specific manner is an open question. The answer may in part lie with the dynamic structure of SHRs combined with allosteric regulation mediated through inter- and/or intradomain cross talk, which can be influenced by hormone binding, DNA architecture, and the binding of coregulatory proteins. In this review article, we discuss the properties of the SHRs that allow this discriminatory action, which is critical to SHR.
transduction of different signals. It is timely in light of recent advances in our understanding of SHR protein structure and the growing interest in targeting regions of the receptor distinct from the ligand binding pocket, with small molecule modulators that may ultimately lead to a new portfolio of tailored therapeutic treatments for hormone-dependent diseases.

II. Functional and Structural Organization of Steroid Hormone Receptors

From early biochemical studies with purified receptor proteins, it was deduced that SHRs would consist of discrete regions important for hormone and DNA binding (52). Limited proteolysis and antibody binding identified a third distinct region, termed the “immunoreactive” domain (53, 54). The cloning of the ERα and GR cDNA in the mid-1980s and subsequent deletion mapping studies refined this domain organization into the picture we have today (Fig. 2A). SHRs are modular proteins that consist of a highly variable N-terminal domain (NTD; also called A/B domain); a 66-amino acid core DNA binding domain (DBD; or C domain); a flexible hinge domain of 40 to 60 amino acids (D domain); a ligand binding domain (LBD) of about 265 amino acids (E domain); and in some receptors a variable stretch of amino acids at the very C terminus referred to as the F domain (Fig. 2A) (24). Two activation domains, termed AF1 and AF2, are located within the NTD and LBD, respectively, and are responsible for mediating the transcriptional activity of SHRs. The two AF can act in an independent manner, although usually full receptor activity requires synergistic cooperation between AF1 and AF2. This cooperativity can be mediated by direct receptor N/C-terminal interactions or the binding of coregulatory proteins to either activation function separately or simultaneously to both AF1 and AF2.

A. Ligand binding domain (LBD) (AF2)

1. Hormone binding

In the absence of hormone, SHRs are present in a heterocomplex with heat shock proteins and immunophilins (Fig. 1). This complex has been particularly well studied for the GR and consists of a dimer of hsp90, together with one copy of hsp70, p23, and the immunophilins FKBP51 and -52; the latter have been proposed to link SHRs to the motor protein dynein and the microtubule retrograde transport network (reviewed in Refs. 55 and 56). The binding site for hsp90 on the GR has been mapped to a seven-amino acid sequence at the N-terminal end of the LBD (residues 529 to 535 for human GR), but the exact sequence of amino acids was found not to be critical for this interaction (57). Importantly, the binding of the hsp90 heterocomplex was essential for maintaining the GR in a

---

Figure 1. General dynamics and mechanism of SHR action. SHR action is spatially, temporally, and structurally highly dynamic. In the absence of hormone, the receptor is complexed with co-chaperone molecules (green) in the cytoplasm. The NTD/AF1 exists in an ID conformation, compared with the well-ordered DBD and LBD. Factors affecting disorder-order transition of the SHR AF1/NTD are illustrated. The binding of hormone (I) causes rearrangement of regions in the LBD (helix 12, AF2), and this leads to translocation to the nucleus and binding to DNA response elements and coregulatory proteins (blue) (II). Under the influence of factors shown, NTD/AF1 undergoes disorder-order transition, resulting in the folding of NTD/AF1. In this conformation, AF1/NTD surfaces are well suited for the interaction with specific coactivators, binding of which further influences the conformation of NTD/AF1 and facilitates the assembly of the transcription initiation complex (blue) in a promoter-specific manner (III). The binding of DNA and/or coregulatory proteins may allosterically regulate ligand binding, which could lead to formation of SHR-transcription complexes (IV). The formation of transcriptionally competent complexes by the DNA-bound SHR (III) must be readily reversible, and the co-chaperone complex (green) may aid in dissociating the SHR-chromatin complex, recycling of the receptor protein to the cytoplasm, and stabilization of the ID NTD/AF1 (V). The SHR is also subject to posttranslational modifications, such as phosphorylation (P), which are likely to fine-tune the allosteric regulation of receptor structure and function.
state primed to bind hormone because reducing levels of hsp90 in yeast cells inhibited receptor action (58). In addition to a role in hormone binding, interactions with hsp90 are important for cycling of SHRs on DNA/chromatin and for receptor turnover (56, 59) (Fig. 1).

The primary role of the LBD is to bind specific steroid ligands, which become buried in the hydrophobic interior. Selectivity of different hormones is based upon the size of the ligand binding pocket and interactions between specific amino acids and the A and D rings of the steroid
molecule (reviewed in Ref. 60). From analysis of the amino acid sequences of SHR-like proteins from different organisms, including invertebrates such as molluscs and annelids, it has been proposed that the ancestral gene for SHRs was ER-like and bound estrogens (reviewed in Ref. 61).

The binding of hormone has dramatic consequences for individual receptor structure-function relationships. Structures of the LBD for each class of SHR bound to either agonists or antagonist are available. The LBD is made up of 11 α-helices and a variable number of β-strands; for example, the ERα-LBD has 11 helices and two β-strands (Fig. 2B) (reviewed in Ref. 60). The agonist-bound receptors share a globular folded structure that has been described as a three-layer helical sandwich, with the hormone buried within the ligand binding pocket and sealed in by the conformational change in helix 12 (note that convention names the helices 1 to 12 after the first structures solved; in the SHR subfamily, helix 2 is absent or disordered). The reorientation of helix 12 in response to hormone binding, together with residues in helices 3, 4, and 5 make up the AF2 surface (62–64). This surface pocket has a hydrophobic interior and typically binds coactivator proteins containing the peptide motif LxxLL (where L is leucine and x is any amino acid) that forms a α-helix, which buries the leucine residues within the groove (62). In addition to the hydrophobic surface, the LxxLL helix is orientated by a lysine residue in helix 4 and a glutamic acid in helix 12, which together form a “charge clamp” (63, 65–67).

Extensive analysis of ERα and ERβ LBD revealed the importance of residues outside the ligand binding pocket that contributed to receptor-specific ligand interactions and allosteric communication with the coactivator binding surface (64). In particular, the position of helix 11 was implicated in ER subtype-selective agonist, partial agonist, or antagonist responses to a particular ligand. An allosteric network linking the ligand binding pocket and the AF2 surface was also elegantly demonstrated through a mutation screen with the GR. A number of mutations were identified that reduced the requirement for hsp90 in hormone binding by the receptor: two of the residues, tyrosine 598 (numbering for human GR) in helices 5–6 and methionine 752 (helix 12) contribute to the AF2 surface, whereas a third residue, methionine 604 (helices 5–6) is part of the ligand binding pocket, which provides a structural basis for the allosteric regulation of ligand and coregulatory protein binding to the LBD (68). This study also revealed a further role for hsp90 in connecting this allosteric network.

The structure of the LBD of ER and AR bound to antagonists or SRMs have also been solved (e.g., Fig. 2B). In these cases, it is interesting that the overall fold of the LBD remains similar to that of the agonist-bound receptor, but there are changes within the ligand binding pocket or in the position or flexibility of helix 12. In the case of the ER SRMs, tamoxifen (Fig. 2B) and raloxifene, displacement of helix 12 prevents coregulatory proteins binding to the AF2 surface (60, 69, 70). The agonist activity of tamoxifen in certain cells has been shown to depend on AF1 because removal of the NTD of the receptor abrogates the activation of a target gene by this ligand (71–73).

Interestingly, other surfaces on the AR-LBD have been shown to bind small molecules that interfered with the binding of coregulatory proteins to the AF2 site (74). It is tempting to speculate that this site, termed “BF3,” is an auxiliary protein-protein binding site or that it could be exploited as an allosteric drug binding surface.

2. Dimerization interfaces

Sequences within the LBD have also been identified as dimerization interfaces. ER-LBD dimerization involves amino acids in helix 10, but the dimerization interface for the GR-LBD comprises residues in helices 1, 3, and 5 (43, 75–77). However, whether the presence of a dimerization site within the LBD is common to all SHRs is less clear. In the case of the AR, this is a subject of debate, and it is noteworthy that in the crystal structures of the AR-LBD, only monomers were observed. A recent comparison of the AR-LBD with the GR sequence and structure identified residues in helices 1, 3, and 5 as a potential dimerization interface (78). The involvement of different regions of the LBD in dimerization suggests that this interaction could be differentially regulated allosterically by hormone binding, which in turn could lead to cell- or gene-selective responses through the exposure of different surfaces for coregulatory protein binding.

3. N/C-terminal interactions

Interestingly, whereas other SHRs interact with coactivator proteins via LxxLL motifs, which bind in the hydrophobic groove of AF2, the AR-LBD preferentially binds the AR-NTD and coactivators with more bulky hydrophobic residues in the sequence F/WxxLF/W/Y (65–67, 79–81). This AR-N/C terminal interaction results in stabilization of the receptor protein, a reduction in the dissociation of hormone and selective gene activation (82–84). There is also growing evidence that N/C-terminal domain interactions occur for other SHRs, for example ERα (85, 86), PR (87, 88), and the MR (89).

The structural basis for the AR-LBD/AF2 binding preference for bulky hydrophobic aromatic residues over leucine has been reported (65–67). X-ray crystallography studies showed that the AR-NTD FxxLF motif forms a charge clamp with glutamic acid 897 in helix 12 and lysine...
720 at the C-terminal end of helix 3, and the bulky hydrophobic side chains fit better into the surface pocket on the LBD. In contrast, an LxxLL motif peptide fails to make hydrogen bond contacts with the glutamic acid residue in helix 12 and makes fewer hydrophobic contacts with the surface of the LBD (65–67).

It can be concluded from the above studies that the binding of hormone acts as an allosteric switch to regulate SHR-DNA and SHR-protein interactions, including interdomain interactions and/or dimerization.

B. Hinge region

The binding of steroid leads to exposure of the nuclear localization signal (RKxK/RK), within the hinge domain, and translocation of the receptor-steroid complex into the nucleus and subsequent DNA binding (Fig. 1). The hinge region is a 40- to 60-residue linker sequence between the DBD and the LBD, which shows little amino acid homology between different SHRs. The hinge region contains the C-terminal extension (CTE) of the DBD, the nuclear localization signal (NLS), and sites for posttranslational modifications (Fig. 2A).

The CTE is defined as the sequence immediately C-terminal of the conserved second zinc finger of the core DBD. In the case of nonsteroid receptors, for example in the VDR and the thyroid hormone receptor β (NR1A2), the CTE forms an α-helix and is important for heterodimerization with RXR and minor groove DNA contacts (reviewed in Ref. 90). In the original structures of the GR and ER and subsequently the AR DBD, the CTE appeared disordered, although mutational studies highlighted a role in DNA binding (91–94). Interestingly, recent structural analysis of a number of GR-DBD-DNA complexes observed a helical conformation for the CTE (95). Electron densities were also observed for the first seven to eight residues of the PR-CTE, which revealed interactions with both the minor groove and the core PR-DBD (96, 97). Mutational analysis demonstrated that the PR-CTE was necessary for recognizing nucleotides flanking the DNA response element (96). More recently, the hinge region of the PR was found to regulate the response of the receptor on different promoters. This involved cytoplasmic/nuclear shuttling and retention in response to acetylation of lysines in the NLS and phosphorylation of serines in the NTD (98).

Structural and functional analysis of the AR-hinge region (residues 629 to 634) showed that the NLS interacted with the armadillo repeats 2, 3, and 4 of the NLS receptor protein α-importin (99). An extended receptor peptide (amino acids 617 to 634) adopted a β-turn-like structure when bound to α-importin, which was unique to the AR NLS (99). Recent mutational analysis of this basic motif in the AR demonstrated a number of functional interrelationships, including nuclear localization, DNA binding, intranuclear dynamics, and transcriptional activity (100). The lysines in this motif are also subject to acetylation (101–103) and methylation (104).

Collectively, the above-mentioned studies strongly suggest that the CTE can adopt different conformations depending upon whether the receptor is bound to DNA, protein partners, or free in solution (99, 105, 106). Furthermore, because the CTE itself appears to be an intrinsically disordered (ID) region, it suggests the possibility that CTE interactions are involved in mediating allosteric coupling with the NTD and LBD (discussed in Sections V.B.2. and V.I).

C. DNA binding domain (DBD)

On the basis of amino acid sequence homology and DNA binding specificity, SHRs can be classified as members of the ER subfamily (ERα, β, and ERRα, β, and γ) and the GR subfamily (GR, MR, PR, and AR) (25). The core DBD of SHRs consists of approximately 66 amino acids and involves the coordination of two zinc ions by eight cysteine residues (90, 107). SHRs typically bind in vitro to palindromic or palindromic-like DNA sequences as homodimers. The consensus sequences for ER have the half-site 5′-AGGTCA-3′, and for GR, PR, MR, and AR the half-site is 5′-AGAACAA-3′; each half-site is separated by a three-nucleotide spacer, completing the 15-bp receptor response element (90, 107). Early pioneering mutagenesis and “domain swapping” experiments identified a sequence of three amino acids, termed the P-box, within the first zinc-finger module as being important for DNA response element recognition and binding. In the case of the ER, these residues are glutamic acid, glycine, and alanine, whereas the corresponding amino acids in the GR subfamily are glycine, serine, and valine (108, 109). In the second zinc-finger module, a five-amino acid sequence, the D-box, was found to mediate dimerization of the receptor on DNA (90, 107).

The first SHR structures to be solved were the GR-DBD by nuclear magnetic resonance (NMR) spectroscopy (91) and the solution and crystal structures of the ERα-DBD (92, 93). These studies revealed that the DBD folded into a compact globular structure containing two principal α-helices perpendicular to each. The recognition helix fitted into the major groove of DNA and contained the P-box residues (Fig. 2C), which made direct and water-mediated hydrogen bonds with the nucleotide sequence. In addition, there are a number of interactions between amino acid side chains and the phosphate backbone of the DNA. The second helix acts to stabilize the recognition helix in place. The five amino acids of the D-box form
an exposed loop structure (Fig. 2C) and mediate the “head-to-head” binding of the receptor monomers on the DNA response element.

Although the overall folding of the DBD is very similar for all receptors studied, there are some differences that are thought to be functionally important. For example, the structure of the AR-DBD forms a more closely packed dimerization interface than that observed for the GR-DBD, and it has been proposed that this, in part, is the basis for AR binding to androgen response elements, a subset of DNA response elements that are selective for this receptor (94). Furthermore, whereas the main determinants for DNA binding reside within the P-box residues, it has become apparent that the region immediately adjacent to the DBD, termed the CTE, also has a role to play in stabilizing DNA binding and/or response element selection by SHRs through contacts with the minor groove of DNA (96, 110).

A recent structural analysis has highlighted the allosteric properties of DNA binding. Receptor binding sites associated with glucocorticoid-regulated genes were shown to influence the composition of coregulatory complexes recruited by the GR (95). The binding of the GR-DBD to different response elements was found to be asymmetric, and a region of the DBD, the “lever arm,” was found to adopt different conformations depending on the DNA architecture (95). The lever arm consists of the amino acids Glu450-Gly-Gln-His-Asn-Tyr455, which are conserved in the PR- and MR-DBD, but shows three and four amino acid changes in the AR and ER, respectively. The histidine at position 453 adopted one of two possible conformations in the DBD dimer: in the upstream monomer the histidine was packed in the DBD core, whereas for the downstream monomer the residue was “flipped out” and conformationally flexible (95). The glutamic acid and the tyrosine participated in contacting DNA; both these residues are found in AR, MR, and PR, but only the tyrosine is found at the same position in the ERα/β. It remains to be investigated more fully to determine whether the lever arm is a common property of the SHR-DBD and to determine the mechanism(s) of allosteric control to the GR-AF1 or -AF2 activities. These studies reveal the structural dynamics imparted by DNA recognition and binding and highlight the potential regulatory role of different DNA sequences or binding partners, which could lead to cell or promoter selective SHR function.

Recent advances in genome-wide analysis of SHR-regulated genes have dramatically increased our ability to identify hormone-regulated gene networks and have allowed for the identification of SHR binding regions (SBRs) throughout the genome. The outcome of these studies has been the revelation that receptor binding sites can have degenerate sequences and different architecture of half-sites (palindromes, inverted palindromes, and repeats), although typically a high-affinity half-site matching the consensus sequence is always present (111–117). The other outcomes from these chromatin immunoprecipitation (ChIP), or ChIP-seq, studies are the observations that: 1) there is a high preponderance of other transcription factor binding motifs associated with receptor response elements (e.g., Foxo A1, ETS1, GATA2); and 2) in some cases the SBR can be more than 10 kb away from the transcription start site of the regulated gene (111–113, 116, 117), although other studies reported SBR within 1.5 kb of the transcription start site (112, 115). Of particular interest was the comparison of genomic binding sites for the GR across four mammalian species, which revealed that sequence conservation for individual sites was greater between species than for receptor binding sites within the same species (116, 118). The latter observation emphasizes the potential importance of DNA architecture of SBR for gene selective regulation by the GR.

D. N-terminal domain (NTD) (AF1)

The NTD of SHRs is highly variable in both amino acid sequence and length, with the MR, PR-B isoform, and AR having NTD of greater than 500 amino acids, the largest of any members of the nuclear receptor superfamily (reviewed in Ref. 50). In contrast to the DBD and LBD, there is little amino acid sequence homology between the different classes of SHR-NTD. However, comparison of the primary sequence for a given receptor across different species does reveal regions of conserved residues (50). Such conserved sequences often colocalize with regions delineated as necessary for receptor-dependent gene regulation, which in turn are involved in protein-protein interactions (49, 50). Deletion of the LBD of SHRs results in a protein that is constitutively active in reporter gene assays, and thus the function of AF1 has been described as hormone-independent (see Ref. 50 and references therein). A significant number of coregulatory proteins have been identified binding to AF1, including chromatin modifying enzymes, basal transcription factors, coactivators, and corepressors (reviewed in Ref. 50). In some cases, the same binding partner has been described as interacting with both AF1 and AF2 through distinct receptor-interacting domains (119, 120).

Deletion studies have defined the AF1 of the ERα (72, 121) and GR (122) to 100 and 200 amino acid regions of the respective NTD. The location and nature of the GR-AF1 was further refined to a 58-residue core domain (amino acids 187 to 244; GR-AF1core) (123). This AF1 core domain retained 60 to 70% of the activity of the full-length AF1. Point mutational studies revealed an im-
portant role for hydrophobic amino acids in AF1core activity (124, 125), whereas mutations that reduced the overall acidity of the AF1 domain led to progressive impairment of transactivation (126). It was speculated that these residues were important structurally and defined the solvent exposed surface of the transactivation domain. It is significant, therefore, that mutating key glutamic acid, phenylalanine, or tryptophan residues in the corresponding enh2 domain (amino acids 108 to 317) of the rat GR impaired transactivation activity, but not the ability to repress transcription through protein-protein interactions with the transcription factor activator protein 1 (127).

Although the term “AF1” is used to describe the transactivation function of the SHR-NTD in some receptors, the activity is highly modular and can map to multiple sequences within the NTD. Deletion of the AR-NTD results in a transcriptionally weak protein, providing evidence for the main transactivation function being located within the NTD (128, 129). The AR-AF1 is modular in nature, and regions important for transactivation have been characterized by deletion analysis (128, 129), by the use of fusion proteins (129), and by point mutation (130, 131). These studies identified two overlapping regions, amino acids 101 to 370 and 360 to 485, as being critical for receptor-dependent transactivation, and highly conserved hydrophobic amino acids within this region have been shown to be important for activity and protein-protein interactions.

Similarly, multiple regions have been identified within the PR-NTD (132) and MR-NTD (133–135) that, when deleted, impaired receptor-dependent transactivation. The AF1 domain or related SHR-NTD activities have been shown to function in a variety of mammalian cell types and in the budding yeast Saccharomyces cerevisiae and when fused with a heterologous DNA binding domain. A striking exception to the later observation is the isolated PR-AF1 domain, which requires the PR-DBD to be functionally active, emphasizing the importance of intradomain communication (132). When taken together, these studies illustrate that different surfaces, even within a single SHR-NTD, can be involved in activation, or indeed repression, of transcription. Furthermore, the strength of the respective SHR-NTD in activating transcription is also variable. A clear correlation between AF1 activity (and inversely AF2 activity) and the length of the NTD has been demonstrated using GalPRDBD-NTD fusion proteins, with the longer AR- and PR-NTD showing the highest activity in a reporter gene assay (66).

III. Structural Analysis of Full-Length Nuclear Receptors

Significant insight has been provided by x-ray crystallography and/or NMR spectroscopy of the isolated DBD and LBD. However, the presence of significant regions of intrinsic disorder (Section V) and overall protein size have hampered attempts to determine the three-dimensional structure of a full-length SHR.

In contrast, a recent x-ray structure of a complex containing the full-length peroxisome proliferator-activated receptor (PPAR) γ and RXRα nonsteroid nuclear receptors has been published (105) (Fig. 2D), whereas solution studies, using small angle x-rays and fluorescence resonance transfer, have described the “shape” of a number of nonsteroid nuclear receptor complexes (136). Isothermal calorimetry and hydrogen deuterium exchange (HDX) experiments have elegantly revealed interdomain and inter-receptor communication for the thyroid hormone receptor (137) and the VDR (138) heterodimer complexes with RXR, respectively.

In the crystal structure of PPARγ and RXRα bound to a DR1 DNA response element (5′-AGGTCA/AGGTCA-3′), the complex is asymmetrical, and the overall conformation of the individual LBD and DBD were remarkably similar to the structures of the isolated domains and those described above for SHRs (Fig. 2D) (105). In the crystal complex, the PPAR monomer adopted a “closed” conformation with extensive interactions between the LBD and the PPAR-DBD and the RXR-LBD, Hinge, and DBD (105). In contrast, the RXR monomer had an “open” conformation with the hinge region extended, creating a surface for PPAR binding. The hinge region of the PPAR also made contact with the DNA and contained two α-helices, whereas the RXR hinge region was more flexible and lacked secondary structure (Fig. 2D) (105). The functional importance of the interaction of the PPAR-LBD with the DBD was supported by the introduction of a F347A mutation. This residue is not directly in contact with either the ligand binding pocket or the AF2 region, but impaired DNA binding and transcriptional activation (105). The NTD of both receptors, although relatively short at 110 and 134 amino acids, respectively, were not observed in the crystal structure of the complex and were highly dynamic as determined by HDX experiments.

However, these different conformations of the hinge region or the closed complex of PPAR were not observed in the solution structures of a number of class II receptors bound with RXR to different response elements (136). Although these studies lack atomic resolution, they resolved as a homogenous complex and generally all show the receptors as an elongated shape in an open conformation, with the hinge regions in extended conformations, permitting the ordering of the LBD over the 5′ half-site of the DNA element (136). Furthermore, the coactivator protein Med1 bound with a 1:1 stoichiometry with the receptor heterodimer complex and only bound to the RXR

In the budding yeast Saccharomyces cerevisiae reporter gene assay (66).
partner (136). This is in contrast to other studies that suggest binding of coregulatory protein fragments to both receptors in similar receptor complexes, although binding to one partner could be modulated by the DNA response element (138).

Zhang et al. (138) have used HDX to investigate the consequences for receptor conformation of ligand binding, DNA response element binding, and coregulatory protein interactions using the VDR-RXR heterocomplex as a model. Binding of either ligand alone caused changes in HDX within the cognate receptor LBD and the LBD of the receptor partner. As might be expected, a number of these changes mapped to regions of the receptors involved in heterodimerization (138). However, there were also perturbations at distant sites; for example, 1,25-vitamin D₃ binding caused changes in helix 3 of the RXR partner, suggesting allosteric interreceptor communication. Most striking was the destabilization of the VDR-DDB by binding of either 1,25-vitamin D₃ or 9-cis-retinoic acid (138). DNA response element recognition and binding also resulted in changes in HDX of both receptors; there was strong protection from solvent exchange for the VDR-DDB/CTE, consistent with more contacts with the DNA (138). DNA binding, and significantly the architecture of the DNA response element, also led to changes in the LBD of both receptors: the regions affected were the dimerization interface and AF2 surface. A 1:1 complex was formed between a fragment of the steroid receptor coactivator-1 (SRC-1) (NCoA1) and VDR-RXR: with VDR binding to the NR3 box and RXR to NR1 box (p160 family of coactivators, NCoA1, 2, and 3, have three LxxLL motifs termed NR box 1 to 3). The nature of the DNA response element was shown to alter the conformation of the AF2 regions such that there was a reduction in VDR and enhancement of RXR interactions with SRC-1 (138).

Taken together, the structural analysis of full-length or two-domain receptor proteins, in complex with different DNA response elements, illustrates the complexity of multiple inter/intradomain interactions and the possible mechanisms of allosteric regulation imparted upon ligand, DNA, or coregulatory protein binding. Furthermore, the above-mentioned structural studies have highlighted the growing importance of the hinge domain in regulating receptor conformation and complex “shape.” In light of these studies with class II nuclear receptors, which form heterodimeric complexes with RXR, it will be vital to solve the structures for SHR complexes containing at least two domains. Because these receptors bind to palindromic-like DNA sequences as homodimers, such studies are required to determine what conformation(s) are adopted by the hinge regions of each monomer and to identify the intra/interdomain communications, which would highlight the potential mechanisms of allosteric regulation in response to hormone and DNA binding.

IV. Steroid Hormone Receptors Function in a Ligand-, Cell-, and Promoter-Specific Manner

Allosteric regulation is a property that arises from the modular structure of SHRs. The consequence of this is that multiple surfaces can engage in protein-protein interactions (i.e., AF1, AF2, lever arm, and the CTE), and accessibility of these sites is regulated by the binding of hormone and the DNA architecture. Thus, the means of gene- and tissue-specific regulation reside within a single transcription factor. The cellular context (SHR protein levels and coregulatory protein levels) and chromatin environment will also influence the SHR transcriptional activity, but in this review we emphasize the features within the receptor protein that underpin this regulation.

A. Selective utilization of AF1 and AF2: cell culture studies

Initial studies mapping the functional domains of the ERα led to the observation of cell and promoter (gene) discriminating functions for the AF1 and AF2 surfaces. Strikingly, using receptor constructs lacking the NTD or the LBD, it has been shown that AF1 and AF2 function in a cell type-specific manner. For example, AF1 was active in CEF (chicken embryonic fibroblast) cells, HepG2 liver cells, and the breast cancer cell lines T47D and MCF-7, but not in HeLa cells (42, 73, 139). Furthermore, dependence on AF1 appeared to correlate with the more differentiated breast and prostate cancer cell lines, whereas AF2-dependent activity showed the opposite trend (73).

The activation functions of ERα also showed promoter selectivity, with AF2 being active on more complex promoters and having the ability to synergize with both AF1 and other transcription factors (71, 139). In contrast, in CEF permissive cells, AF dependence did not distinguish between “complex” and “simple” (response element + TATA-box) promoters and failed to act cooperatively with other transcription factors (139).

Evidence for differential utilization of the GR-AF1 and -AF2 domains comes from studies using the human osteosarcoma cell line U2OS expressing wild-type or variants of the GR with mutations in either AF1 (E219K/F220L/W234R) or AF2 (E773R) (95, 140, 141). After expression profiling to identify GR-regulated genes in U2OS cells, a panel of nine genes were selected for investigation. On the basis of expression in response to the synthetic glucocorticoid dexamethasone, the genes were classified as: AF1-dependent (IGFBP); AF2-dependent
shown that ER has only recently become available. SHR activation functions, but compelling mental in highlighting the potential for differential use of ovariectomized animals treated with estradiol (17). ER but not cortical bone (17). Together, these AF1 was more tissue specific, having a role in trabecular, cortical and trabecular bone; crucially the requirement for it was demonstrated that AF2 was important for both necessary for maintaining bone density and structure, and 84). Collectively, these studies illustrate the importance of different regions of SHRs (AF1/NTD, AF2/LBD) for target gene expression and provide a starting point for evaluating the mechanisms for this selectivity, which is likely to involve specific protein-protein interactions and posttranslational modifications.

B. Selective utilization of AF1 and AF2: in vivo studies

The above-mentioned in vitro studies have been instrumental in highlighting the potential for differential use of SHR activation functions, but compelling in vivo evidence has only recently become available. In vivo it has been shown that ERα-AF1 was not required for vasoprotective actions of estradiol but was important for ERα function in reproductive tissues (144). A similar transgenic mouse model has been used to investigate the role of ERα activation functions in bone: comparing the consequences of a complete ERα gene deletion with either a deletion of the ERα-NTD(AF1) or a deletion impairing AF2 function in ovariectomized animals treated with estradiol (17). ERα is necessary for maintaining bone density and structure, and it was demonstrated that AF2 was important for both cortical and trabecular bone; crucially the requirement for AF1 was more tissue specific, having a role in trabecular, but not cortical bone (17). Together, these in vivo studies could show that AF1 activity was important for estradiol action in the uterus, but not for the increase in liver weight in response to hormone treatment (17, 144). The significance of these findings lies not just in providing evidence for tissue/cell-specific selectivity of AF1 and AF2, but also in illustrating the possibility of therapeutically targeting these receptor functions to achieve tissue restricted effects. Thus, blocking ERα-AF1 to treat uterine cancer should not impact upon ERα activity in maintaining cortical bone.

The above-mentioned studies demonstrate the potential for different SHR surfaces to be used in a cell- and promoter-specific manner. Further experiments are clearly needed to extend these observations and to determine the molecular mechanisms regulating the display and function of different SHR surfaces.

V. Intrinsically Disordered Structure and Steroid Hormone Receptor Action

The traditional view in biology is that the specific function of a given protein is determined by its unique three-dimensional structure, the so-called “lock and key” hypothesis. However, in recent years, it has become quite evident that many biologically important proteins possess large stretches of amino acid sequences that do not adopt a well-defined three-dimensional structure (145–150). These unstructured proteins/protein regions have been termed “intrinsically disordered” and exist as dynamic ensembles of interconverting conformers that are capable of undergoing disorder/order transition under specific physiological conditions (Fig. 3) (151–154). However, in contrast to structured or ordered proteins whose conformation is relatively stable with occasional cooperative conformational switches, they do not automatically adopt a classical fully folded and well-defined functional structure and typically undergo conformational changes (155–159). Both random coil-like and collapsed (partially folded or molten globule-like and premolten globule-like) conformations with poorly packed side chains are the features of ID proteins/regions (Fig. 3) (160, 161).

In recent years, it has become evident that eukaryotic genomes are highly enriched in ID proteins, which appear to promote molecular recognition (162). Based on biophysical and computational analyses, ID regions/domains are prevalent in a majority of transcription factor proteins, including the hinge and NTD of SHRs (49, 50, 163–166). Characterization of the conformational propensities and function of such nonglobular protein sequences represents a major challenge.

A. What does it mean to be intrinsically disordered?

The complex and highly coordinated interactions of proteins play a fundamental role in the control of cellular physiology, where different functions can be achieved via recognition of specific and unique identification sequences frequently found inside ID regions (145–150, 154). It has been predicted that a number of signaling proteins (~65%) possess long ID regions/domains, which play a
critical role in cell cycle control, transcriptional and translational regulation, and signal transduction; this number is predicted to be much higher (~75%) for cancer-associated proteins (165, 167). The significance of such ID domains/regions in signaling molecules is that their conformational flexibility creates large interaction surfaces that allow macromolecular interactions with high specificity and low affinity through coupled binding and folding, an important property of signaling proteins promoting molecular recognition that provide specific, but reversible, interactions with target molecules (151–154). The function-associated conformational changes and disorder-to-order transitions may be brought about by alterations in environmental or cellular conditions. The ID nature of proteins also provides a possible mechanism for their actions to be regulated through posttranslational modifications, such as kinase-dependent phosphorylation (153, 180). Furthermore, sites of other posttranslational modifications such as acetylation, hydroxylation, ubiquitination, methylation, and sites of regulatory proteolytic attack are frequently associated with ID regions (168). The functional importance of conformational changes and disorder-to-order transitions in ID proteins stems from a large decrease in conformation entropy, which in turn can uncouple specificity from binding strength. This enables highly specific interactions to be easily reversible, which is beneficial for proteins involved in signaling and transcriptional regulation.

B. Coupled folding and binding of the steroid hormone receptor NTD

The primary amino acid sequences of the NTD of the SHR, which contain AF1, are much less conserved than are DBD and LBD regions (50, 60, 169–171). Despite having poor sequence homology, these regions are highly enriched in charged amino acids and have low hydrophobicity, which is a signature for ID (49, 50). Studies from several laboratories have confirmed the ID nature of several SHR-NTD or AF1 domains using circular dichroism-, NMR-, Fourier transformed infrared-, and fluorescence emission- spectroscopies and proteomic methods (135, 172–176).

Table 1 summarizes the secondary structure content for different SHR-NTD and AF1 activities under different experimental conditions. From these studies, it can be concluded that the NTD typically lack stable secondary structure but have the propensity to form α-helical conformation. Interestingly, the AF1b region of the MR-NTD appears to be more stably folded than the rest of the domain and to have predominantly β-strand secondary structure (135). SHR-NTD therefore exist as an ensemble of conformations having more or less stable secondary and tertiary structures (Fig. 3).

It is well known that the SHR AF1/NTD generally work in conjunction with other coregulatory proteins and by multiple mechanisms (177–182). This raises the question: what is the structural basis of the functional activity of the ID NTD/AF1 in the context of full-length SHR action? As discussed in Section III, studies carried out so far, including analysis of the crystal structure of the full-length PPARγ and RXRα (105), suggest that the AF1 may be unstructured even in full-length receptor. Furthermore, in the case of the GR, PR, and AR, the NTD is ID within the context of a two-domain fragment containing the entire NTD and DBD (174, 183, 184). It can be hypothesized that one of the reasons why the holo-receptor is associated with several chaperone proteins in the cytosol, before ligand binding, is protecting this large ID region from degradation. However, available data indicate that all the known chaperone proteins bind to the LBD of the receptor (see Section II.A), except in the case of the AR-NTD interactions with the C terminus of the Hsp70-interacting protein (246) and the co-chaperone protein Bag1L (185, 186). That raises another question: if these proteins bind...
to the LBD, how do they protect NTD structure? The answer to this question may lie within the dynamic structure of the SHR protein that allows sensing of the intracellular environment and efficient intramolecular domain interactions, which in turn may induce structure in the otherwise ID NTD. In fact, this hypothesis is supported by the finding that certain chemical chaperones known to protect/stabilize protein structure are capable of inducing a compact structure in the AF1/NTD of the AR, GR, and MR (135, 176, 187–190) (Table 1 and Fig. 4) and the growing evidence for N/C-terminal interactions among SHRs (see Section II.A.3).

Generally, under physiological conditions, proteins must have specific structure to carry out their proper functions. The GR-AF1 domain was the first region of any nuclear receptor to be characterized as being ID (172, 173). It was further reported that this ID region can be induced to adopt a helical structure when incubated in the presence of trifluoroethanol (TFE) (Table 1). At that time, it was hypothesized that the GR-AF1 might not require an ordered conformation and act through the so-called “acid blob” concept (172). However, several mutagenesis stud-

### Table 1. Measured secondary structure in nuclear receptor NTD and or AF1 domains

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Conditions</th>
<th>α-helix</th>
<th>β-strand</th>
<th>β-turn</th>
<th>Bend</th>
<th>Coil</th>
<th>Method</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR-AF1</td>
<td>Buffer</td>
<td>13</td>
<td>20</td>
<td>32</td>
<td>36</td>
<td>CD</td>
<td>FTIR</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>Buffer</td>
<td>16</td>
<td>24</td>
<td>17</td>
<td>19</td>
<td>24</td>
<td>CD</td>
<td>176</td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>40</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>CD</td>
<td>FTIR</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>TMAO + TFIIF</td>
<td>37</td>
<td>17</td>
<td>11</td>
<td>11</td>
<td>24</td>
<td>FTIR</td>
<td>188</td>
</tr>
<tr>
<td>AR-NTD</td>
<td>Buffer</td>
<td>14</td>
<td>27</td>
<td>24</td>
<td>34</td>
<td>CD</td>
<td>FTIR</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>58</td>
<td>12</td>
<td>9</td>
<td>22</td>
<td>CD</td>
<td>FTIR</td>
<td>188</td>
</tr>
<tr>
<td>GR-AF1</td>
<td>Buffer</td>
<td>26.7</td>
<td>11.6</td>
<td>12.3</td>
<td>10</td>
<td>39.4</td>
<td>FTIR</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td>Buffer</td>
<td>26</td>
<td>13</td>
<td></td>
<td>26</td>
<td>CD</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P-serine</td>
<td>42.1</td>
<td>11.8</td>
<td>12.4</td>
<td>9.3</td>
<td>24.4</td>
<td>FTIR</td>
<td>243</td>
</tr>
<tr>
<td></td>
<td>MR-AF1a</td>
<td>Buffer</td>
<td>11</td>
<td>29</td>
<td>25</td>
<td>36</td>
<td>CD</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>63</td>
<td>5</td>
<td>10</td>
<td>22</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>MR-MD</td>
<td>Buffer</td>
<td>13</td>
<td>22</td>
<td>24</td>
<td>40</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>66</td>
<td>13</td>
<td>7</td>
<td>15</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>MR-AF1b</td>
<td>Buffer</td>
<td>19</td>
<td>32</td>
<td>31</td>
<td>19</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>13</td>
<td>31</td>
<td>23</td>
<td>33</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>MR-NTD</td>
<td>Buffer</td>
<td>20</td>
<td>21</td>
<td>24</td>
<td>35</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>79</td>
<td>2</td>
<td>3</td>
<td>15</td>
<td>CD</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>PR-AnDBD</td>
<td>Buffer</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td>CD</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ JPD2</td>
<td>58</td>
<td>11</td>
<td></td>
<td></td>
<td>CD</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>Ecr-NTD</td>
<td>Buffer</td>
<td>8</td>
<td>31</td>
<td>23</td>
<td>89</td>
<td>CD</td>
<td>244</td>
<td></td>
</tr>
<tr>
<td>dHR38-NTD</td>
<td>Buffer</td>
<td>11.1</td>
<td>11.7</td>
<td>10.3</td>
<td>66.6</td>
<td>CD</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TFE</td>
<td>72.3</td>
<td>4.4</td>
<td>8.7</td>
<td>17.8</td>
<td>CD</td>
<td>245</td>
<td></td>
</tr>
</tbody>
</table>

JPD2, Jun-protein dimerization 2; P-serine, phosphorylated serine(s); TMAO, trimethylamine N oxide; Ecr, ecdysone receptor; CD, circular dichroism; FTIR, Fourier transform infrared spectroscopy.
ies suggest that negative charges per se are not sufficient for AF1 activity; indeed, key hydrophobic amino acids may be crucial for AF1 activity (124, 125). The pattern of proteolytic degradation products in cell-free extracts suggested that the GR-AF1 might be structured in vivo (126). Furthermore, helix-destroying mutations of amino acids in three hypothetical helices of the GR-AF1 strongly reduced GR transcriptional activity (172, 191). Thus, it could be hypothesized that the GR-AF1 may be structured in vivo, at least when directly involved in transcriptional activation. Subsequent studies with other SHRs such as AR, PR, ER, and MR showed similar trends (135, 174–176, 189). In recent years, it has been shown that the NTD/AF1 of SHRs undergo coupled binding/folding events. It has been well established that like many other transcription factors, during transcription regulation the NTD/AF1 shifts to a conformational space in which more structured conformers exist (Table 1 and Fig. 4).

The crowded conditions inside the cell have been suggested to cause ID proteins to fold into an ordered three-dimensional structure. Molecular crowding in cells may include small molecule solutes and macromolecules including proteins, DNA, and RNA. Because SHRs are known to function in a cell/tissue-specific manner and AF1-mediated transcriptional activity is particularly sensitive to cellular environment (see Section IV), it can be hypothesized that one of the reasons for cell-specific effects of AF1 may be influenced by binding/folding events governed by molecular crowding in specific cells. Several mechanisms important in inducing ordered conformations of the NTD/AF1 under physiological conditions have been described, which in turn may facilitate receptor interaction with specific coactivators and subsequent transcriptional activity.

1. Role of small molecules (osmolytes)

Functionally folded structures of most macromolecules are sensitive to changes in cellular environmental conditions. The native functional activities and stability of macromolecules are fine-tuned through accumulation of high concentrations of small organic molecules known as osmolytes or chemical chaperones. Cells regulate many biological processes such as protein folding and protein-protein interactions via accumulation of...
specific osmolytes (190, 192–195). In recent years, the mechanism of osmolyte compatibility and osmolyte-induced stability has attracted considerable attention. Osmolytes that occur naturally represent different chemical classes such as amino acids (proline and glycine), methylamines (betaine and trimethylamine-N-oxide), and polyols and sugars (sorbitol and trehalose) (190). Because osmolytes predominantly affect the protein backbone, the balance between osmolyte-backbone interactions and amino acid side chain-solvent interactions determines protein folding (192–195). In ID proteins, the hydrophobic amino acids are insufficient, relative to the charged side chains, to cause spontaneous folding (156, 188, 205–207). Addition of the solvophobic effect of a profolding osmolyte can tip the energy balance in favor of folding because the backbone collapses to avoid solvent interactions with the osmolyte (192–195). The quantity of osmolyte required depends on both its inherent solvophobic properties in the interaction with the peptide backbone and the free energy balance provided by the sum of all backbone-osmolyte interactions and the sum of all amino acid side chain-solvent interactions (192–195).

The ID NTD of SHRs, or isolated AF1 regions, can be induced to fold spontaneously to native, functional forms by addition of certain organic osmolytes to the solvent (135, 173, 176, 196, 197) (Table 1 and Fig. 4). Incubation with several osmolytes causes the ID AF1 domain of the GR to fold into a form that can bind strongly to specific coregulatory proteins that are critical for the transcriptional activity of the receptor (196). The folding curves were cooperative and typical of a protein that folds spontaneously via the classic “two-state” model. This strongly suggests that a natural conformation had been reached (196). Osmolytes also reduced limited proteolysis of the GR-AF1 polypeptide and resulted in the tryptophan and tyrosine residues becoming less solvent exposed, which is consistent with folding or increased conformational stability of the receptor domain (196). Similarly, the ID AF1 domain of the AR has been shown to have the propensity to form helical structures in the presence of the osmolyte, trimethylamine N-oxide (188), and to adopt a more folded conformation that was resistant to limited proteolysis (176, 189) (Table 1). Furthermore, this osmolyte-induced conformation in the AR-AF1 domain significantly facilitates its interaction with a critical coregulatory protein (188, 189). Trimethylamine N-oxide has also been shown to stabilize a more folded conformation for regions of the MR-NTD and, in the case of central 139 amino acids, increase binding of a number of coactivator and corepressor proteins (Table 1) (135).

Recent studies suggest that trehalose-induced folding of the GR-NTD facilitates binding to a glucocorticoid response element (GRE) in the context of a two-domain GR polypeptide containing the entire NTD plus the DBD, termed GR500 (190). It has been previously reported that GR500 is capable of binding to GRE and can stimulate AF1-mediated GR activity similar to that observed with the hormone-bound full-length receptor. Taken together, the above results suggest that osmolyte-induced folding of the NTD or AF1 of SHRs may be important for the interaction of the receptor protein with both coregulatory proteins and response element DNA.

2. Binding of DNA and protein partners

According to the classic model of SHR action (Fig. 1), receptor bound to specific hormone response element (HRE) by virtue of high-affinity interactions modifies chromatin structure and/or contacts the multiprotein transcription machinery to regulate transcription from a target promoter (59, 95, 119, 120, 230).

a. DNA architecture and SHR-NTD folding. In recent years, it has become evident that SHR:HRE binding also influences the three-dimensional configuration of the receptor (95). Consequently, in an HRE-specific manner, through conformational rearrangements, the surfaces of the SHRs are modified such that various critical ancillary factors can bind (95). Furthermore, DNA-binding can also be a trigger for an active intramolecular communication that can change the conformation of the SHRs in a site-specific manner (199). Because transcriptional control of a specific gene depends upon the interactions of the SHR with DNA/chromatin, the exact DNA sequences and architecture of the response elements in the regulatory regions of the gene could help determine the hormone response. Specific sequences within estrogen response element (ERE) have been shown to modulate the conformation of ER (200, 201), and conformational changes resulted in altered recruitment of coactivator proteins to the ER:ERE complex (199). This provides an intellectually satisfying rationale for the varied transcriptional effects of specific sequences found in the SHR genomic binding sites controlling specific genes. Furthermore, as discussed in Section III, recent HDX experiments demonstrated that binding of DNA response elements led to changes in the dimerization and AF2 surfaces of each partner in a heterodimer receptor complex (138). However, it also raises the question: what are the regions/domains of the receptor that may be most prone to these conformational changes due to HRE binding? Several studies have shown that dramatic folding effects have been seen on the AF1/NTD of SHRs when DNA binding takes place (174, 183, 184, 201). Moderate conformational changes have also been observed in other domains, namely, the DBD
Several coregulatory protein binding and SHR-NTD folding. Various contexts.

In order to understand receptor-mediated transcription in the functional structure of the SHR necessary for positive regulation of transcription depends on receptor-DNA binding and DNA architecture (Fig. 4). Thus, part of the functional structure of the SHR necessary for positive regulation of transcription depends on receptor-DNA binding. This structure may differ from the conformation at genes negatively regulated by SHRs, where the receptor is not necessarily bound to DNA—for example, receptor interactions with other transcription factors to repress the expression of target genes. At such sites, the SHR is tethered to DNA indirectly via binding to a heterologous transcription factor. For example, the effect of estrogen at the pS2 gene is mediated through a cross talk between the ER, ERE, and an activator protein 1 response element (202).

More extensive analyses are needed to reveal all the structure-function variations in the SHR and its NTD/AF1 in order to understand receptor-mediated transcription in various contexts.

b. Coregulatory protein binding and SHR-NTD folding. Several coregulatory binding partner proteins are involved in the action of the SHR on target gene transcription (49, 119, 203, 204). There are several examples in which binding sites within the ID regions of many transcription factors contain molecular recognition features, which consist of short stretches of amino acid sequences that undergo a disorder-to-order transition and are stabilized by binding to a partner protein(s) that act as coregulators (153, 156, 188, 205–207). Thermodynamics predicts that if one or more conformers of an ID protein/region bind specific proteins with high affinity, interaction of the ID region with that partner at appropriate concentrations may cause its structure to stabilize (208, 209). It has been predicted that these segments may have advantages for cell signaling by allowing the decoupling of partner binding specificity and affinity, due to which the strength and duration of signaling events can evolve (208, 209).

Coregulatory proteins are known to modulate the transcriptional activity of the SHRs through multiple mechanisms. The SHR:coregulator complexes can act to modify chromatin, influence RNA polymerase II phosphorylation, and bind mediators and proteins of the basal transcription complex. Many of the known key SHR coregulators bind to both the NTD/AF1 and AF2, using distinct receptor binding sites (210–213). In recent years, the mechanism of action and the structure of AF2:coregulator interaction and its effects on the regulation of SHR target genes have been explained in depth (5, 45, 46, 63, 231). However, how the choice of AF1:coregulator interaction is made and its functional significance in gene regulation by SHRs remain poorly understood. There are examples in which the ID regions of a transcription factor protein take full shape upon interaction with protein binding partner(s) through an induced-fit model of folding (154, 214). Based on this model, it has been hypothesized that NTD/AF1 is not fully structured in vivo until it binds to one or more coregulatory protein(s), and this induced conformation or limited set of conformations in NTD/AF1 is a prerequisite for its interaction with specific sets of other coregulatory proteins in a cell- and promoter-specific manner.

The NTD/AF1 of SHRs is known to recruit proteins from the basal transcriptional machinery. For example, the TATA box binding protein (TBP) can directly bind to the NTD/AF1 domain of several SHRs (135, 175, 182), and RAP74, a subunit of the transcription factor IIF (TFIIF) complex, interacts with the AR AF1 domain (179, 215). Significantly, the TBP binding induces secondary/tertiary structure formation in the GR-AF1 (182) (Table 1) and ERα-NTD (175). Using several biophysical and partial proteolytic digestion experiments, the GR-AF1 has been shown to assume a three-dimensional fold with significant helical content upon interaction with TBP (182). The TBP binding-induced folding in the GR-AF1 significantly enhances its interaction with SRC-1, and subsequent AF1-mediated, GRE-driven promoter-reporter ac-
tivity (182). Similar studies have been reported for the ER-NTD (175). TBP has a central role in the basal transcription machinery, and it is interesting to note that it directly binds to the NTD of the ERα but failed to bind ERβ-NTD (175). This difference in TBP binding implies differential recruitment of target proteins by the NTD of ERα and ERβ. The affinity of the ERα and GR-AF1/NTD: TBP interaction was determined to be in the micromolar range, as assessed by surface plasmon resonance spectroscopy (175, 182). Based on these results, it has been proposed that the interaction between the AF1/NTD and TBP may proceed in a two-step manner with initial very fast, low-affinity association, followed by a slow, folding event and tighter association (175, 182). The initial association may occur by electrostatic interactions between the acidic residues of highly negatively charged AF1/NTD and the positively charged TBP. However, this initial unstable protein complex subsequently may convert into a more stable form by the folding of the ID AF1/NTD and the formation of specific contacts between the two proteins (216–218).

Furthermore, the dissociation of this binding interaction suggested a complex behavior, with a rapid dissociation for AF1/NTD molecules that did not undergo proper folding and a slower dissociation for those molecules that did fold successfully upon physical interaction with the TBP (175). Such a two-step binding mechanism is consistent with the change in protein conformation that accompanies the AF1/NTD:TBP interaction. Based on the binding and consequent folding of the AF1/NTD, it can be hypothesized that the interaction between SHR-NTD/AF1 domains and protein(s) from basal transcription machinery may be a unified mechanism, through which these ID AF1/NTD and protein(s) from basal transcription machinery bind to different coregulatory proteins are solved and can be compared.

In addition to DNA binding and dimerization, the DBD has also been highlighted as a site for protein-protein interactions and may act as a hub for the transmission of allosteric signals as a result of receptor binding of both hormone and DNA. Significantly, binding of coregulatory proteins to the PR-DBD CTE has been described, which results in allosteric regulation and folding of the PR-NTD (97, 106, 219) (Table 1).

3. Posttranslational modifications and SHR-NTD folding

Posttranslational modification is an important phenomenon that controls the functions of transcription factors in cells by regulating their DNA-binding affinity, interaction with components of the transcription initiation complex, and the shuttling between the cellular compartments (220). The SHRs are the target for a plethora of posttranslational modifications, including phosphorylation, acetylation, methylation, and sumoylation, and these modifications have a complex effect on SHR signaling (221). In some cases, one type of posttranslational modification can influence another type (198, 222). Acetylation of the lysine residues in the AR hinge domain plays an important role in AR-mediated gene regulation, and mutation of these residues can impair phosphorylation of serine 94 in the NTD (103, 198). Similarly, methylation of these residues enhances AR-dependent transactivation by increasing N/C-terminal interactions and recruitment of the receptor to target genes (104, 223). The interplay of phosphorylation and sumoylation in the regulation of ERR-NTD transcriptional activity has also been observed, and these modifications form a conserved phosphosumoyl switch that exists within a larger synergy control motif (224).

Recent studies have suggested that signaling via phosphorylation-regulated protein-protein interaction often involves ID regions, and these regions have a much higher frequency of known phosphorylation sites than ordered regions, suggesting a strong preference for locating phosphorylation sites in the ID regions (153). One of the main reasons for such propensity is to facilitate extensive formation of hydrogen bonding between the backbones and/or side chains that can occur through disorder-order transition within the ID region. The formation of these hydrogen bonds would be difficult if the sites of phosphorylation were located within ordered regions (153, 209). In terms of structural consequences of site-specific phosphorylation, both disorder-to-order and order-to-disorder conformational transitions have been ob-
served to follow the phosphorylation event, and these conformational changes often affect protein function (153). Lack of well-defined structure is generally associated with specific sites of enzyme-catalyzed phosphorylation and ubiquitination sites (209), which facilitates the inclusion or exclusion of specific binding partners in the complex (209). Because a significant part of the binding energy is required to fold an ID region, high specificity coupled with low affinity provides the basis for easily reversible interactions (209).

Like many other transcription factors, the SHRs are phosphoproteins, and it has been suggested that kinase-mediated phosphorylation plays an important role in the regulation of their activities (225). Phosphorylation of specific residues within the NTD/AF1 of the GR and ER regulates the transcriptional activity of these receptors (226–228). There are also reports suggesting that phosphorylation may affect SHR stability and thus alter receptor activity (180). Analysis of known phosphorylated residues within the AR-, ERα-, GR-, and PR-NTD with predicted ID reveals a strong correlation (Table 2). In the case of the ERβ and the GR, site-specific phosphorylation has been associated with conformational changes in the ID AF1 region (180, 229). Based on a simulation modeling approach, it was proposed that phosphorylation of serine 211 in the human GR-AF1 may result in the formation of more compact structure in the surrounding peptide within the AF1 (amino acids 187–244), which may expose novel protein surface for cofactor interaction (228). Furthermore, it was shown that common surfaces within GR affected by phosphorylation may be responsible for influencing the regulation of selective genes (228). More recent studies show that when GR-AF1 is phosphorylated at serine 211, using the p38 MAPK, secondary/tertiary structure is formed in the AF1 domain (Table 1), suggesting that under physiological conditions, site-specific phosphorylation may play a crucial role in allowing the ID AF1 domain of the GR to adopt functionally active conformation(s) (180). Furthermore, it was found that the resulting structurally modified forms of AF1 facilitated interactions with critical coregulatory proteins (CBP, TBP, and SRC-1), and possibly additional factors, resulting in the assembly of multiprotein complexes involved in GR-mediated regulation of transcription (180).

These findings demonstrate a mechanism through which the ID NTD of SHRs may adopt a functionally active conformation under physiological conditions in response to phosphorylation. More generally, it can be concluded that the SHR-NTD may be structured in vivo, at least when phosphorylated and/or directly involved in interactions with DNA and binding partner proteins (Fig. 4).

### VI. Intrinsic Disorder Can Optimize Allosteric Coupling

Allosteric coupling is a common phenomenon associated with proteins with modular structures (230) and can be defined by propagation of conformational perturbation in one domain/region of the protein due to signals passed from distant sites in the same molecule. As discussed above, due to their modular structures and regulation through specific ligands interacting with different domains (i.e., coregulatory proteins for NTD and LBD, DNA for DBD, and steroid hormone for LBD), SHRs are tightly regulated through allosteric coupling (230, 231). Each of these interactions can act as signals that modulate interactions with other molecules and/or establish intramolecular and interdomain communications (Fig. 5). As a result, SHRs can switch from one functional state to another by selective stabilization of different conformations, and the ID NTD/AF1 may be the region that undergoes most sig-

### TABLE 2. Location of known phosphorylated residues in regions of predicted intrinsic disorder

<table>
<thead>
<tr>
<th>Receptor</th>
<th>Residue</th>
<th>PONDR</th>
<th>RONN</th>
<th>GlobPlot</th>
<th>Fold index</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR-NTD</td>
<td>S16</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S81</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S94</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S210</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S256</td>
<td>Border</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S308</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S424</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S515</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERα-NTD</td>
<td>S106</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S118</td>
<td>D</td>
<td>Border</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S167</td>
<td>D</td>
<td>D</td>
<td>Border</td>
<td>D</td>
</tr>
<tr>
<td>GR-NTD</td>
<td>S134</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>Border</td>
</tr>
<tr>
<td></td>
<td>S203</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S211</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S226</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>PR-NTD</td>
<td>S404</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S20</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S25</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S81</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S102</td>
<td>D</td>
<td>D</td>
<td>Border</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S130</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S162</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S190</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S213</td>
<td>D</td>
<td>D</td>
<td>Border</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S294</td>
<td>D</td>
<td>Border</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S345</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S400</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>T430</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>S554</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>

D, Intrinsic disorder; Border, residue is at the border of region of predicted intrinsic disorder. ID was predicted from the primary amino acid sequences using the algorithms PONDR (247), RONN (248), GlobPlot (249), and fold index (250).
significant structural rearrangement and so plays the major role in the transmission of information to distal sites.

In the above description (Fig. 5) allosteric coupling is proposed to induce an ordered conformation in the ID domain through domain-domain interactions, and structure formation in these ID domains/regions is linked to ligand binding either in the ID domain itself or in other domains of the molecule. Thus, the disorder/order transition is coupled to long-range allosteric communication within the molecule and is therefore important for its function (232). Significantly, it has recently been argued that an additional functional consequence of ID is to maximize allosteric interactions (232). This has led to a new model whereby allosteric coupling does not necessarily require a conformational link between the two sites, but is rather a consequence of the “energetic balance” and overall stability of the protein (232). In this manner, SHRs sample various conformations and are primed to respond to changes in different intracellular compartments. It is interesting therefore that the removal of the NTD-flanking sequences on both sides of the GR-AF1, by deleting amino acids, that produces a peptide with AF1 immediately upstream from the DBD leads to acquisition of an ordered conformation in AF1. In this construct, interaction of AF1 with CBP/p300, but not SRC-1, is significantly increased and concomitantly, the GRE-mediated AF1 activity (233). Thus, proximity to the DBD or loss of flanking sequences enhances certain folding properties of AF1 and subsequent activity, suggesting that these flanking sequences play an inhibitory role in regulating GR-AF1 structure and activity (233). Unlike, the two-domain GR fragment (GR500), binding of this peptide to a GRE fails to induce any further structure formation in AF1 (233). These position-specific structural effects suggest that AF1 structure (or propensity for structure formation) is strongly influenced by the amino acids that surround it as well as by influences from the DBD:GRE interaction. Strikingly, these findings also support the process of allosteric coupling through an intrinsic disorder mechanism, and the NTD amino acid sequences that flank AF1 must influence these signaling processes (232). These results also provide a mechanistic explanation for why certain N-terminally
truncated GR, and possibly other SHRs, splice variants differ in their activities (10).

Allosteric mechanisms have evolved in many transcription factor proteins, and this flexibility may allow SHRs to mediate diverse genetic responses directly by modifying their malleable surfaces for specific interactions. Identification of different ligand binding events and subsequent allosteric coupling is likely to play an important role in understanding the diverse target function of SHRs and could be critical in developing improved therapeutic interventions in several pathological conditions associated with SHRs. In cells, SHR function in a very dynamic fashion, existing as an ensemble of conformations that can sample the cellular environment in a rapid manner, and subsets of these conformations can be selected and stabilized by hormone and DNA binding.

VII. Structural Dynamics and the Design of Selective Small Molecule Modulators of Steroid Hormone Receptor Function

It is generally accepted that receptor function is dependent upon ligand-induced conformational changes (i.e., helix 12 orientation), together with the cellular complement and levels of coregulatory proteins. As discussed in Section V, it seems certain that the DNA architecture of receptor binding sites and the folding of the AF1/NTD domain will also contribute to cell-specific receptor actions. SHRs are well-validated drug targets and are prognostic indicators in a number of pathophysiological conditions, including inflammation, hormone-dependent cancers, osteoporosis, and cardiovascular disease (see Section I). Classical antagonists of SHRs act as competitive inhibitors of the natural hormone by binding to the ligand binding pocket and inducing a conformation that is unable to bind coactivator proteins, or retains the receptor complex in the cytosol, and/or targets the receptor for degradation. However, there is increasing interest in the development and characterization of compounds that exhibit mixed agonist/antagonist properties depending on the cellular context; these are best described as “selective receptor modulators.” The “Holy Grail” for hormone therapies is to be able to target SHR signaling in a receptor- and tissue-selective manner to maximize therapeutic benefits and to minimize deleterious outcomes in other receptor target tissues.

The general problem with traditional antagonists or SRMs is the development of resistance and/or off-target side effects. It therefore makes sense to investigate the possibility of identifying inhibitors that act outside of the ligand binding pocket, which could complement or replace existing SRMs. A number of groups have screened different classes of molecules (natural products, biologics, or synthetic compounds) for such inhibitors, and a number of exciting and surprising outcomes have been reported. For example, there are several examples where SRM binding influences the conformation of the receptor (234, 235), and these conformational changes collectively determine the specific biological activity. Thus, a number of peptide probes that recognize different conformations of a SHR bound to different ligands have been identified. Although these peptide libraries have proved their worth in characterizing ligand-binding profiles, initially for the ERα and ERβ, they also open the possibility for developing selective receptor peptide inhibitors (for example, Refs. 236–238) that will selectively disrupt receptor-SRM-coregulatory protein complexes in a tissue-specific manner.

The potential for allosteric inhibitors of receptor function was emphasized by the finding that a number of small molecules (e.g., flufenamic acid, 3,3',5-triiodothyroacetic acid) disrupted protein-protein interactions involving AR-AF2 and inhibited AR-dependent transactivation (74). The surprise came when x-ray structures revealed these molecules bound to a hydrophobic surface of the LBD, termed BF3 (involving residues in helix 1, helices 3 to 5, a loop, and helix 9), which is distinct from the ligand binding pocket, and crucially the AF2 region (74). This finding, together with the information from peptide binding studies, suggests that other surfaces of the LBD may participate in protein-protein interactions and could therefore be novel drug targets. Allosteric regulation could also stabilize a specific conformer that reduces the affinity of the second binding site for its ligand, and this dynamic could be explored as a novel approach to structure-based SRM design to produce differential responses.

In a separate study, a number of hits were identified from a screen of two small-molecule libraries that inhibited ER-dependent DNA binding (239). One of these molecules, theophylline, 8-[(benzylthio)ethyl]-(7Cl,8Cl), selectively inhibited ERα-dependent transcription and acted outside of the ligand binding pocket. Similarly, two compounds, pyrvinium pamoate and a natural product, harmol hydrochloride, have been described that inhibit AR activity, independent of hormone binding (240). Harmol appeared to disrupt DNA binding, whereas pyrvinium acted at a subsequent step in receptor signaling, possibly disrupting protein-protein interactions (240). Although the binding sites on the ER or AR for these small molecules have yet to be identified, they illustrate the possibilities for developing inhibitors that target other domains of SHRs.

The structural flexibility of the SHR-AF1/NTD may have appeared to be a rather unattractive target for small-
molecule inhibitors of receptor function. However, the uniqueness and functional importance of the NTD makes it a worthwhile, even if technically challenging, drug target. It is therefore highly significant that a small molecule, identified in a screen of compounds isolated from a marine sponge, was recently shown to be highly effective and specific at inhibiting AR activity in vitro and in vivo (241). The compound, termed EPI-001, is a bisphenol A ester derivative (BADGE), and it interacted with the AR-AF1 domain, differentially disrupted protein-protein interactions, and reduced binding to androgen-regulated genes (241). Overall, this study provides “proof-of-principle” that small-molecule inhibitors can be identified that target the structurally plastic SHR-NTD.

Traditionally, it has been difficult to design or identify small molecules that target protein-protein interfaces, given the hydrophobic nature and relatively large surface area of such sites. However, the above-mentioned studies would challenge this preconception and open up the potential to develop tissue-specific modulators of SHR action that would complement or replace existing therapeutics.

VIII. Conclusions and Future Challenges

The last 25 yr have seen tremendous progress in our understanding of the structure-function relationships of members of the nuclear receptor superfamily, including SHRs. This has included the isolation of receptor cDNA; the solving, at atomic resolution, of the structures of the isolated DBD and LBD; the identification of binding partners; and a clearer appreciation of the role of ID and allosteric coupling in receptor action. The future will prove to be just as challenging and exciting, and we may reasonably expect progress in a number of areas.

Solving the three-dimensional structure of a full-length SHR has proved a major technical challenge. However, our increased understanding of the functional and structural properties of the ID NTD, together with the allosteric coupling through hormone and DNA binding, opens up the possibility of investigating the conformation of a SHR-DNA-coregulatory protein(s) complex. In the absence of a structure for a full-length SHR complex, it will be important at the very least to obtain structural information, perhaps through small angle x-rays or DHX experiments, for a two-domain SHR in the presence and absence of DNA binding site and protein binding partners. Such studies will be invaluable if we are to further understand the structural basis for SHR-dependent gene regulation and allosteric control. Such information will also be vital to our understanding of the impact of point mutations in receptor-dependent diseases and in the design of novel inhibitors and SRM ligands to modulate tissue-specific SHR action.

In conclusion, SHRs are highly modular proteins, and this property has been exploited in the regulation of these molecules by the binding of different ligands (hormone, DNA, coregulatory proteins). The result is the exquisite control of normal physiology by steroid hormones and the serious consequences that arise through misregulation in the pathophysiology of certain cancers, inflammatory disease, and cardiovascular disorders. It has become increasingly clear that the NTD, containing the major transactivation function AF1, is naturally ID, and that this structural plasticity is functionally important for SHR activity. The SHR-NTD exists as an ensemble of conformers, having more or less structure, which prime this region of the receptor to rapidly respond to changes in the intracellular environment through hormone binding and postranslation modifications. It is interesting that the NTD itself is often highly modular, with multiple regions contributing to receptor-dependent gene regulation. This seems particularly the case for the SHRs with the largest NTD: AR, PR, and MR. The consequences of this are unclear at present, but it may be significant that the length of the NTD has been correlated with AF1 “strength” in a reporter gene assay. A picture is emerging where the induced folding and/or stabilization of a functional conformation in the NTD is achieved through coregulatory protein binding and/or allosteric coupling with the DBD and DNA response element binding, and we are beginning to understand how different surfaces within the SHR protein may be created and used to control target gene expression.

Acknowledgments

Address all correspondence and requests for reprints to: Iain J. McEwan, School of Medical Sciences, University of Aberdeen, Forsterhill, Aberdeen AB25 2ZD, Scotland, United Kingdom. E-mail: iain.mcewan@abdn.ac.uk. Or, Raj Kumar, Department of Basic Sciences, The Commonwealth Medical College, Scranton, Pennsylvania 18510. E-mail: rkumar@tcmedc.org.

Work in the authors’ laboratories is supported by National Institutes of Health Grant R01DK058829 (to R.K.) and The Chief Scientist Office, Scottish Government, Prostate Cancer Charity (United Kingdom) and The Kosterlitz Center, University of Aberdeen (to I.J.M.).

Disclosure Summary: The authors have nothing to disclose.

References


42. Berry M, Metzger D, Chambon P 1990 Role of the two activating domains of the oestrogen receptor in the cell-type and promoter-context dependent agonistic activity of the anti-oestrogen 4-hydroxysteramifinen. EMBO J 9:2811–2818
51. McEwan IJ, Lavery D, Fischer K, Watt K 2007 Natural disordered sequences in the amino terminal domain of nuclear receptors: lessons from the androgen and glucocorticoid receptors. Nucl Recept Signal 5:e001


71. Tzukerman MT, Esty A, Santiso-Mere D, Danielian P, Parker MG, Stein RB, Pike JW, McDonnell DP 1994 Human estrogen receptor transactivational capacity is determined by both cellular and promoter context and mediated by two functionally distinct intramolecular regions. Mol Endocrinol 8:21–30


78. Centenera MM, Harris JM, Tilley WD, Butler LM 2008 The contribution of different androgen receptor domains to receptor dimerization and signaling. Mol Endocrinol 22:2373–2382


85. Metivier R, Penot G, Flourirot G, Pakdel F 2001 Synergism between ERα transactivation function 1 (AF-1) and AF-2 mediated by steroid receptor coactivator protein-1: requirement for the AF-1 α-helical core and for a direct interaction between the N- and C-terminal domains. Mol Endocrinol 15:1953–1970


88. Dong X, Challis JR, Lye SJ 2004 Intramolecular interactions between the AF3 domain and the C-terminus of the human progesterone receptor are mediated through two LXXLL motifs. J Mol Endocrinol 32:843–857


95. Meising SH, Pufall MA, So AY, Bates DL, Chen L, Yamamoto KR 2009 DNA binding site sequence directs
glucocorticoid receptor structure and activity. Science 324: 407–410
115. Massie CE, Adrian B, Barbosa-Morais NL, Lynch AG, Tran MG, Neal DE, Mills IG 2007 New androgen receptor genomic targets show an interaction with the ETS1 transcription factor. EMBO Rep 8:871–878
123. Dahlman-Wright K, Almlöf T, McEwan IJ, Gustafsson JA,
Wright AP 1994 Delineation of a small region within the major transactivation domain of the human glucocorticoid receptor that mediates transactivation of gene expression. Proc Natl Acad Sci USA 91:1619–1623


125. Almlöf T, Wallberg AE, Gustafsson JA, Wright AP 1998 Role of important hydrophobic amino acids in the interaction between the glucocorticoid receptor τ 1-core activation domain and target factors. Biochemistry 37:9386–9594


form, BAG-1L, enhances oestrogen-dependent transcription. Oncogene 22:4973–4982


188. Kumar R, Betney R, Li J, Thompson EB, McEwan IJ 2004 Induced α-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. Biochemistry 43: 3008–3013


193. Street TO, Bolten DW, Rose GD 2006 A molecular mechanism for osmolyte-induced protein stability. Proc Natl Acad Sci USA 103:13997–14002


196. Kumar R, Serrette JM, Khan SH, Miller AL, Thompson EB 2007 Effects of different osmolytes on the induced folding of the N-terminal activation domain (AF1) of the glucocorticoid receptor. Arch Biochem Biophys 465:452–460


211. Benecke A, Champon P, Gronemeyer H 2000 Synergy between estrogen receptor α activation functions AF1 and AF2 mediated by transcription intermediary factor TIF2. EMBO Rep 1:151–157

212. Onate SA, Boonyaratanaokornkit V, Spencer TE, Tsai SY, Tsai MJ, Edwards DP, O’Malley BW 1998 The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF2 domains of steroid receptors. J Biol Chem 273:12101–12108


216. Shoemaker BA, Portman JJ, Wolynes PG 2000 Speeding
molecular recognition by using the folding funnel: the fly-casting mechanism. Proc Natl Acad Sci USA 97:8868–8873


220. Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE 2002 Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. J Biol Chem 277:22424–22428


233. Kumar R, Thompson EB 2010 Influence of flanking sequences on signaling between the activation function AF1 and DNA-binding domain of the glucocorticoid receptor. Arch Biochem Biophys 496:140–145


244. Nocula- żogowska M, Rymarczyk G, Lisowski M, Ozzyar A 2009 Isoform-specific variation in the intrinsic disorder
of the ecdysteroid receptor N-terminal domain. Proteins 76:291–308


Save the Date for Endocrine Board Review Course,
September 11-12, 2012, Miami, Florida.
www.endo-society.org/CEU