From Antibodies to Crystals: Understanding the Structure of the Glucocorticoid Receptor and Related Proteins

Iain J. McEwan

Institute of Medical Sciences, University of Aberdeen, Foresterhill Health Campus, Aberdeen AB25 2ZD, Scotland, UK; iain.mcewan@abdn.ac.uk

Abstract: The steroid/thyroid hormone or nuclear receptor superfamily is quickly approaching its 40th anniversary. During this period, we have seen tremendous progress being made in our understanding of the mechanisms of action of these physiologically important proteins in the field of health and disease. Critical to this has been the insight provided by ever more detailed structural examination of nuclear receptor proteins and the complexes they are responsible for assembling on DNA. In this article, I will focus on the contributions made by Jan-Åke Gustafsson and colleagues at the Karolinska Institute (Sweden) and, more recently, the University of Houston (USA), to this area of nuclear receptor research.

Keywords: glucocorticoid receptor; NMR spectroscopy; X-ray crystallography; cryo-EM; liver X receptor

1. Introduction

A full understanding of the molecular basis of living processes—from reproduction to metabolism to how we age—requires knowledge of protein structures and the mechanistic insights this provides. The development of the disease and drug discovery field is also strongly dependent on knowledge of the three-dimensional shape of proteins, and nowhere is this truer than for the nuclear receptor superfamily [1]. The human genome codes for 48 members of this family of ligand-activated transcription factors. They play central roles from the very beginning of life through development (retinoic acid, vitamin D3, thyroid hormones) and sexual differentiation (sex steroid hormones) to how we, as adults, digest our food (bile acids), regulate metabolism (steroid and thyroid hormones) and maintain healthy bones and muscle strength (steroid hormones).

Over the last half-century, our understanding of steroid hormone action and receptor structure has evolved dramatically. Starting with the recognition that these lipophilic molecules could directly regulate gene transcription through binding intracellular receptor proteins, our knowledge advanced until reaching the acceptance of the two-step model of receptor action proposed by Elwood Jensen in the early 1960s. The structural analysis of isolated receptor domains, and, more recently, full-length proteins, has provided valuable insights into the roles which hormone and DNA response elements play in allosteric regulation, as well as mechanisms for receptor dimerization and co-regulatory protein recruitment (for reviews, see [2,3]).

For more than four decades, Professor Jan-Åke Gustafsson and colleagues have made seminal discoveries that have advanced our appreciation of how nuclear receptor architecture underpins mechanisms of hormone action in health and disease. In this review, I will look back at how our structural understanding of steroid and related receptors has progressed from biochemical models to atomic resolution structures, with a specific emphasis on the contributions from the Gustafsson laboratory.

2. Early Models of Steroid Receptor Structure and Function

The principal circulating glucocorticoid in humans is the hormone cortisol, which was first isolated by Kendall and Hensch in 1946 [4]. Forty years later, the cDNA for...
the human glucocorticoid receptor (GR) was cloned by the Evans and Chambon groups independently [5–7], and the cDNA for the rat receptor was cloned by the Gustafsson and Yamamoto groups, working together [8]. The glucocorticoid receptor is the protein that binds cortisol (in humans) and mediates the actions of the hormone in a cell- and tissue-selective manner. The key to studying receptor function, prior to the availability of the cDNA, was the purification of the receptor from target tissues, for example, rat liver. This work was greatly facilitated by radiolabeled forms of natural and synthetic glucocorticoids and the generation of receptor-specific antibodies. Combining purification with selective partial proteolysis and a panel of antibodies, Gustafsson and co-workers proposed a three-domain model for the GR: functional domains that bound steroids and DNA, and an immune modulatory domain (IMM) (Figure 1) [9–11]. Chymotrypsin’s digestion of the 94 kDa purified GR resulted in a large fragment that bound both steroids and DNA (A+B domains), and a fragment recognized by a panel of monoclonal antibodies raised against the full-length protein (Domain C) (Figure 1).

![Figure 1. Schematic representation of the glucocorticoid receptor purified from a rat liver. Domain organization of the rat GR based on steroid (A), DNA (B), and immunoglobulin (C) binding and partial proteolysis with trypsin or α-chymotrypsin.](image)

Further digestion with trypsin released a 25 to 27 kDa fragment that retained steroid binding (Domain A). Concomitant with these biochemical studies, the purified GR was shown to bind to discrete regions upstream of the start site of transcription of the mouse mammary tumor virus (MMTV) DNA [12]. Importantly, having antibodies and partial amino acid sequences ultimately led to the cloning of the cDNA for the rat GR [8]. The function of domain C was initially unclear, but thought to be important for GR activity; the IMM region or domain C is now recognized as the amino-terminal domain (NTD), and in the case of the GR and other steroid receptors, it is important for gene regulation (see [13]).

3. Structure of the Glucocorticoid Receptor

3.1. GR-DBD

The initial biochemically-derived domain structure for the GR (Figure 1) was subsequently validated and expanded upon by deletion and point mutation studies, as well as domain-swapping experiments exploiting the available receptor cDNAs (Figures 2 and 3) (see [14–20]). The cloning of steroid hormone receptor cDNA was also an important driver for the identification of related proteins and the concept of the nuclear receptor superfamily [1]. In addition to increasingly detailed functional analysis, the receptor cDNA opened up new avenues for recombinant protein expression in bacteria, yeast, and insect cells. The ability to produce and purify large quantities of the individual receptor domains also made it feasible to investigate the receptors’ structure and folding at atomic resolution using NMR and X-ray crystallography.
Receptors 2023, 2

Figure 2. Structural analysis of the glucocorticoid receptor DNA-binding domain. Upper panel shows the distribution of amino acids and the coordination of Zn ions (large circles). Small blue circles represent individual amino acids. The amino acids representing the P-box (red) and D-box (green) residues are highlighted. Lower panel shows the 3D solution structure of the GR-DBD: circles represent Zn ions. The location of the recognition helix, with the P-box residues (Helix 1), is indicated along with α-Helix 2 and 3; the D-box residues involved in dimerization are also highlighted (Pbd 1gdc).

Figure 3. Structural analysis of the human glucocorticoid receptor. The upper panel shows the domain organization of the hGR: NTD/AF1, N-terminal domain containing activation function (AF) 1; DBD, DNA-binding domain; hinge region (amino acids 489 to 529); and LBD, ligand-binding domain. Lower panels summarize available structural information for the helical regions of NTD/AF1, the DBD homodimer (Pdb 1gdc), and agonist- (Pdb 1P93) and antagonist- (Pdb 1NHZ) bound GR-LBD (GR4 and GR1, respectively, see text for details): Helix 12 and the β-turn involving strands 3 and 4 are indicated.
The first structural model to be reported was the solution structure of the DNA-binding domain (DBD) for the rat GR by Gustafsson and colleagues (Figure 2) [21,22]. The DBD is the most conserved region of steroid receptors, spanning 66 to 68 amino acids, and is a defining feature of the nuclear receptor superfamily. The importance of Zn ions and a scheme for coordination by cysteine residues was proposed by Yamamoto and colleagues [23], and the role of Zn coordination for the elegant globular folding of the DBD was revealed by the NMR structure [21] and mutagenesis [24] (Figure 2).

The 3D model of the isolated GR-DBD involved the packing of hydrophobic residues to form a core, with two α-helices at right angles to each other (residues Ser459 to Glu469 and Pro493 to Gly504). The N-terminal helix formed the DNA recognition helix, and had P-box residues that determined the DNA binding specificity [19,20,25]. Other elements of the identified secondary structure included type I and type II turns and a short element of β-sheet (residues Leu 475 to Gly478) (Figure 2). From the NMR structure, it was possible to model a dimer of the GR-DBD bound to a DNA response element, showing how the D-box residues [25] in the second zinc finger module interacted to create a dimerization interface [26–28]. This structure of the DBD is now recognized as a canonical fold within the nuclear receptor superfamily, with all subsequent structures exhibiting the same general geometry. From this very first structure of an isolated steroid receptor domain, Gustafsson and collaborators have continued to provide novel insights into the structure and function of the GR (Figure 3).

3.2. GR-NTD

The amino-terminal domain (NTD) of steroid receptors contains regions important for gene transactivation (reviewed in [29]). In the case of the human GR, mutagenesis studies by Evans and co-workers identified the region between amino acids 77 and 262, termed τ1 (or AF1), as important for GR-dependent transactivation [15,30]. The independent function of this domain to act as a transactivation domain was further demonstrated in yeast when fused to the GR DNA-binding domain in vivo [31] or in an in vitro cell-free transcription assay [32]. Early work also highlighted the role that the GR-NTD played in protein–protein interactions, notably with components of the general transcription machinery such as the TATA-binding protein [32–34]. In the case of GR-AF1 (τ1), the critical region of activity was mapped to a core 58-amino-acid sequence (residues 187 to 244) [35,36].

Members of the steroid receptor subfamily of nuclear receptors have long NTDs, which have varying degrees of intrinsically disordered structure (IDS) and the propensity to adopt an α-helical structure [29]. The IDS of the GR-NTD plays a role in facilitating allosteric regulation and interdomain communication [37], and as a site for phosphorylation, leading to structural and functional changes in receptor activity [38–41]. For example, phosphorylation of serines 211 and 226 in the GR-AF1 core region enhanced the binding of the coregulator CREB-binding protein (CBP) [42]. More recently, the IDS of steroid receptors has been shown to be a key element in the formation of liquid–liquid condensates and the assembly of transcriptional active complexes (with Mediator) at enhancer regions [43]. However, Stortz et al. [44] did not find evidence supporting an essential role of GR IDS in condensate formation, but, rather, highlighted the importance of the LBD. The contrasting findings may reflect the different methodological approaches used, as one study was primarily in vitro using purified proteins [43], while the other used a cell model for condensate formation [44]. Interestingly, in the case of the AR, the roles of both the intrinsically disordered structure (NTD) [45,46] and of other domains [47] have also been reported to mediate condensate formation. Importantly, condensate formation involving the AR-NTD has been associated not only with the assembly of transcriptional complexes (Med1, open chromatin histone marks), but also with the emergence of resistance of prostate cancer cells to classical antiandrogen drugs [48,49]. Collectively, these studies highlight the functional importance of the NTD and the intrinsic structural flexibility of this region of the GR, as well as other steroid receptors, for gene regulation and allosteric fine-tuning of receptor activity.
The first evidence for the IDS within the NTD of a steroid receptor came from biophysical analysis (circular dichroism) and 2D-homo- and hetero-nuclear NMR and mutational analysis of the core region of τ1 (AF1core) [50]. This was confirmed and further investigated by Kumar and Thompson and colleagues, who showed changes in terms of folding in response to natural osmolytes and TBP binding [34,51,52]. Importantly, the structural flexibility was also seen for the full-length GR-AF1 (τ1) [50]. In the presence of a hydrophobic solvent or a natural osmolyte, an increase in the α-helical structure was observed, and the disruption of the helix by the introduction of proline residues disrupted transcriptional activity. The locations of three helical segments were then mapped by NMR: 189 to 200, 216 to 226, and 234 to 239 (Figure 3) [50]. Importantly, structural changes were observed upon binding the TATA-binding protein (TBP), consistent with the model of induced folding of transactivation domains [34]. Varying degrees of intrinsically disordered structure and secondary structure elements have now been reported for the NTD of the androgen [53,54], estrogen (α and β) [55], mineralocorticoid [56], and progesterone receptors [57].

3.3. GR-LBD

The ligand-binding domain of steroid receptors, region A, represents the C-terminal third of the receptor protein (Figure 3), based on earlier partial proteolysis experiments (Figure 1). The overall folding of the agonist-bound GR-LBD conforms to the canonical α-helical rich globular conformation seen for all members of the nuclear receptor superfamily. It is characterized by a three-layer α-helical sandwich structure (Figure 3) [58–60]. In the agonist-bound complex, helix 12 adopts a conformation that seals off the ligand-binding pocket (LBP) and creates the activation function 2 surface, important for co-activator protein binding (Figure 3). However, it is notable that all the reported structures of the GR-LBD had mutations introduced (e.g., Phe602Ser, Cys638Asp) in order to stabilize the expression of the recombinant protein and aid in the production of diffractable crystals [58,59]. The structure of the cortisol-bound LBD had mutations introduced in helix 9 (Glu684Ala/Glu688Ala) to aid in crystallization [60]. The agonist (dexamethasone or cortisol)-bound structures were almost identical, and showed the activated orientation of helix 12 as well as the creation of a hydrophobic pocket to which an SRC-2 (TIF2) NR-box peptide (LxxLL) was bound (Figure 3) [58–60].

The LBP of the GR has a volume of 578 Å³ to 599 Å³, and the binding of the synthetic glucocorticoid dexamethasone involves extensive hydrophobic and hydrogen bond interactions [58,59]. A comparison of the structures with dexamethasone or cortisol reveals a highly similar network of hydrogen bonding involving the amino acids Arg611, Gln570 (steroid A ring), Asn564, Gln642, and Thr739 (steroid D ring), as well as an additional bond between residues Tyr735 and Asn642 and the D-ring sidechain of cortisol [58–60]. Glutamine 642 is GR-specific, and hydrogen bonds to the ligand and plays a role in steroid recognition and selection; however, other amino acid agonist (dexamethasone) contacts involve residues conserved in PR and AR.

The functional conformation of steroid receptors is known to be a homo-dimer of receptor monomers, with the sequences in the DBD and LBD forming dimerization interfaces. There is a good consensus that a common role of the D-box residues exists (DBD; Figures 2 and 3) in dimer formation. In contrast, there is considerable variation and debate as to the sequences in the LBD involved in dimerization (reviewed in [3]). In the original structure, according to Xu and co-workers, residues isoleucine 628 and proline 625 were identified as forming a unique dimerization interface involving an intermolecular β-sheet formed by the β-turn structure and strands 3 and 4 (Figure 3) [58]. In contrast, in the four crystal structures reported by Gustafsson, Carlquist, and co-workers, there appeared to be variation in the dimer surface [59], although one of the observed dimer structures agreed with the earlier findings of Bledsoe et al. [58].

In addition to the agonist-bound GR-LBD structure, Kauppi et al. [59] also solved a structure with the RU486 antagonist bound (Figure 3). Strikingly, in one crystal (GR1), that structure was comparable to the earlier structure for ERα-LBD, which bound with
the selective estrogen ligand, raloxifene [61]. However, in another crystal (GR3), there was significant disruption of helix 12 and part of helix 11, such that this region interacted with the AF2 pocket of a neighboring molecule (Figure 3) [59]. In both GR1 and GR3, the coactivator binding pocket (AF2) was obstructed. In the antagonist conformation, helix 12 was observed to block the AF2 surface and binding of co-activator proteins [61]. A comparison of the structures, solved by Kauppi et al. [59], also showed how Tyr-735 (phenylalanine in the AR and tyrosine in PR) adopts different rotamer conformations in the agonist and antagonist structures. This residue, towards the C-terminus of helix 11, which has hydrogen bonds with cortisol (see above) had previously been identified as being important for receptor-dependent transactivation without disrupting ligand binding [62].

Collectively, over the last forty years, the Gustafsson laboratory and collaborators have led the way in the structural analysis of the GR. This has revealed the three-domain organization of the full-length receptor and the globular folding of the DBD and LBD, as well as novel insights into the secondary structure and folding of the intrinsically disordered GR-NTD. In addition, work from this group has contributed to our understanding of the structure and folding properties of the ERα, β LBDs, and NTDs. More recent studies have focused on analyzing full-length receptor complexes bound to DNA.

4. LXR-RXR Complex Bound to DNA

The liver X receptor (LXR) is an important regulator of lipid and cholesterol metabolism, inflammation, neural development, and cancer (for recent reviews, see [63–65]). LXR is an oxysterol-dependent nuclear receptor that acts as a heterodimer with RXRs and shows a preference for direct repeats of half-site AGGTCA spaced by four nucleotides [66]. The Gustafsson and Webb groups solved the crystal structure for LXRβ (amino acids 72 to 461): the RXRα (amino acids 98 to 462) complex bound to the DNA sequence 5′taAGGTCActtcAGGTCA-3′ and the SRC-2 LxxLL peptide [67]. Each receptor was truncated for the first 71 or 97 N-terminal amino acids, respectively, and sequence amino acids 98 to 128 of the RXR were not resolved in the crystal structure. The structure revealed an asymmetric organization, biased towards the 3′ half site of the DNA response element (Figure 4) [67]. In the resolved structure, the RXRα-DBD occupies the 5′ half site with LXRβ-DBD binding the 3′ half site. This extended architecture of the nuclear receptor–DNA complex has also been seen for other RXR heterodimers using the methods cryo-EM [68] and solution angle scattering (SAXs) [69–71]. The exception is the PPARγ/RXR crystal bound to a DR1 element structure, which had a more closed and compact conformation [72]. Thus, the overall structure of the RXRα-LXRβ DNA bound complex is one of an open conformation, with the LXRβ-LBD positioned over the DBDs and the RXRα-LBD found 3′ over the LXRβ-DBD (Figure 4). Interestingly, in contrast to other RXR-NR complexes, there is little evidence for interdomain communication in the RXRα-LXRβ complex, but there is a large, buried surface representing the dimerization interface in the LBDs [67]. Importantly, the overall folding of the LBD (see, for example, the earlier crystal structure for LXRβ-LBD [73]) and DBD in all the multi-domain complexes solved so far resemble the globular structures of the isolated domains. However, what these structures reveal is how the LBDs and DBDs orientate with respect to the nature of the DNA response element and the importance of the hinge region between the DBD and LBD to accommodate different geometries. Another notable feature of the RXRα-LXRβ complex was evidence that the amino acid sequences flanking the core DBD participated in DNA binding. Thus, although the RXRα-NTD was not resolved in the crystal structure, intriguingly, region residues 75 to 80 of LXRβ, together with the C-terminal extension, were found to make minor groove interactions with the four-nucleotide spacer sequence (Figure 4) [67]. Interestingly, the recent cryo-EM structures for the ER [74] and AR [75] proposed interesting conformations of the NTD based on antibody epitope exposure and receptor monomer orientation on DNA. In the ERα, the NTD was located adjacent to, and created a platform with, the LBD allowing recruitment of the coactivator SRC3 [74]. In contrast, in the AR-DNA complex, the
NTD domains encircled the LBDs, but again created a platform for co-regulatory protein recruitment (SRC3, p300) [75].

Figure 4. Crystal structure of the LXRβ-RXRα complex bound to DNA. The structure of RXRα (residues 98 to 462; green) and LXRβ (residues 72 to 461; turquoise) bound to a DR4 response element (5′-TAAGGTCACTTCAGGTCA-3′) was resolved at 3.1 Å (Pdb 4nqa). Each receptor was ligand-bound and interacted with a LXXLL motif from SRC2 (magenta).

5. Future Studies

Glucocorticoid hormones and the glucocorticoid receptor are known to be key regulators of human physiology, playing roles in immunity, metabolic regulation and pathological conditions including inflammation, rheumatoid arthritis, asthma, and cancer [76]. There is, therefore, considerable interest in understanding the structure–function relationships of the receptor proteins from scientists, clinicians, and the pharmaceutical industry. A key gap in our knowledge is the lack of structural information for the full-length hGR bound to DNA. The recent successes with ERα and the AR suggest this is now more feasible using advances in cryo-EM. Such a structure would provide valuable insights into the overall architecture of the receptor–DNA complex. Having an additional structure of the full-length steroid receptor complex will also provide a better understanding of the folding and arrangement of the intrinsically disordered GR-NTD in the complex, and will allow for comparison with existing structures. One of the early successes from the Gustafsson group was the generation of a panel of monoclonal antibodies that specifically recognized the GR-NTD; these would be valuable probes in any future cryo-EM structure of a transcriptionally active complex. Other important areas that would benefit from structural analysis are the mechanism of trans-repression, whereby GR negatively regulates transcription through interactions with other transcription factors, such as NFκB, and, secondly, the impact of post-translational modifications on receptor folding and conformation. In conclusion, this brief overview illustrates how Jan-Ake Gustafsson and colleagues have contributed widely to our understanding of the structure–function relationships of the GR and related nuclear receptors. However, there is always more to learn.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

5. Govindan, M.V.; Devic, M.; Green, S.; Gronemeyer, H.; Chambon, P. Cloning of the human glucocorticoid receptor cDNA. *Nucleic Acids Res.* 1985, 13, 8293–8304. [CrossRef]


47. Ahmed, J.; Meszaros, A.; Lazar, T.; Tompa, P. DNA-binding domain as the minimal region driving RNA-dependent liquid–liquid phase separation of androgen receptor. *Protein Sci.* **2021**, 30, 1380–1392. [CrossRef]


65. Song, X.; Wu, W.; Warner, M.; Gustafsson, J.A. Liver X Receptor Regulation of Gial Cell Functions in the CNS. *Biomedicines* 2022, 10, 2165. [CrossRef]


67. Lou, X.; Toresson, G.; Benod, C.; Suh, J.H.; Philips, K.J.; Webb, P.; Gustafsson, J.A. Structure of the retinoid X receptor α-liver X receptor β (RXRa-LXRβ) heterodimer on DNA. *Nat. Struct. Mol. Biol.* 2014, 21, 277–281. [CrossRef]

68. Orlov, I.; Rochel, N.; Moras, D.; Klaholz, B.P. Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. *EMBO J.* 2012, 31, 291–300. [CrossRef]


Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.