Review

CellPress

Advances in our structural understanding of orphan nuclear receptors

Nerea Gallastegui[1], Jonathan A.G. Mackinnon[1], Robert J. Fletterick[2], and Eva Estébanez-Perpiñá[1]

[1] The Institute of Biomedicine of the University of Barcelona (IBUB), Department of Biochemistry and Molecular Biology, University of Barcelona (UB), Baldiri Reixac 15-21, 08028 Barcelona, Spain
[2] The Department of Biochemistry and Biophysics, University of California San Francisco (UCSF), San Francisco, CA, USA

Nuclear receptors (NRs) are key players in the regulation of gene expression, coordinating protein assemblies upon their surfaces. NRs are regulated by ligand binding, which remodels the interaction surfaces and subsequently influences macromolecular complex formation. Structural biology has been instrumental in the discovery of some of these ligands, but there are still orphan NRs (ONRs) whose bona fide ligands have yet to be identified. Over the past decade, fundamental structural and functional breakthroughs have led to a deeper understanding of ONR actions and their multidomain organization. Here, we summarize the structural advances in ONRs with implications for the therapeutic treatment of diseases such as metabolic syndrome and cancer.

Nuclear receptors: scaffolding proteins in transcription

NRs act as master regulators of reproduction, development, and metabolism, and their key role is embodied by their central position in transcriptional regulation [1–3]. NRs work together with a diverse group of proteins termed 'coregulators' (coactivators and corepressors), which define NR tissue-specific actions. NRs require coregulators to modulate gene transcription, typically via chromatin modifications and interactions with the basal transcription machinery [4]. NRs are key elements in these large multimeric assemblies, whose structure evokes the image of the mythological Titan Atlas holding up the celestial heavens on his back. Structurally, NRs comprise an N-terminal (NTD) domain, a DNA-binding domain (DBD), a hinge, and a ligand-binding domain (LBD) (Box 1, Figure S1 and S2 in the supplementary material online) [5]. Coregulators have been shown to bind to different NR domains [6]. Their recruitment is affected by the conformational changes that many NRs undergo in response to small hydrophobic ligands binding to the nested ligand-binding pocket (LBP) found in the LBD. This incredibly adaptable pocket can vary in volume from almost nonexistent to spacious cavities of up to 1600 Å^3 (Figure 1, Table S1 in the supplementary material online) [6,7]. The LBP has also been shown to vary in size, not only between different NR subfamilies and isoforms, but also within the same receptor complexed to different ligands. Ligand-directed NR actions have been the focus of many successful pharmacological efforts to develop synthetic modulators due to their involvement in myriad diseases, such as metabolic diseases and cancer, thus making them prime biomedical targets. As a result, more than 10% of current medicines target NRs, an achievement made possible only due to decades of structural and functional research [1,8,9].

Surprisingly, no bona fide ligands have been identified for about half of the human NRs. Such NRs are termed 'ONRs' (see Glossary; for additional information, see www.nursa.org) and, once their natural ligand has been validated, the receptor is regarded as 'adopted' [6,9–11]. Finding the unidentified physiological ligands of ONRs has crucial pharmacological implications and this process is referred to as 'deorphanization'. There are many examples where NR structures have been pivotal for ligand identification [6,10]. Here, we present insights into the ONR structural breakthroughs of the past decade, including the only full-length (FL) ONR crystal structure solved thus far, hepatocyte nuclear factor 4 alpha (HNF4-α; Boxes 1 and 2;

Glossary

Constitutive NR: a NR with a basal activity devoid of a ligand requirement.
Inverse agonist: ligand that binds to the receptor as an agonist but induces a pharmacological response of that of an antagonist.
NR ligands: the classical molecular endocrinology view of NR ligands comprised small lipophilic endocrine molecules within the cell, such as hormones and lipid derivatives. However, functional and structural studies have broadened the range of chemically distinct ligands able to bind to this protein superfamily. For instance, deorphaned examples exist that have more atypical ligands, such as a heme group or environmental chemicals that can exert an endocrine-disruptor function (e.g., bisphenol A).
NR standard structural orientation: the standard orientation for the overall description of NR LBD structures features a frontal view on the LBD body displaying H3 vertically and dividing the LBD in two. To the right of H3, H1 and H2 are easily visible. To the left of H3, one can place H12, and right above H12, the coactivator binding groove or AF-2 pocket. Figure 1 (main text) shows the labeling and location of H1–H12 as well as the sliced surface representation of a NR depicting the enclosed the LBP.
Orphan nuclear receptor (ONR): NRs whose natural ligands have not been identified or have not been agreed upon to be of physiological relevance.

Corresponding author: Estébanez-Perpiñá, E. (evaestebanez@ub.edu).
Keywords: orphan nuclear receptors; structure–function relations; ligands; full-length; deorphanization.

0968-0004/
 http://dx.doi.org/10.1016/j.tibs.2014.11.002

Box 1. NR domain structure

N-terminal domain (NTD)
The NTD is the least conserved domain of NRs both in sequence and length and can range from a few residues to huge modules [76]. The NTD strongly drives transactivation and coregulator binding, but its intrinsic disorder hinders atomic characterization. The mechanistic importance of the structural flexibility of the NTD and its mutation in human diseases reinforces the potential of this domain as a drug target [76].

DBD
The DBD was initially considered to be a passive anchor for specific DNA sequences, but more recently it has been shown to have an active role in NR interdomain allosterism and DNA crosstalk [77]. NRs recognize DNA sequences with specificity and affinity through a conserved DNA-binding fold within the DBD (Figure S3 in the supplementary material online). NRs can be classified by their mode of DNA binding: (i) steroid homodimers (GR, PR, ER, AR, and MR); (ii) nonsteroidal heterodimers (TR, VDR, RAR, RXR, PPAR, Nurr1, NOR1, TR2, and TR4) [1]; (iii) nonsteroidal homodimers (TR, REV-ERB, RXR, HNF4-α, PNR, TLX, and COUP-TF); and (iv) monomers (ERR, SF-1, NGF-1, LRH-1, TR2, TR4, ROR, and TLX) [78,79,82–85]. However, some NRs exhibit distinct characteristics: TR2, TR4, and TLX can bind DNA as both monomers and dimers, GCNF binds as oligomers, and DAX1 and SHP are DBD-absent NRs.

Hinge
The hinge is a poorly conserved region connecting the DBD with the LBD and its flexibility makes it prone to protease cleavage. Although initially regarded as a passive polypeptide linker, the hinge is a hotspot for post-translational modifications that influence translocation, DNA binding, and transactivation. It is also linked to disease. Structural studies that have shed light on this region have been achieved through the determination of the FL crystal structures of the PPARγ–RXRα–complex, VDR–RXRα, RXRα–LXRβ, RARα–RXRα and the HNF4-α homodimer [39,83–85]. These studies have revealed that the hinge has extensive contact with DNA and the other NR modules (Box 2, and Figure S3 in the supplementary material online). Consequently, this has exposed unforeseen connectivity between NR domains, which may modulate NR actions [8,22,39,78–81].

LBD
The LBD recruits coregulators and aids in transcriptional regulation [5]. Most importantly, many NRs are activated by small specific hydrophobic ligands, which bind to a nested expandable LBP, whose volume can range from almost nonexistent to large and spacious cavities [5]. Ligand-directed NR actions have triggered many successful efforts in the development of synthetic modulators, making these proteins prime therapeutic targets in biomedicine [6].

Figures 1–3). Furthermore, we summarize the insights that these structures have given for designing novel therapeutics. We present a comprehensive overview of ONR structures for both structural biologists and the general reader with emphasis on their distinctive traits that challenged the initial view of a canonical NR LBD structure (Figures 1 and 2).

Current knowledge of the canonical LBD

X-ray crystallography has been the most powerful technique for unraveling NR-LBD structures. Despite only providing static snapshots of NRs, X-ray crystallography has provided invaluable mechanistic insights into the flexibility and adaptability of NRs (Figures 1–3) [5,6,8]. A prototypical LBD features 10–12 helices and one to two short β turns (Figure 1A) assembled in a three-layer helical sandwich with a cryptic LBP. Functionally, ligands act as switches for coregulator binding by triggering major movements of the most C-terminal helix (H12). The ligand-dependent conformational variability of H12, relative to the rest of the LBD, defines whether the LBD displays an active conformation or one of the repressed conformations (Figure 2) [6]. In an active LBD, H12 is approximately perpendicular to H3 and H5, forming a triangular arrangement that has a fully functional and accessible coactivator-binding groove (AF-2) [12]. This groove permits coactivator binding through conserved LxxLL motifs found in coactivators [13]. In the absence of ligand, the H12 has been observed in an array of poses that range from an autorepressed conformation, which blocks coactivator assembly by folding back against the LBD to cover the preshaped H3 and H5, to a H12 that projects freely into solution to leave the AF-2 incomplete [6]. Interestingly, only the active conformation permits coactivator binding while other conformations have been shown to permit corepressor binding. To bind to corepressors, NRs differentially associate with amino acid sequences that resemble the coactivator motifs (LxxI/HIxxxI/L) [14–17]. Several studies have shown that coactivator and corepressor docking surfaces do overlap but are distinct [18–20].

The common model of what NR ligands are and what they do is derived from studies using classical endocrine NR–endogenous ligand complexes [21]. However, knowledge from different ONRs has challenged this strict canonical view. For example, the LBP of some ONRs has been shown to be either practically nonexistent (hydrophobic side chains fill up the cavity) or have a large empty volume (Figure 1, Table S1 in the supplementary material online) [6,10]. The variability in the LBP shape and volume results from the overall fold of the LBD, caused by rearrangement of the major helices, the absence or presence of particular helices and β-turns, and the influence or absence of ligands (Figure 1B). Additionally, coregulators bound to specific surface patches further impact the internal LBP cavity due to structural adaptations elicited by their binding [7]. Furthermore, some ONRs depict H12 in a constitutive agonist conformation despite the lack of ligands. Combined, these observations raise the question of whether these ONRs are truly ligand independent. Also, if a ligand is found, the question remains as to whether all NR ligands act as conformational switches or whether they merely bind to a more fixed structure than was originally believed.

Crystal structures of several ONRs were instrumental in the initial deorphanization stages. Serendipitously identified ligands have been observed in the crystal structures derived from protein expression, purification, and crystallization. These ligands have guided scientists in the process of deorphanization, providing key cues for proceeding ligand validation strategies [6]. Here, we describe the different structural traits of each ONR in each subfamily (categorized according to sequence homology) and how those traits may aid in future ONR deorphanization [11].

Subfamily 1: thyroid hormone receptor-like

NR subfamily 1 comprises of ten members: the thyroid (TR), retinoic acid (RAR), peroxisome proliferator-activated (PPAR), oxysterol (LXR), farnesoid (FXR), vitamin D (VDR), RAR-related orphan (ROR), REV-ERB, pregnane X