CellPress

Nuclear receptor full-length architectures: confronting myth and illusion with high resolution

Fraydoon Rastinejad[1], Vincent Ollendorff[2], and Igor Polikarpov[3]

[1] Sanford-Burnham Medical Research Institute, Metabolic Disease Program, 6400 Sanger Road, Lake Nona, FL 32827, USA
[2] INRA, UMR866 Dynamique Musculaire et Métabolisme, F-34060 Montpellier Université Montpellier 1, F-34000 Montpellier - Université Montpellier 2, F-34000 Montpellier, France
[3] Instituto de Física de São Carlos, Universidade de São Paulo, Av. Trabalhador São-Carlense, 400, São Carlos, SP, 13560-970, Brazil

The crystal structures of three nuclear receptor (NR) complexes have emerged to reveal their multidomain architectures on DNA. These pictures provide unprecedented views of interfacial couplings between the DNA-binding domains (DBDs) and ligand-binding domains (LBDs). The detailed pictures contrast with previous interpretations of low-resolution electron microscopy (EM) and small angle X-ray scattering (SAXS) data, which had suggested a common architecture with non-interacting DBDs and LBDs. Revisiting both historical and recent interpretations of NR architecture, we invoke new principles underlying higher-order quaternary organization and the allosteric transmission of signals between domains. We also discuss how NR architectures are being probed in living cells to understand dimerization and DNA-binding events in real time.

A brief history of single domain structures

NRs are metazoan transcription factors that regulate metabolism, development, homeostasis, and reproduction. In humans, the 48 NRs can be divided into four groups based on their receptor dimerization patterns and DNA-type preferences. The first group forms homodimers, binds to DNA inverted repeats, and includes steroid receptors, such as glucocorticoid receptor (GR), estrogen receptor (ER), progesterone receptor (PR), androgen receptor (AR), and mineralocorticoid receptor (MR). A second group heterodimerizes with retinoid X receptor (RXR), binds to DNA direct repeats, and includes receptors such as peroxisome proliferator-activated receptor (PPAR), retinoic acid receptor (RAR), vitamin D receptor (VDR), and thyroid hormone receptor (TR). A third group comprises homodimers that bind to DNA direct repeats, such as hepatocyte nuclear factor 4 (HNF-4α) and Rev-Erb. The fourth group contains monomers that bind to extended single DNA half-sites, including receptors such as RAR-related orphan receptors (RORs) and NURR family members [1–3]. Consensus half-sites are typically 5′-AGGTCA-3′ sequences for nonsteroid receptors, and 5′-AGAACA-3′ sequences for steroid receptors.

When viewed from their N to their C terminus, NR polypeptides exhibit a modular organization with five to six segments, designated A–F. Only two domains had been well characterized through high-resolution structural methodologies. These are the DBD that specifically contacts response elements, and the LBD that recognizes endogenous small-molecule ligands and coregulator regions [4–6]. Crystallographic studies on DBD–DNA complexes revealed the basis for half-site recognition, and the roles of inter-half-site spacing and half-site repeat nature as selectivity features [2]. Crystallography later revealed how ligands are bound in the LBD structures, beginning with the TR and RAR [6–8]. The binding of different types of ligand to a single NR was subsequently shown for the ER through a series of detailed structure–function studies [9,10].

Most NR LBDs have the capacity to bind coactivator segments with LXXLL sequences, and corepressor segments with LXXXLXXX [I/L] sequences (where L = leucine, I = isoleucine, and X = any amino acid) [11,12]. These short elements interact at the LBD surface in a manner that depends on the ligand occupied inside the LBD pocket. Components of coregulator complexes modify the histone tails in chromatin, favoring either the activation or repression of target genes [13]. Early crystallographic studies addressed how coactivator LXXLL segments recognize the surfaces of LBDs, focusing on PPARγ and ER LBDs [10,14].

These and subsequent structural studies of isolated DBDs and LBDs provided a deep understanding of the molecular interactions within each of these domains [6]. However, our understanding was incomplete because these studies did not reveal how the many different domains and segments of a NR cooperate in the context of a quaternary architecture with functional relevance. These missing insights prevented the field from fully considering allosteric communications, such as how ligand binding may lead to changes in DNA binding and vice versa. Now, three published reports reveal the detailed, higher-order molecular architectures of NR complexes using X-ray crystallography [15–17]. These pictures show surprisingly complex domain–domain interconnections, also providing new insights into how signals can be

Corresponding author: Rastinejad, F. (frastinejad@sanfordburnham.org).
Keywords: nuclear receptor; transcription factor; allostery; X-ray crystallography; SAXS; electron microscopy; BRET.

0968-0004/
 http://dx.doi.org/10.1016/j.tibs.2014.10.011

communicated between domains in an allosteric fashion. Previously, a different picture was proposed for full-length NRs, based only on solution-based and low-resolution techniques. That picture was based on the two conserved domains (DBDs and LBDs) of NRs having no direct contact with one another when the receptor was bound to its DNA element, being organized instead as isolated beads (domains) positioned on opposite ends of an extended string (hinge region). That disconnected domain architecture now seems inconsistent with both the recent set of crystallographic findings based on multiple NR complexes, and the larger body of structural, biophysical, and cell-based studies that support NR quaternary structures on DNA that have highly coupled DBD–LBD interfaces for allosteric communications between these domains. In this review, we discuss both the historical and newly reported findings that mechanistically examine how NR–DNA complexes use their complex molecular architectures to sense and transmit signals through their domains.

Carefully revisiting the mousetrap

One of the critical early goals of NR structural biology was to define the LBD conformations that could be reliably described as both the inactive and active states. To this end, the structure of the unliganded RXRα LBD structure was compared to a subsequent structure of the RARγ LBD with the activating ligand *all-trans* retinoic acid [8]. That comparison led the authors to propose a mousetrap mechanism for ligand activation of NRs [8]. As shown in Figure 1A, ligand binding was suggested to induce an altered position in Helix-12 (H12). H12 was described as a stable helix located away from the LBD body in the apo-state (deemed to be the inactive conformation). Upon ligand binding, H12 moves to a new position on the surface the LBD, entrapping the ligand (active conformation), hence the 'mousetrap' mechanism. However, further analysis of the mousetrap mechanism using those original crystallographic coordinates suggests that this interpretation was misguided (Figure 1B). The H12 position in the apo-state is positioned through artificial crystal-packing interactions.

An alternative, better-supported model for ligand activation, proposed by Schwabe and colleagues, was derived from their fluorescence spectroscopic studies [18]. This mechanism, known as H12 dynamic stabilization, instead characterizes the inactive LBD state as one with relatively high mobility and lack of structural order in H12. A disorder-to-order transition is induced with binding of activating ligands. Similarly, nuclear magnetic resonance (NMR) studies of several NRs revealed ligand-induced stabilization of NR LBDs and the correspondent H12s [19–22]. Supporting data for the H12 dynamic stabilization mechanism is strong. These data come from studies conducted on RXRα LBD, as well as several other NR LBDs [23–26]. These reports used hydrogen-deuterium exchange mass-spectrometry (H/D ex MS) studies, and consistently found faster hydrogen exchange in H12 in the apo-state, confirming its relative lack of structural order. This is inconsistent with the idea of H12 as a stable helix, as had been proposed in the mousetrap model. Instead, the binding of activating ligands produces a stable helical conformation in H12 residues and also adds global stability to the LBD fold.

The so-called 'inactive' H12 appears similarly misinterpreted in the estrogen-related receptor (ERR) LBD structure [27]. Here again the authors describe the location of H12 as both ordered and positioned away from the receptor LBD, so as to be consistent with their original notion of the mousetrap mechanism [27]. However, as indicated in Figure 1C, the position of H12 is again strongly influenced by crystal-packing interactions. Intriguingly, a second ERR subunit in the asymmetric unit, ignored in that report, shows a disordered state (not visible electron density) predicted by the H12 dynamic stabilization model [27].

A related question has been how ligands enter and exit the LBD. One hypothesis, that different parts of the NR LBD body can harbor a gate for ligand entry or exit from the ligand-binding cavity, goes back to the first liganded NR LBD structure [7]. The H1–H3 loop and neighboring β-sheets were speculated to be a ligand entry site to the TR LBD in work from the Fletterick lab [7]. This hypothesis received further support from molecular dynamics simulations of the NR LBDs. The entry/exit channel for the ligand was initially believed to be only on the side of the receptor LBD where H12 is located [28], but several other competing ligand exit pathways were also identified [29–31], including one predicted by the analysis of the earlier TR LBD structure [7]. Only subtle protein conformational adaptations were shown to be required for ligand binding to the TR LBD irrespective of the entrance pathway, further indicating that H12 might not be the only, or even preferred, route for ligand association/dissociation with the NR LBDs [32]. Consistent with the previous studies of radioactive estradiol dissociation from ER preparations [33], molecular dynamics (MD) simulations of ligand dissociation from the ER LBD revealed that preferred pathways of ligand dissociation from the LBD are mediated by the NR quaternary state [34]. Thus, combined evidence from crystallographic structures, NMR, MD simulations, and biochemical studies all call for the questioning of the mousetrap model mechanism of NR activation.

Illusions of allostery

Most NR LBDs can bind alternatively to receptor-specific coactivators or corepressors, with the ligand acting as the switch for their coregulator exchange. For dimeric NRs, an important question has been whether two coregulators motifs bind equivalently to both subunits. Establishing the true binding stoichiometry between coregulator motifs and receptor dimers has proved to be particularly confusing in the case of RAR-containing dimers. Studies with isolated LBDs of RXR heterodimers (such as RXR–RAR) were interpreted to indicate that only one subunit in some RXR heterodimers can bind to the coactivator-derived LXXLL motif [35]. A combination of SAXS and X-ray crystallographic studies were applied in a study to understand how this asymmetric binding of coregulators is established with isolated RAR LBDs [35]. However, in this study, the homodimer of RARβ LBD was used instead of the functional RXR–RAR heterodimer, and the authors proposed an allosteric mechanism to account for 1:2