Journal of Structural Biology 185 (2014) 355–365

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Local motifs involved in the canonical structure of the ligand-binding domain in the nuclear receptor superfamily

Motonori Tsuji *

Institute of Molecular Function, 2-105-14 Takasu, Misato-shi, Saitama 341-0037, Japan

ARTICLE INFO

Article history:
Received 2 September 2013
Received in revised form 14 December 2013
Accepted 16 December 2013
Available online 19 December 2013

Keywords:
Agonism
Antagonism
Local motif
Charge–dipole interaction
Signal amino acid
Nuclear receptor ligand-binding domain

ABSTRACT

Structural and sequence alignment analyses have revealed the existence of class-dependent and -independent local motifs involved in the overall fold of the ligand-binding domain (LBD) in the nuclear receptor (NR) superfamily. Of these local motifs, three local motifs, i.e., AF-2 fixed motifs, were involved in the agonist conformation of the activation function-2 (AF-2) region of the LBD. Receptor–agonist interactions increased the stability of these AF-2 fixed motifs in the agonist conformation. In contrast, perturbation of the AF-2 fixed motifs by a ligand or another protein molecule led the AF-2 architecture to adopt an antagonist conformation. Knowledge of this process should provide us with novel insights into the 'agonism' and 'antagonism' of NRs.

1. Introduction

The nuclear receptor (NR) is a transcription factor that controls many crucial biological events such as morphogenesis and homeostasis. NRs constitute a superfamily and are widespread among eukaryotes. In mammalian species, this superfamily is composed of retinoic acid (RAR and RXR), steroid hormone (AR, ER, GR, MR, and PR), thyroid hormone (TR), vitamin D3 (VDR), prostaglandin or lipid mediator (PPAR), and other ligand-orphan receptors. NRs exhibit a modular structure with five or six functionally separable domains. Of these functional domains, the ligand-binding domain (LBD) acts as a molecular switch functioning in a ligand-dependent manner, i.e., the ligand-dependent activation function (AF-2) of transcription (Bourguet et al., 1995; Renaud et al., 1995).

For the last two decades, structural information on NRs has been rapidly accumulating through the use of X-ray crystallography and nuclear magnetic resonance (NMR) experiments, with over 700 three-dimensional (3D) structures of LBDs having been reported. In addition, the 3D structures of both unliganded and liganded LBDs have been solved for the same NRs (Bourguet et al., 1995; Renaud et al., 1995; Egea et al., 2000; Gampe et al., 2000b). These structures have been found to reflect important structural changes between the apo- and holo-LBDs, wherein the structures of their C-terminal regions (helices H11–H12), i.e., their respective AF-2 architectures, differ significantly from each other (Fig. 1a and b) (Wagner et al., 1995). Agonist and antagonist conformations of the AF-2 architecture have also been obtained for

Abbreviations: AncCR, ancestral corticoid receptor; AncGR1, ancestral glucocorticoid receptor 1; AncGR2, ancestral glucocorticoid receptor 2; hAR, human androgen receptor; hCAR, human constitutive androstane receptor; hCOUP-TFI, human chicken ovalbumin upstream promoter-transcription factor-I; hCOUP-TFII, human chicken ovalbumin upstream promoter-transcription factor-II; ssDAF-12, *Strongyloides stercoralis* DAF-12; hDAX-1, human dosage-sensitive sex-reversal adrenal hypoplasia congenital region on the X chromosome gene 1; mDAX-1, mouse dosage-sensitive sex-reversal adrenal hypoplasia congenital region on the X chromosome gene 1; dmDHR38, *Drosophila melanogaster* drosophila hormone receptor 38; hEAR2, human v-erbA related protein-2; hvEcR, *Heliothis virescens* ecdysteroid receptor; hERα, human estrogen receptor-α; mERRγ, mouse estrogen-related receptor-γ; dmFTZ-F1α, *Drosophila melanogaster* fushi tarazu factor 1-α; FTZ, a 19-residue peptide containing the LXXLL motif that binds to FTZ-F1; rFXR, rat farnesoid X-activated receptor; hGR, human glucocorticoid receptor; GRIP1, glucocorticoid receptor-interacting protein 1; hHMR, human testicular receptor 3; hHNF4α, human hepatocyte nuclear factor 4-α; mLRH-1, mouse liver receptor homolog-1; hLXRα, human liver X-receptor-α; hMR, human mineralocorticoid receptor; N-COR1, nuclear receptor co-repressor 1; hNGFI-B, human nerve growth factor-induced B; hNOR, human neuron-derived orphan receptor; hNURR1, human nuclear receptor related 1; hPNR, human photoreceptor cell-specific nuclear receptor; hPPARγ, human peroxisome proliferator-activated receptor-γ; hPR, human progesterone receptor; hPXR, human pregnane-X receptor; hRARγ, human retinoic acid receptor-γ; hRev-erbβ, human Rev-erb-β; hRev-erbAα, human Rev-erbA-α; rRORβ, rat retinoic acid-related orphan receptor-β; hRXRα, human retinoid-X receptor-α; hSF-1, human steroidogenic factor 1; SRC1, steroid receptor co-activator 1; TIF2, transcriptional intermediary factor 2; hTRβ, human thyroid hormone receptor-β; hTR2, human testicular receptor 2; hTR4, human testicular receptor 4; dmUSP, *Drosophila melanogaster* ultraspiracle protein; rVDR, rat vitamin D receptor.

* Fax: +81 48 956 6985.
E-mail address: motonori@molfunction.com

1047-8477/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jsb.2013.12.007

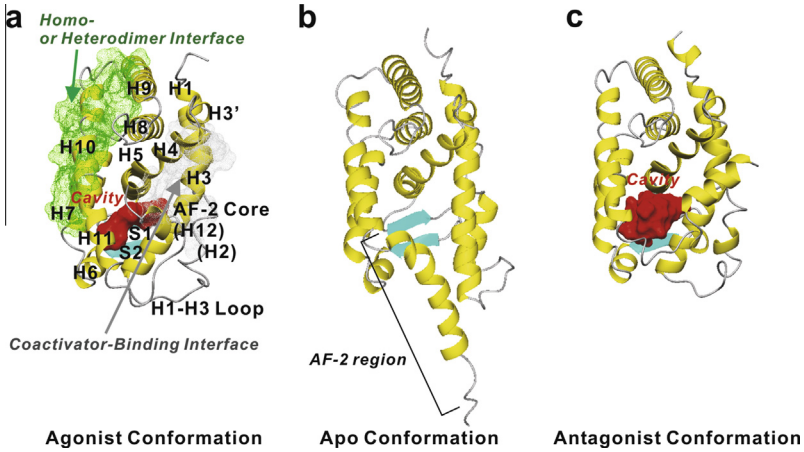

Fig.1. Canonical structures of NR LBDs. (a) hRXRα holo-LBD (1FM9) (Gampe et al., 2000a) in the agonist conformation. (b) hRXRα apo-LBD (1LBD) (Bourguet et al., 1995) in the apo conformation. **c:** mRXRαF318A holo-LBD (1DKF) (Bourguet et al., 2000) in the antagonist conformation.

the same NRs in the presence or absence of a cofactor (i.e., a coactivator or corepressor) fragment (Figs. 1a and c) (Brzozowski et al., 1997; Shiau et al., 1998; Xu et al., 2002; Greschik et al., 2004). In the case of the agonist complex, the AF-2 architecture has adopted a so-called agonist conformation, and the resulting AF-2 core (helix H12) simultaneously participates in the cofactor-binding site, thus creating a coactivator-binding site (Renaud et al., 1995). In contrast, in the case of the antagonist complex, the AF-2 architecture adopts a so-called antagonist conformation, in which the AF-2 core interacts with its own cofactor-binding site composed of the helices H3, H3′, and H4; this conformation of the AF-2 architecture prevents interactions with the cofactor's LXXLL motif (where X can be any amino acid), which is known to be the consensus sequence (Nolte et al., 1998). Moreover, non-crystallographical homo- and heterodimers showing either the agonist or antagonist conformation of the AF-2 architecture in the presence or absence of cofactor fragments have been obtained, and some of their dimer interfaces have been confirmed to date (Fig. 1a) (Bourguet et al., 1995, 2000; Brzozowski et al., 1997; Nolte et al., 1998; Shiau et al., 1998; Williams and Sigler, 1998; Gampe et al., 2000a; Bledsoe et al., 2002; Billas et al., 2003; Svensson et al., 2003; Watkins et al., 2003; Williams et al., 2003; Duda et al., 2004; Greschik et al., 2004; Xu et al., 2004; Kruse et al., 2008; Zhou et al., 2011). The results of these previous studies suggest that the AF-2 region is in a dynamic state, and that the formation of a receptor–ligand complex strongly biases this equilibrium toward either an agonist or antagonist conformation. Thus, the NR LBDs generate or degenerate a cofactor-binding site, and are thereby able to transmit a signal to the basal transcription machinery via the cofactor-binding interface.

On the other hand, as a result of genome analyses of a large variety of organisms, the number of sequences assigned to the NR superfamily has rapidly increased, and more than 1,000 such sequences (i.e., more than 300 variants) have been observed to date (UniPort release 2013_07). In contrast, the 3D structures of LBDs have been reported for only 34 species (48 variants) to date (Protein Data Bank: Berman et al., 2000). Nonetheless, all crystallographically solved 3D structures of the LBDs have revealed the presence of essentially only one common fold (Fig. 1), i.e., a three-layered helical sandwich fold consisting of twelve helices (helices H1–H12) and small, antiparallel β-sheets (sheets S1 and S2) (Bourguet et al., 1995; Renaud et al., 1995).

Brelivet and co-workers have revealed that NRs can be categorized into two different classes (classes 1 and 2) using a structure-based sequence analysis of a large number of NR sequences (Brelivet et al., 2004). Of especial interest in this context is the finding that these two class NR LBDs can clearly be distinguished based on the class-dependent signature motifs related to their oligomeric behavior.

In the present study, we have attempted to reveal the conserved local motifs (including the known signature motifs) involved in the common overall fold of the LBDs. As a result, we have found that these local motifs are formed by conserved amino acids at particular positions, i.e., signal amino acids. In other words, the signal amino acids construct an individual local motif, and their assemblies determine the common overall fold of the LBD. In addition, we discuss the NR 'agonism' and 'antagonism' in terms of the stability of three local motifs in the AF-2 architecture, i.e., the AF-2 fixed motifs.

2. Material and methods

2.1. Alignment analyses

All 41 (168 in Supplementary Fig. S1) sequences examined in this study were manually aligned using both the signal and the conserved amino acid residues of the LBDs using our program (Tsuji, M., Homology Modeling Professional for HyperChem, Institute of Molecular Function, Saitama, Japan) (Tsuji, 2007). These signal amino acids have in part been reported by Wurtz and co-workers (Wurtz et al., 1996). Other signal amino acids were obtained by structural alignment (see the Extraction of Conserved Local Motifs). Sequence alignment was performed for individual NR classes (discussed in the Results section). The class 1 NRs were aligned for the sequence of hRXRα LBD, and the class 2 NRs were aligned for that of hRARγ LBD. Because the signal amino acids were distributed at suitable positions over the full-length LBD and were primarily located at the terminal positions of the secondary structures, these alignments could be used for exclusive identification. The results of the sequence alignments are summarized in Fig. 2 (see also Supplementary Fig. S1).

2.2. Extraction of conserved local motifs

The crystallographically solved 3D structures of the NR LBDs registered in the Protein Data Bank (PDB) (Berman et al., 2000) database were analyzed using Homology Modeling Professional for HyperChem. Interaction analyses (partners, distances, and angles) were carried out for all of the residues of each 3D structure.