Design and in vitro characterization of RXR variants as tools to investigate the biological role of endogenous rexinoids

Albane E Le Maire, Martial Rey, Valérie Vivat, Laura Guée, Pauline Blanc, Christian Malosse, Julia Chamot-Rooke, Pierre Germain, William Bourguet

Society for Endocrinology
JOURNAL OF MOLECULAR ENDOCRINOLOGY

Design and in vitro characterization of RXR variants as tools to investigate the biological role of endogenous rexinoids

Journal:	*Journal of Molecular Endocrinology*
Manuscript ID	JME-22-0021.R1
mstype:	Research paper
Date Submitted by the Author:	24-Jun-2022
Complete List of Authors:	le Maire, Albane; CNRS, Centre de biologie structurale Rey, Martial; Pasteur Institute, Biologie Structurale et Chimie Vivat, Valérie; IGBMC, Génomique fonctionnelle et cancer Guée, Laura; Inserm, Centre de biologie structurale Blanc, Pauline; Inserm, Centre de biologie structurale Malosse, Christian; Institut Pasteur, Biologie Structurale et Chimie Chamot-Rooke, Julia; Institut Pasteur, Biologie Structurale et Chimie Germain, Pierre; Inserm, Centre de biologie structurale Bourguet, William; Inserm, Centre de Biologie Structurale
Keywords:	Nuclear receptors, Retinoids, Ligand, Structure/function, Protein design

Design and in vitro characterization of RXR variants as tools to investigate the biological role of endogenous rexinoids

Albane le Maire[1], Martial Rey[2], Valérie Vivat[3,*], Laura Guée[1], Pauline Blanc[1], Christian Malosse[2], Julia Chamot-Rooke[2], Pierre Germain[1], and William Bourguet[1]

[1]CBS (Centre de Biologie Structurale), Univ Montpellier, CNRS, Inserm, Montpellier, France

[2]Institut Pasteur, Université de Paris, CNRS USR2000, Mass Spectrometry for Biology Unit, 75015 Paris, France

[3]IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Univ Strasbourg, CNRS, Inserm, Illkirch, France

Correspondence to P Germain or W Bourguet: germain@cbs.cnrs.fr or bourguet@cbs.cnrs.fr

***Current address:** Flare Therapeutics, 215 First Street, Suite 150, Cambridge, MA 02142, USA

Keywords: Nuclear receptors, ligand, retinoids, structure-function, protein design

Short Title: Design and characterization of RXR-based tools

Abstract

Retinoid X receptors (RXRα, β, and γ) are essential members of the nuclear receptor (NR) superfamily of ligand-dependent transcriptional regulators that bind DNA response elements and control the expression of large gene networks. As obligate heterodimerization partners of many NRs, RXRs are involved in a variety of pathophysiological processes. However, despite this central role in NR signaling, there is still no consensus regarding the precise biological functions of RXRs and the putative role of the endogenous ligands (rexinoids) previously proposed for these receptors. Based on available crystal structures, we introduced a series of amino acid substitutions into the ligand-binding pocket of all three RXR subtypes in order to alter their binding properties. Subsequent characterization using a battery of cell-based and *in vitro* assays led to the identification of a double mutation abolishing the binding of any ligand while keeping the other receptor functions intact, and a triple mutation that selectively impairs interaction with natural rexinoids but not with some synthetic ligands. We also report crystal structures that help understand the specific ligand-binding capabilities of both variants. These RXR variants, either fully disabled for ligand binding or retaining the property of being activated by synthetic compounds, represent unique tools that could be used in future studies to probe the presence of active endogenous rexinoids in tissues/organs and to investigate their role *in vivo*. Last, we provide data suggesting a possible involvement of fatty acids in the weak interaction of RXRs with corepressors.

Introduction

The three retinoid X receptor subtypes (RXRα, β, and γ; NR2B1-3 according to the official nomenclature) are essential members of the nuclear receptor (NR) superfamily acting as obligatory heterodimerization partners of many family members (Evans and Mangelsdorf 2014; Germain, et al. 2006; Gilardi and Desvergne 2014). Similar to other NRs, RXRs exhibit a modular structure with several domains, notably a central DNA-binding domain (DBD), and a C-terminal ligand-binding domain (LBD). Numerous 3D structures of NR LBDs, including in the context of the nearly full-length receptor (Chandra, et al. 2017), have corroborated the multifunctional nature of this domain, which controls not only ligand binding into the ligand-binding pocket (LBP) and receptor dimerization but also the interaction with transcriptional corepressors or coactivators through a ligand-dependent activation function (Weikum, et al. 2018).

RXRs are expressed in virtually every tissue of the body but disparities are observed in their distribution (Germain et al. 2006). RXRα is predominantly expressed in skin, epidermis, intestine, liver, lung, and kidney. RXRβ is ubiquitously expressed, whereas RXRγ is mostly restricted to the skeletal muscle and certain part of the brain. By virtue of their ability to form heterodimers with other NRs such as retinoic acid (RARs), thyroid hormone (TRs), peroxisome proliferator-activated (PPARs), or vitamin D3 (VDR) receptors, RXRs control a variety of pathophysiological processes (Evans and Mangelsdorf 2014; Mark, et al. 2006; Mark, et al. 2015). The central role of RXRs makes it a potential therapeutic target for treatment of diseases like cancer and metabolic diseases (Dawson and Xia 2012). Accordingly, major research efforts have been directed towards the identification of potent synthetic molecules acting as modulators of RXR transcriptional activity. They led to the development of several classes of compounds with a panel of activities ranging from agonists to antagonists through partial agonists, with some of them even used in clinical trials (Dominguez, et al. 2017; Perez, et al. 2012; Schierle and Merk 2019).

Although RXRs occupy a central position in NRs signaling, the existence of RXR endogenous ligands and the putative physiological functions of such ligands have remained a matter of debate. Initial studies led to the discovery of 9-*cis*-retinoic acid (9CRA) as a ligand for all three RXR and RAR subtypes (Heyman, et al. 1992; Levin, et al. 1992). But, except for the pancreas, it is not detected in cells

(Arnold, et al. 2012; Jones, et al. 2015). A number of fatty acids and retinoic acid analogues (**Fig. 1**) have since been proposed to be physiologically relevant endogenous ligands of RXRs, including phytanic acid (Kitareewan, et al. 1996), methoprene acid (Harmon, et al. 1995), and a series of unsaturated fatty acids such as docosahexaenoic acid (DHA) (de Urquiza, et al. 2000; Goldstein, et al. 2003; Lengqvist, et al. 2004; Niu, et al. 2017). However, their status as *bona fide* endogenous RXR ligands could be questioned with regard to their low abundance, very low affinity for RXR and/or weak transactivation capacity, or because they were identified in very specific tissues. The saturated analogue of 9CRA, 9-*cis*-13,14-dihydroretinoic acid (9CDHRA) present at high concentrations in mice, particularly in the serum and liver (Ruhl, et al. 2015), was recently proposed as an endogenous physiologically relevant mammalian RXR ligand. Note that 9CDHRA also activates RARs at high concentration (de Lera, et al. 2016; Krezel, et al. 2021; Krzyzosiak, et al. 2021; Ruhl, et al. 2018; Ruhl et al. 2015). *In vivo* data in mice suggest that the molecule could be involved in the control of stress-adaptation and depressive-like behaviors through RXR-mediated signaling (Krzyzosiak et al. 2021; Ruhl et al. 2015). However, the putative role of 9CDHRA in a broader physiological context has not been explored yet, so the questions regarding the existence, the biological function and the target tissues of a universal RXR physiological ligand still remain.

Based on structural information and using biochemical and cell-based assays, we designed two RXR variants that will help answer several crucial questions regarding its endogenous ligand(s). Does RXR principally serve as a heterodimerization partner allowing DNA binding specificity and signaling through the partner receptors or does it require natural activating ligands to achieve its physiological functions? What are the tissues/organs in which an endogenous RXR ligand is functionally required? What are the biological functions of endogenous rexinoids? Prototypical rexinoids (*i.e.* RXR synthetic or natural ligands) proposed so far to act as endogenous ligands contain a carboxylic head group involved in a salt bridge with an arginine residue of the helix H5 in the RXR LBD, and a long aliphatic chain making numerous van der Waals contacts with essentially non-polar residues distributed all over the LBP of RXR (see examples in **Fig. 1**). In contrast, synthetic rexinoids generally harbor aromatic rings, notably one bearing the carboxylic moiety such as in bexarotene, BMS649 or LG268 (**Fig. 1**). By introducing a series of amino acid substitutions into the LBP of RXRs, we identified a variant with fully

impaired ligand-binding activity (hereafter termed 2mRXRs), and a second one responding to some synthetic ligands only (hereafter termed 3mRXRs). Extended structural and functional analyses of these RXR variants revealed the structural bases for their novel ligand-binding properties, and showed that the other functions of the receptor, including DNA-binding and dimerization, remain mostly unchanged. We propose that mouse lines expressing these RXR variants that have partially (the transcriptional activity of 3mRXR can be rescued by synthetic compounds) or totally (2mRXR) lost their ligand-binding capacity would represent unique tools to probe the pathophysiological function(s) and the site(s) of action of endogenous RXR ligand(s) in the whole animal. We also provide data suggesting a possible role of fatty acids in the weak interaction between RXRs and corepressors.

Materials and Methods

Ligands and peptides

Ligands (CD3254, BMS649, LG268, Bexarotene and TTNPB were purchased from Tocris Bioscience. 9CRA, DHA, phytanic acid and methoprene acid were purchased from Sigma-Aldrich. 9CDHRA was a kind gift from Angel de Lera (University of Vigo, Spain). All compound stock solutions were prepared at 10 mM in DMSO. The SRC-1 peptide labelled with FITC (FITC-LTERHKILHRLLQEGSP), the TIF-2 peptide (KHKILHRLLQDSS) and the corepressor NCoR peptide labelled with FITC (FITC-DPASNLGLEDIIRKALMGSFD) were purchased from EZbiolab.

Plasmids

The DNA encoding mouse wtRXRα, 2mRXRα, 3mRXRα, mouse wtRXRβ, 2mRXRβ, 3mRXRβ, and mouse wtRXRγ, 2mRXRγ, 3mRXRγ were cloned (InFusion cloning kit) into a pSG5 puro vector between the BamHI and BglII restriction sites.

Transactivation assays

COS-1 cells were grown in DMEM with Glutamax and 10% (v/v) FCS, and co-transfected with the DR1-tk-Luc reporter gene and the expression vector of one of the different RXRs studied using JetPei transfectant (Ozyme). After 24 h, the medium was changed to a medium containing the indicated ligands

or vehicle. Cells were lysed and assayed for reporter expression 48 h after transfection. The luciferase assay system was used according to the manufacturer's instruction (Promega). In each case, results were normalized to coexpressed β-galactosidase. Each transfection was performed in duplicate and repeated three to six times.

Preparation of RXRα variants for *in vitro* studies

For the preparation of 2mRXRα and 3mRXRα variants, the DNA encoding mouse RXRα including the two (R321E/C437Q) and the three (I315A/F318A/A332T) mutations, respectively, optimized for bacterial expression, purchased to Integrated DNA Technologies were cloned using InFusion cloning kit into a pDB vector between the NdeI and XhoI restriction sites. pDB-2mRXRα and pDB-3mRXRα constructs code for proteins containing a hexa histidine-tag and a HRV 3C cleavage site (Leu-Glu-Val-Leu-Phe-Gln/Gly-Pro) at the N terminus, where a specific cleavage occurs between Gln and Gly.

Expression and purification of proteins

The wtRXRα LBD protein was purified as previously described (Nahoum, et al. 2007). The 2mRXRα and 3mRXRα proteins were overproduced in *E. coli* BL21 (DE3) grown at 37°C until OD600nm reaches 0.6 and overnight at 20° (for 2mRXRα) or for additional three hours at 37°C (for 3mRXRα), after expression induction with 0.5 mM IPTG. Cells were harvested and resuspended in 500 mM NaCl, 50 mM Tris pH 7.5, plus one tablet of Complete EDTA free protease inhibitor cocktail and lysozyme at 10 µg/ml and sonicated on ice. The lysate was centrifuged at 18,000 rpm at 4 °C for 30 min. The supernatant was clarified and loaded on a 5 ml HisTrap Column (Cytiva) equilibrated with 500 mM NaCl, 50 mM Tris pH 7.5, 10 mM imidazole. The column was washed with the latter buffer during 10 column volumes (CV) and with 500 mM NaCl, 50 mM Tris pH 7.5, 50 mM imidazole during 10 CV and the protein was eluted with 500 mM NaCl, 50 mM Tris pH 7.5, 250 mM imidazole. Fractions containing the protein were pooled and incubated with 3C protease overnight. The cleaved protein was concentrated and loaded on a HiLoad 26/60 Superdex 75 PG (GE Healthcare) equilibrated with 150 mM NaCl, 20 mM Tris pH 7.5, 10% glycerol and 2mM DTT. The pure fractions containing proteins were pooled and concentrated

by centrifugation using a Vivaspin™ protein concentrator with a 10 kDa cut-off until concentration, determined by UV absorbance at 280 nm, reaches 8 mg/mL.

Size-exclusion chromatography-multi angle light scattering (SEC-MALS)

SEC-MALS experiment was performed at 25°C using a Superdex 200 10/300 GL column (GE HealthCare) connected to a miniDAWN-TREOS light scattering detector and an Optilab T-rEX differential refractive index detector (Wyatt Technology, Santa Barbara, CA). The column was equilibrated with 0.1 m filtered 50mM Tris pH7.5, 150mM NaCl buffer. A sample of 25 µL of RXRα protein (wtRXRα, 3mRXRα or 2mRXRα LBDs), alone or pre-equilibrated with an equal molar excess of purified TRα1 LBD, at 1 mg/mL (around 50 µM) was injected at 0.5 mL/min. Data acquisition and analyses were performed using the ASTRA software (Wyatt). Based on measurement on BSA sample under the same conditions, we have estimated an experimental error in molar mass around 5%.

Thermal shift assay (TSA)

Solutions of 25 µL containing 5 µM of RXRα protein, in the presence of DMSO (condition without ligand) or in the presence of three molar excess of ligands, and 1X Sypro® Orange in 50 mM Tris pH 8.0, 200 mM NaCl, 5% glycerol were added to the wells of a 96-well PCR plate. The plates were sealed with an optical sealing tape (Bio-Rad) and heated in a 7500 Real Time PCR system (Applied Biosystem) from 25 to 95°C at 1°C intervals. Fluorescence changes in the wells were monitored with a photomultiplier tube. The wavelengths for excitation and emission were 545 nm and 568 nm, respectively. The melting temperatures (Tm) were obtained by fitting the fluorescence data with a Boltzmann model using the GraphPad Prism software.

Steady-state fluorescence anisotropy

Measurements of the binding affinities of the fluorescent SRC-1 NR2 peptide for RXRα LBDs in the presence or absence of ligands was performed using a Safire2 microplate reader (TECAN). The excitation wavelength was set at 470 nm and emission was measured at 530 nm. The buffer solution for assays was 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, and 10% (v/v) glycerol.

The measurements were initiated at 10 µM of protein, and the sample was then diluted successively by a factor of 2 with the buffer until the lowest protein concentration at 2.4 nM. Fluorescent peptides were added to protein samples at 4 nM, allowing establishment of the titration curve. Ligands, when added, were at a final concentration of three molar excess relative to the highest concentration of protein. The reported data are the average of at least three independent experiments, and error bars correspond to standard deviations.

Native mass spectrometry

Proteins were mixed with each ligand at a 1:5 molar ratio in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT. Protein-ligand samples were then buffer exchanged twice using a biospin 6 (Bio-Rad) against 200 mM ammonium acetate buffer. Then protein-ligand complexes were nanoelectrosprayed in native conditions using a Triversa NanoMate (Advion) coupled to a Tribrid Eclipse Orbitrap mass spectrometer (ThermoFisher Scientific). Spectra were recorded at 15K resolution (at m/z 200) with 20 microscans per scan using the intact protein mode in normal pressure mode. The In-Source Collision Induced Dissociation (IS-CID) energy was varied from 0 V to 60 V with 10 V steps to promote dissociation of the ligand from the protein. Ratios between the ligand-bound and free monomeric protein (10+ ion) were measured for each IS-CID energy and compared for each ligand. Spectra were acquired between 600 and 3000 m/z.

Crystallization and structure resolution

The 2mRXRα and 3mRXRα proteins concentrated at 8 mg/ml were mixed with a five-fold molar excess of LG268 or BMS649, respectively, and a three-fold molar excess of the TIF-2 NR2 peptide. We mixed the complexes in a 1:2 ratio with reservoir solution consisting of 26% (w/v) PEG 3350, 0.2 M NaF or 18% (w/v) PEG 3350, 0.2 M Na Nitrate for 2mRXRα and 3mRXRα complexes, respectively, and crystallized the complexes using the sitting-drop vapor diffusion technique. We mounted a single crystal from mother liquor onto a cryoloop, soaked in the reservoir solution containing an additional 25% (v/v) glycerol and flash-frozen it in liquid nitrogen. We collected diffraction data at the PXI beamline of SLS at 1.58 Å resolution for 2mRXRα complex and at XALOC beamline of ALBA at 3.30 Å resolution for

3mRXRα complex. We processed diffraction data using XDS (Kabsch 2010) and scaled them with SCALA from the CCP4 program suite (Winn, et al. 2011). We obtained the initial phases by molecular replacement with Phenix (Liebschner, et al. 2019) using the structure of the wt RXRα in complex with BMS649 (Egea, et al. 2002) (PDB 1MVC). Initial Fo – Fc difference maps had strong signal for the ligand and for the peptide, which we could fit accurately into the electron density. We built the model with Coot (Emsley and Cowtan 2004) and refined it with Phenix. The RXRα-RARα LBD heterodimer was purified and crystallized as described previously (Bourguet, et al. 2000a). The structure was solved by molecular replacement using the structure of the mutant RXRαF318A-RARα LBD heterodimer (PDB 1dkf). (Bourguet, et al. 2000b). Data collection and refinement statistics are summarized in **Supplementary Table 1**. Figures were prepared with PyMOL (http://pymol.org/).

Results

Design of RXR variants

Guided by the many RXR LBD crystal structures available and our structure of the RXRα-RARα LBD heterodimer where the RXRα and RARα protomers are bound to a fortuitous bacterial fatty acid and the antagonist BMS614, respectively (**Supplementary Fig. 1**), we introduced a series of mutations in the mouse RXRα LBP with the aim to disable ligand-binding without altering other receptor functions (**Fig. 2A**). Note that there is no species or subtype difference in the amino acid composition of RXR LBP so that the results obtained here on mouse RXRα can be directly extended to all mouse and human receptor subtypes. Single and multiple substitutions were introduced in wild-type (wt) RXRα so as to remove essential ligand/receptor interactions or to generate steric hindrances precluding the formation of a stable ligand/RXRα complex. Two RXRα variants for which full functional and structural characterizations are provided below appeared of particular interest to us with regard to their specific ligand binding profiles. One, referred to as 2mRXRα, contains two mutations, namely R321E and C437Q (mouse numbering). Located in helix H5, R321 plays a key role in forming a salt bridge with the carboxylic head group of rexinoids (**Figs 1** and **2A**) so that its replacement by a glutamic acid residue was expected to reduce ligand binding affinity. On the other side of the LBP (helix H11), the small cysteine residue to larger glutamine amino acid substitution was introduced so as to generate steric hindrance and further decrease ligand binding. This second mutation was chosen based on previous work showing that the mutation of human RXRα C432 (equivalent to C437 in mouse) into bulky aromatic residues (*i.e.* tryptophan or tyrosine) resulted in constitutively active RXRα variants, while the mutation into a glutamine residue had no impact on the basal activity of the receptor but was detrimental to 9CRA binding (Zeng, et al. 2015). The second RXRα variant, referred to as 3mRXRα, contains three mutations, namely I315A, F318A, and A332T (**Fig. 2A**). The two small alanine residues were introduced at positions 315 and 318 to remove the essential hydrophobic stabilizing contacts provided by the bulky isoleucine and phenylalanine amino acids, respectively. Last, located at the tip of the S1/S2 β-turn and in very close proximity to the carboxylic head group of rexinoids, alanine residue 332 was substituted for a slightly larger threonine amino acid with the goal to negatively interfere with ligand binding. A

battery of cell-based and *in vitro* assays was subsequently used to perform a full characterization of the ligand responsiveness of these RXRα variants.

Functional characterization of the 2m- and 3mRXR variants

The ligand-binding function is fully impaired in 2mRXRs

First, reporter gene assays were conducted in COS-1 cells to assess the transcriptional activity of wt and RXRα variants in the presence of various synthetic and natural compounds previously proposed as RXRs endogenous ligands. As shown in **Fig. 2B**, wtRXRα could be potently activated by LG268 and CD3254, slightly less by bexarotene, and to a lower extent by BMS649, 9CRA and finally 9CDHRA. Note that 9CDHRA was almost as active as 9CRA at 10 μM, whereas all the other natural compounds displayed no activity at this concentration (**Supplementary Fig. 2A**). A similar experiment using the 2mRXRα variant revealed that none of the compounds tested were able to induce a substantial transcriptional response below 1μM, although little RXRα activation could be observed with LG268 at 1 μM (**Fig. 2C**). As expected, the mutations produced the same functional alteration when introduced in RXRβ and RXRγ (**Supplementary Fig. 2B**).

Using fluorescence anisotropy measurements, we then evaluated the ability of the various ligands to induce the recruitment by wtRXRα and 2mRXRα of a fluorescein-labeled peptide derived from the coactivator SRC-1 and encompassing one LxxLL nuclear receptor interacting motif (hereafter referred to as SRC-1). Consistent with transactivation assays, the synthetic ligands and 9CRA were able to strongly increase the affinity of wtRXRα for SRC-1, whereas 9CDHRA or phytanic acid (PA) strengthened the interaction only slightly (**Fig. 2D**). Results obtained with 2mRXRα were also in line with cell-based data as all the compounds used failed to induce a robust recruitment of the coactivator-derived peptide (**Fig. 2E**).

Because the aforementioned data suggested that 2mRXRα is unable to bind any kind of ligand, we evaluated the direct interaction of the compounds with wtRXRα and 2mRXRα using thermal shift assay (TSA) and differential scanning fluorimetry (nanoDSF) that monitor the thermal stabilization of the receptor upon ligand binding. In full agreement with cell-based and fluorescence assays, all the synthetic ligands were able to induce an increase of the melting temperature (Tm) of wtRXRα, whereas every

natural compound but 9CRA, and to a lesser extent 9CDHRA, failed to do so (**Fig. 2F** and **Supplementary Fig. 2C**). Regarding 2mRXRα stabilization, none of the synthetic or natural compounds induced a stabilization of the protein, even at high ligand/protein ratios (**Fig. 2G** and **Supplementary Fig. 2D**). Although 2mRXRα and wtRXRα have similar secondary structure content (**Supplementary Fig. 3**), the decreased Tm observed for the unliganded 2mRXRα relative to that of wtRXRα could be partly due to a slight destabilizing effect of the mutations, but more likely to the fact that the bacterially expressed wtRXRα LBD co-purifies, and co-crystallizes (**Fig. 2A**), with a fatty-acid of bacterial origin that stabilizes the protein domain. Such binding is precluded in 2mRXRα. Last, the interaction of wtRXRα and 2mRXRα with the synthetic compounds BMS649, LG268 and CD3254 was characterized by using native mass spectrometry. Native mass spectrometry (MS) has the capability to directly detect ligand binding to protein, determine stoichiometry, relative binding affinities and specificity (Heck 2008). As expected from the previous experiments, incubation of the rexinoids with wtRXRα resulted in the formation of binary protein-ligand complexes, but not with the 2mRXRα variant (**Fig. 3A, B** and **Supplementary Fig. 4**). Note that nonspecific binding outside the RXR LBP may account for the small amount of binary complex detected with 2mRXRα and CD3254, as a very similar signal was obtained with 3mRXRα (**Fig. 3B, C**).

As a whole, these data show that while synthetic ligands act as potent wtRXRα agonists with EC50s ranging from 3 nM to 58 nM, all the natural compounds but 9CRA (EC50 of 130 nM) and 9CDHRA (1 μM < EC50 < 10 μM) exert weak activation properties, if any (EC50s > 10 μM). They also demonstrate that 2mRXRs have lost most of their capabilities to bind natural and synthetic rexinoids.

3mRXRs respond exclusively to benzoic acid-containing synthetic compounds

As with 2mRXRα, the cell-based, *in vitro* assays and native MS experiments described above were applied to 3mRXRα in order to reveal its ligand responsiveness. Reporter gene assays in COS-1 cells revealed that LG268, bexarotene and BMS649 retained a strong agonistic activity, whereas all the natural compounds, including 9CRA, and unexpectedly the synthetic compound CD3254 did not induce any measurable transcriptional response, up to a concentration of 1 μM (**Fig. 4A**). As expected, the mutations produced the same effects when introduced in RXRβ and RXRγ (**Supplementary Fig. 5A**).

Similar observations were made in coactivator recruitment experiments using fluorescence anisotropy where only the three synthetic compounds LG268, bexarotene and BMS649 were able to increase the affinity of 3mRXRα for SRC-1 (**Fig. 4B**). CD3254 and all the natural compounds failed to induce a recruitment of the coactivator-derived peptide above that obtained with unliganded 3mRXRα. Tm measurements in presence of the various compounds confirmed that all the synthetic compounds but CD3254 interact with 3mRXRα, thereby increasing its thermal stability. In contrast, none of the natural compounds tested bind to this RXRα variant (**Fig. 4C** and **Supplementary Fig. 5B**). Last, MS analysis confirmed that in contrast to wtRXRα, 3mRXRα is unable to form a stable binary complex with CD3254, whereas a good interaction with BMS649 and LG268 is preserved (**Fig. 3C**).

Interestingly, a close look at the chemical differences between 3mRXRs binders (LG268, bexarotene and BMS649) and non-binders (natural compounds and CD3254), revealed that members of the latter group of compounds contain an aliphatic chain next to the carboxylate moiety, whereas compounds belonging to the other group harbor an aromatic benzoic acid instead (**Fig. 1**). Altogether, these functional data show that we could identify (i) RXRα, β, γ variants whose ligand-binding function is fully impaired (2mRXRs), and (ii) RXRα, β, γ variants that fail to bind natural compounds, whose hallmark is the presence of an aliphatic chain bearing the carboxylate moiety, but can be activated by synthetic compounds harboring a benzoic acid group (3mRXRs).

2mRXRs and 3mRXRs retain their DNA-binding and dimerization capabilities

In addition to providing information regarding the ligand responsiveness of wtRXRα and its variants, the transactivation assays, that uses full-length receptors and the direct repeat 1 (DR1) response element of RXR homodimers, also reveals their DNA-binding and homodimerization capabilities. Mutations located in the LBP of RXRs are unlikely to impact these functions, and accordingly, the DNA- and dimerization-dependent signals observed for 3mRXRα in COS-1 cells with some synthetic ligands (**Fig. 4A**) and to a lesser extent for 2mRXRα with LG268 (**Fig. 2C**) suggested preserved DNA-binding and self-association functions in RXR variants.

Having shown the ligand-binding specificities and the DNA-binding and homodimerization capabilities of the two RXR variants, we then evaluated the impact of the mutations on the ability of

2mRXRα and 3mRXRα to form heterodimers. To this end, the oligomeric states of wtRXRα, 2mRXRα and 3mRXRα LBDs, alone and in complex with TRα1 LBD, were determined using electrophoresis under non-denaturing conditions (**Fig. 5A**). For individual RXRα proteins, two predominant bands corresponding to monomers and homotetramers were observed. An additional band indicating the presence of homodimers was apparent with 3mRXRα. In the presence of the heterodimerization partner TRα1, bands of similar intensities corresponding to heterodimers were clearly visible for wtRXRα, 2mRXRα and 3mRXRα. Independent confirmation of the heterodimerization abilities of wtRXRα and the RXRα variants was brought by size-exclusion chromatography-multi angle light scattering (SEC-MALS) analysis (**Fig. 5B**). All three RXRα/TRα1 mixtures eluted similarly at the elution volume of heterodimers, in addition to the elution peak corresponding to monomers.

As expected, these results show that the mutations do not modify the DNA-binding and dimerization properties of 2mRXRα and 3mRXRα, and confirm that heterodimerization does not require the presence of RXR ligand to occur.

Structural characterization of the 2m- and 3mRXRα variants

To gain insights into the structural basis of RXR variants ligand-binding specificities, we solved the crystal structures of the ternary complexes containing the 2mRXRα or 3mRXRα LBDs bound to a coactivator (TIF-2)-derived peptide and the agonists LG268 or BMS649, respectively (**Supplementary Table 1**). Both structures show the canonical active conformation with the C-terminal activation helix H12 capping the LBP and the TIF-2 peptide bound to the surface formed by residues from helices H3, H4 and H12 (**Supplementary Fig. 6**). The ligands LG268 and BMS649 could be unambiguously positioned in their respective electron densities (**Supplementary Fig. 6**). Importantly, our ability to obtain these structures is an additional proof that the introduced mutations do not alter proper protein folding.

Structure of LG268-bound 2mRXRα

Although our functional analysis demonstrated that the ligand-binding capacity of 2mRXRα is largely impaired, the high protein and ligand concentrations used during the crystallization process allowed

formation of the complex. Comparison of the structure with that of human wtRXRβ in complex with LG268 (PDB 1H9U) (Love, et al. 2002) shows that the two LG268 ligands superimpose well with no major repositioning (**Fig. 6A**). In contrast to wtRXRβ R387 (equivalent to R321 in mouse RXRα) that points towards the LBP to form a salt bridge with the carboxylate moiety of the ligand, E321 in 2mRXRα turns around and points to the solvent. Water molecules occupy the space left vacant by the R321E mutation. In helix H11, Q437 that replaces a cysteine residue in wtRXR resides at a van der Waals distance to LG268 and appears to be well accommodated following a few side chain rearrangements, notably L438 and L441 (**Fig. 6A**). Clearly, the R321E mutation is the major contributor to the binding impairment of LG268 to 2mRXRα. However, when superimposing the 2mRXRα structure with that of human wtRXRα bound to 9CRA (PDB 1FBY) (Egea, et al. 2000), one can see that the latter occupies more space in the helix H11 region than the synthetic compounds so that the presence of a large glutamine residue at the position normally occupied by a small cysteine amino acid generates a severe steric clash with the ligand (**Fig. 6B**). A similar situation can be seen with several ligands, including CD3254 (PDB 3FUG) (Perez Santin, et al. 2009) (**Fig. 6C**) or 9CDHRA (PDB 4ZSH) (Ruhl et al. 2015) (**Fig. 6D**). Overall it appears that both R321E and C437Q mutations play a role in the ligand-binding disability of 2mRXRα, their relative involvements varying according to the chemical structure of the compounds.

*Structure of BMS649-bound 3mRXR*α

Our functional analysis demonstrated that benzoic acid-containing synthetic compounds such as BMS649 can bind to and activate 3mRXRα. Comparison of the structure with that of human wtRXRα in complex with BMS649 (PDB 1MVC) (Egea et al. 2002) shows that the BMS649 molecules in the two structures are not fully superimposable. Both the position and the conformation of the ligands are slightly different in the two structures (**Fig. 7A**). Due to the steric hindrance generated by the A332T mutation, and to the space made available by the I315A and F318A mutations, BMS649 in 3mRXRα is shifted towards helices H5, H7 and H11, and the benzoic acid group is tilted by roughly 50° relative to its orientation in wtRXRα. This rotation appears essential to maintain important stabilizing interactions between the ligand and surrounding residues from helices H3 and H5 (**Fig. 7B**). Such adaptation is not

possible for natural compounds and CD3254 where the phenyl ring of the benzoic acid group is replaced by a rigid aliphatic chain. Furthermore, by superimposing the 3mRXRα structure with that of human wtRXRα bound to 9CRA (PDB 1FBY) (Egea et al. 2000), we also noticed that the latter is longer than BMS649 and the other synthetic compounds except CD3254 (**Fig. 7C**). This length is incompatible with the 1.5 Å ligand shift towards helices H7 and H11 measured for BMS649 and imposed by the A332T mutation present in 3mRXRα. Indeed, ligands like 9CRA, 9CDHRA, CD3254 or DHA are rather large and markedly more constrained within the wtRXRα LBP than bexarotene, LG268 or BMS649. In conclusion, our structural analysis reveals that 3mRXRα discriminates between natural and synthetic rexinoids on the basis of their sizes and chemical structures (*i.e.* aliphatic acids *vs* benzoic acids).

Discussion

At least one RXR subtype is present in every mammalian cell and participates in diverse essential physiological processes by heterodimerizing with fifteen different nuclear receptors (Evans and Mangelsdorf 2014; Gilardi and Desvergne 2014). Nevertheless, it is still unknown if RXR actually needs a ligand to fulfill its biological roles in all cell types, in a subset of tissues only, or if it primarily serves as a ligand-independent dimerization partner. No conclusive study has been reported to date on the ligand dependency of RXR because none of them has used a mutated form of RXR in which the ligand-binding function of the receptor has been specifically disabled. Many *in vivo* data in mouse, including from RXR knock-outs (Kastner, et al. 1994; Kastner, et al. 1996) or from the less drastic mutation where the RXRα C-terminal helix H12 (activation helix required for ligand-dependent transcription) was deleted (Mascrez, et al. 1998), have provided insight into the physiological functions of these receptors. Owing to the numerous developmental abnormalities observed in the latter study, it was concluded that the RXR ligand-dependent transactivation, and implicitly the involvement of an endogenous ligand, is instrumental during development (Mascrez et al. 1998). However, later studies demonstrated that deletion of the activation helix of RXR induces a strong interaction with corepressors whose binding site is most likely masked by helix H12 in the wild-type receptor (Hu, et al. 2001; Zhang, et al. 1999). Thus, rather than the abolishment of the RXR ligand-dependent activation function, the enhanced recruitment of corepressors by the RXRα-RAR heterodimer (under investigation in the study), and its

consecutive transcriptional repression are most likely at the origin of the phenotypes observed in the mutant mouse.

In this regard, our structure of the fatty acid-bound wtRXRα-RARα heterodimer (**Supplementary Fig. 1**) reveals a receptor conformation in which the helix H12 of RXR is docked in the groove serving as a primary binding site of corepressors (le Maire and Bourguet 2014; le Maire, et al. 2010). The structure provides a rationale for the poor interaction of wtRXRs with corepressors, and its enhancement upon helix H12 deletion (Hu and Lazar 1999; Hu et al. 2001; Zhang et al. 1999). Accordingly, an increase in corepressor interaction was observed when two mutations (M454A/L455A, human numbering, **Supplementary Fig. 1**), intended to destabilize the interaction between helix H12 and the corepressor binding site, were introduced in the RXR sequence (**Supplementary Fig. 7A**). In addition, 2mRXRα and 3mRXRα that are unlikely to copurify with a bacterial fatty acid were shown to interact more avidly than wtRXRα with a fluorescein-labeled peptide containing the LxxIIxxxL interacting motif of the corepressor NCoR (**Supplementary Fig. 7B**). Further investigations will be necessary to validate these data suggesting a possible role of fatty acids in the weak interaction between RXRs and corepressors by inducing a stabilization of helix H12 in the corepressor interacting groove.

A fundamental question that remains is whether an endogenous ligand is actually required for RXR knowing that the transactivation function of all RXR heterodimers can be finely tuned by cognate ligands of the partner receptors. In this regard, while some heterodimers can be activated by ligands of the partner receptor, but not by RXR ligands alone (*e.g.* RXR-RAR), other heterodimers are permissive and allow activation by RXR ligands in the absence of a partner agonist (*e.g.* RXR-PPAR) (Germain, et al. 2002; Kojetin, et al. 2015; le Maire, et al. 2019; Shulman, et al. 2004). In a physiological context such RXR-mediated activation by endogenous ligands could complicate the precise regulation of permissive heterodimers and generate uncontrolled signaling pathway promiscuity.

To our surprise, all the natural compounds used in this study but 9CRA, and to a lesser extent 9CDHRA, did not displayed measurable RXR agonistic activity in our cell-based and *in vitro* assays. One possible explanation for this observation is that these natural substances might be relevant in particular cellular environments or in some physiological conditions as reviewed in (Krezel, et al. 2019). Nevertheless, and whatever the nature of the ligand, we believe that the 2m- and 3mRXR variants

transposed in mouse models will be critical to evaluate the biological outcome of ligand binding alteration *in vivo* and should provide a wealth of information regarding the requirement of *bona fide* physiological RXR ligand(s), their target organs or tissues, and their pathophysiological importance. Interestingly, treatments of pregnant mothers or adult mice bearing the 3mRXR variant with BMS649 or LG268 could be performed with the aim to rescue the pathological outcome of the mutations.

Declaration of interest

The authors declare no conflict of interests.

Funding

The Centre de Biologie Structurale is a member of the France-BioImaging (FBI) and the French Infrastructure for Integrated Structural Biology (FRISBI), two national infrastructures supported by the French National Research Agency (ANR-10-INBS-04-01 and ANR-10-INBS-05, respectively). The native MS experiments were performed on an Eclipse Tribrid Orbitrap Mass Spectrometer funded by the Region Ile de France (DIM 1HEALTH).

Author contribution statement

AlM, PG and WB conceived the project. AlM and WB performed crystallographic analyses. VV and PG performed cell-based assays. AlM, LG and PB provided the purified proteins and performed biophysical analyses. AlM, PG and WB analyzed data. MR, CM and JCR performed mass spectrometry and data analyses. WB wrote the manuscript. All authors contributed to editing the manuscript.

Acknowledgements

We acknowledge the experimental assistance from the staff of the Swiss Light Source (SLS, Villigen, Switzerland), the Spanish synchrotron facility ALBA (Barcelona, Spain) and the European Synchrotron Radiation Facility (ESRF, Grenoble, France) during data collection.

Accession codes: The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 7PDT (3mRXRα), 7PDQ (2mRXRα), and 7QAA (RXRα-RARα).

References

Arnold SLM, Amory JK, Walsh TJ & Isoherranen N 2012 A sensitive and specific method for measurement of multiple retinoids in human serum with UHPLC-MS/MS. *J Lipid Res* **53** 587-598.

Bourguet W, Andry V, Iltis C, Klaholz B, Potier N, Van Dorsselaer A, Chambon P, Gronemeyer H & Moras D 2000a Heterodimeric Complex of RAR and RXR Nuclear Receptor Ligand-Binding Domains: Purification, Crystallization, and Preliminary X-Ray Diffraction Analysis. *Protein Expr Purif* **19** 284-288.

Bourguet W, Vivat V, Wurtz JM, Chambon P, Gronemeyer H & Moras D 2000b Crystal structure of a heterodimeric complex of RAR and RXR ligand-binding domains. *Mol Cell* **5** 289-298.

Chandra V, Wu D, Li S, Potluri N, Kim Y & Rastinejad F 2017 The quaternary architecture of RARβ-RXRα heterodimer facilitates domain-domain signal transmission. *Nat Commun* **8** 868.

Dawson MI & Xia Z 2012 The retinoid X receptors and their ligands. *Biochim Biophys Acta* **1821** 21-56.

de Lera AR, Krezel W & Ruhl R 2016 An Endogenous Mammalian Retinoid X Receptor Ligand, At Last! *ChemMedChem* **11** 1027-1037.

de Urquiza AM, Liu S, Sjoberg M, Zetterstrom RH, Griffiths W, Sjovall J & Perlmann T 2000 Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain. *Science* **290** 2140-2144.

Dominguez M, Alvarez S & de Lera AR 2017 Natural and Structure-based RXR Ligand Scaffolds and Their Functions. *Curr Top Med Chem* **17** 631-662.

Egea PF, Mitschler A & Moras D 2002 Molecular recognition of agonist ligands by RXRs. *Mol Endocrinol* **16** 987-997.

Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P & Moras D 2000 Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *Embo J* **19** 2592-2601.

506 Emsley P & Cowtan K 2004 Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol*
507 *Crystallogr* **60** 2126-2132.

508 Evans RM & Mangelsdorf DJ 2014 Nuclear Receptors, RXR, and the Big Bang. *Cell* **157** 255-266.

509 Germain P, Chambon P, Eichele G, Evans RM, Lazar MA, Leid M, De Lera AR, Lotan R, Mangelsdorf
510 DJ & Gronemeyer H 2006 International Union of Pharmacology. LXIII. Retinoid X receptors.
511 *Pharmacol Rev* **58** 760-772.

512 Germain P, Iyer J, Zechel C & Gronemeyer H 2002 Coregulator recruitment and the mechanism of
513 retinoic acid receptor synergy. *Nature* **415** 187-192.

514 Gilardi F & Desvergne B 2014 RXRs: collegial partners. *Subcell Biochem* **70** 75-102.

515 Goldstein JT, Dobrzyn A, Clagett-Dame M, Pike JW & DeLuca HF 2003 Isolation and characterization
516 of unsaturated fatty acids as natural ligands for the retinoid-X receptor. *Arch Biochem Biophys* **420** 185-
517 193.

518 Harmon MA, Boehm MF, Heyman RA & Mangelsdorf DJ 1995 Activation of mammalian retinoid X
519 receptors by the insect growth regulator methoprene. *Proc Natl Acad Sci U S A* **92** 6157-6160.

520 Heck AJ 2008 Native mass spectrometry: a bridge between interactomics and structural biology. *Nat*
521 *Methods* **5** 927-933.

522 Heyman RA, Mangelsdorf DJ, Dyck JA, Stein RB, Eichele G, Evans RM & Thaller C 1992 9-cis retinoic
523 acid is a high affinity ligand for the retinoid X receptor. *Cell* **68** 397-406.

524 Hu X & Lazar MA 1999 The CoRNR motif controls the recruitment of corepressors by nuclear hormone
525 receptors. *Nature* **402** 93-96.

526 Hu X, Li Y & Lazar MA 2001 Determinants of CoRNR-dependent repression complex assembly on
527 nuclear hormone receptors. *Mol Cell Biol* **21** 1747-1758.

528 Jones JW, Pierzchalski K, Yu J & Kane MA 2015 Use of fast HPLC multiple reaction monitoring cubed
529 for endogenous retinoic acid quantification in complex matrices. *Anal Chem* **87** 3222-3230.

530 Kabsch W 2010 Xds. *Acta Crystallogr D Biol Crystallogr* **66** 125-132.

531 Kastner P, Grondona JM, Mark M, Gansmuller A, LeMeur M, Decimo D, Vonesch JL, Dolle P &
532 Chambon P 1994 Genetic analysis of RXR alpha developmental function: convergence of RXR and
533 RAR signaling pathways in heart and eye morphogenesis. *Cell* **78** 987-1003.

534 Kastner P, Mark M, Leid M, Gansmuller A, Chin W, Grondona JM, Decimo D, Krezel W, Dierich A &
535 Chambon P 1996 Abnormal spermatogenesis in RXR beta mutant mice. *Genes Dev* **10** 80-92.

536 Kitareewan S, Burka LT, Tomer KB, Parker CE, Deterding LJ, Stevens RD, Forman BM, Mais DE,
537 Heyman RA, McMorris T, et al. 1996 Phytol metabolites are circulating dietary factors that activate the
538 nuclear receptor RXR. *Mol Biol Cell* **7** 1153-1166.

539 Kojetin DJ, Matta-Camacho E, Hughes TS, Srinivasan S, Nwachukwu JC, Cavett V, Nowak J, Chalmers
540 MJ, Marciano DP, Kamenecka TM, et al. 2015 Structural mechanism for signal transduction in RXR
541 nuclear receptor heterodimers. *Nat Commun* **6** 8013.

542 Krezel W, Rivas A, Szklenar M, Ciancia M, Alvarez R, de Lera AR & Ruhl R 2021 Vitamin A5/X, a
543 New Food to Lipid Hormone Concept for a Nutritional Ligand to Control RXR-Mediated Signaling.
544 *Nutrients* **13**.

545 Krezel W, Ruhl R & de Lera AR 2019 Alternative retinoid X receptor (RXR) ligands. *Mol Cell*
546 *Endocrinol* **491** 110436.

547 Krzyzosiak A, Podlesny-Drabiniok A, Vaz B, Alvarez R, Ruhl R, de Lera AR & Krezel W 2021 Vitamin
548 A5/X controls stress-adaptation and prevents depressive-like behaviors in a mouse model of chronic
549 stress. *Neurobiol Stress* **15** 100375.

le Maire A & Bourguet W 2014 Retinoic acid receptors: structural basis for coregulator interaction and exchange. *Subcell Biochem* **70** 37-54.

le Maire A, Teyssier C, Balaguer P, Bourguet W & Germain P 2019 Regulation of RXR-RAR Heterodimers by RXR- and RAR-Specific Ligands and Their Combinations. *Cells* **8**.

le Maire A, Teyssier C, Erb C, Grimaldi M, Alvarez S, de Lera AR, Balaguer P, Gronemeyer H, Royer CA, Germain P, et al. 2010 A unique secondary-structure switch controls constitutive gene repression by retinoic acid receptor. *Nat Struct Mol Biol* **17** 801-807.

Lengqvist J, Mata De Urquiza A, Bergman AC, Willson TM, Sjovall J, Perlmann T & Griffiths WJ 2004 Polyunsaturated fatty acids including docosahexaenoic and arachidonic acid bind to the retinoid X receptor alpha ligand-binding domain. *Mol Cell Proteomics* **3** 692-703.

Levin AA, Sturzenbecker LJ, Kazmer S, Bosakowski T, Huselton C, Allenby G, Speck J, Kratzeisen C, Rosenberger M, Lovey A, et al. 1992 9-cis retinoic acid stereoisomer binds and activates the nuclear receptor RXR alpha. *Nature* **355** 359-361.

Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, et al. 2019 Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75** 861-877.

Love JD, Gooch JT, Benko S, Li C, Nagy L, Chatterjee VK, Evans RM & Schwabe JW 2002 The structural basis for the specificity of retinoid-X receptor-selective agonists: new insights into the role of helix H12. *J Biol Chem* **277** 11385-11391.

Mark M, Ghyselinck NB & Chambon P 2006 Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. *Annu Rev Pharmacol Toxicol* **46** 451-480.

Mark M, Teletin M, Vernet N & Ghyselinck NB 2015 Role of retinoic acid receptor (RAR) signaling in post-natal male germ cell differentiation. *Biochim Biophys Acta* **1849** 84-93.

574 Mascrez B, Mark M, Dierich A, Ghyselinck NB, Kastner P & Chambon P 1998 The RXRalpha ligand-
575 dependent activation function 2 (AF-2) is important for mouse development. *Development* **125** 4691-
576 4707.

577 Nahoum V, Perez E, Germain P, Rodriguez-Barrios F, Manzo F, Kammerer S, Lemaire G, Hirsch O,
578 Royer CA, Gronemeyer H, et al. 2007 Modulators of the structural dynamics of the retinoid X receptor
579 to reveal receptor function. *Proc Natl Acad Sci U S A* **104** 17323-17328.

580 Niu H, Fujiwara H, di Martino O, Hadwiger G, Frederick TE, Menendez-Gutierrez MP, Ricote M,
581 Bowman GR & Welch JS 2017 Endogenous retinoid X receptor ligands in mouse hematopoietic cells.
582 *Sci Signal* **10**.

583 Perez E, Bourguet W, Gronemeyer H & de Lera AR 2012 Modulation of RXR function through ligand
584 design. *Biochim Biophys Acta* **1821** 57-69.

585 Perez Santin E, Germain P, Quillard F, Khanwalkar H, Rodriguez-Barrios F, Gronemeyer H, de Lera
586 AR & Bourguet W 2009 Modulating retinoid X receptor with a series of (E)-3-[4-hydroxy-3-(3-alkoxy-
587 5,5,8,8-tetramethyl-5,6,7,8-tetrahydronaphthalen-2-y l)phenyl]acrylic acids and their 4-alkoxy isomers.
588 *J Med Chem* **52** 3150-3158.

589 Ruhl R, Krezel W & de Lera AR 2018 9-Cis-13,14-dihydroretinoic acid, a new endogenous mammalian
590 ligand of retinoid X receptor and the active ligand of a potential new vitamin A category: vitamin A5.
591 *Nutr Rev* **76** 929-941.

592 Ruhl R, Krzyzosiak A, Niewiadomska-Cimicka A, Rochel N, Szeles L, Vaz B, Wietrzych-Schindler M,
593 Alvarez S, Szklenar M, Nagy L, et al. 2015 9-cis-13,14-Dihydroretinoic Acid Is an Endogenous Retinoid
594 Acting as RXR Ligand in Mice. *PLoS Genet* **11** e1005213.

595 Schierle S & Merk D 2019 Therapeutic modulation of retinoid X receptors - SAR and therapeutic
596 potential of RXR ligands and recent patents. *Expert Opin Ther Pat* **29** 605-621.

Shulman AI, Larson C, Mangelsdorf DJ & Ranganathan R 2004 Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell* **116** 417-429.

Weikum ER, Liu X & Ortlund EA 2018 The nuclear receptor superfamily: A structural perspective. *Protein Sci* **27** 1876-1892.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al. 2011 Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67** 235-242.

Zeng Z, Sun Z, Huang M, Zhang W, Liu J, Chen L, Chen F, Zhou Y, Lin J, Huang F, et al. 2015 Nitrostyrene Derivatives Act as RXRalpha Ligands to Inhibit TNFalpha Activation of NF-kappaB. *Cancer Res* **75** 2049-2060.

Zhang J, Hu X & Lazar MA 1999 A novel role for helix 12 of retinoid X receptor in regulating repression. *Mol Cell Biol* **19** 6448-6457.

612 **Figure legends**

613

614 **Figure 1:** Chemical structures of natural and synthetic compounds used in this study.

615

616 **Figure 2:** The ligand-binding function is impaired in 2mRXRs. (A) Close-up view of the mouse RXRα
617 LBP in the wtRXRα-RARα LBD heterodimer structure (PDB 7QAA). The bound fatty acid (FA) is
618 shown as grey sticks together with the light grey 2Fo-Fc map contoured at 1 σ. The mutated residues in
619 2mRXRα and 3mRXRα are shown as grey and black sticks, respectively. (B) Compound activity in
620 transactivation assays in COS-1 cells. Cells were exposed to increasing concentrations of test
621 compounds. Data are expressed as mean (± s.e.m). (C) Same experiment as (B) using 2mRXRα. (D)
622 Titration of fluorescein-labeled SRC-1 peptide by wtRXRα LBD in the absence of ligand (DMSO) or
623 in the presence of the indicated compounds. Assays were performed in triplicate and data are expressed
624 as mean (± s.e.m). (E) Same experiments as (D) using 2mRXRα. (F) Melting temperatures (Tm)
625 determined for wtRXRα LBD in the absence of ligand (DMSO) or in the presence of the indicated
626 compounds. Assays were performed in triplicate and data are expressed as mean (± s.e.m). (G) Same
627 experiments as (F) using 2mRXRα.

628

629 **Figure 3:** Analysis of ligand binding by native mass spectrometry. Characterization of the interaction
630 between wtRXRα LBD (A), 2mRXRα LBD (B) and 3mRXRα LBD (C) and the synthetic compounds
631 CD3254, BMS649 and LG268. Ratios between the ligand-bound and unliganded proteins were
632 measured for each energy and compared for each ligand. Data were normalized by considering the signal
633 obtained for each protein-ligand complex at the lowest dissociation energy (0V) as 100%. Assays were
634 performed in triplicate and data are expressed as mean (± S.D.).

635

636 **Figure 4:** 3mRXRs respond exclusively to benzoic acid-containing synthetic compounds. (A)
637 Compound activity in transactivation assays in COS-1 cells. Cells were exposed to increasing
638 concentrations of test compounds. Data are expressed as mean (± s.e.m). (B) Titration of fluorescein-
639 labeled SRC-1 peptide by 3mRXRα LBD in the absence of ligand (DMSO) or in the presence of the

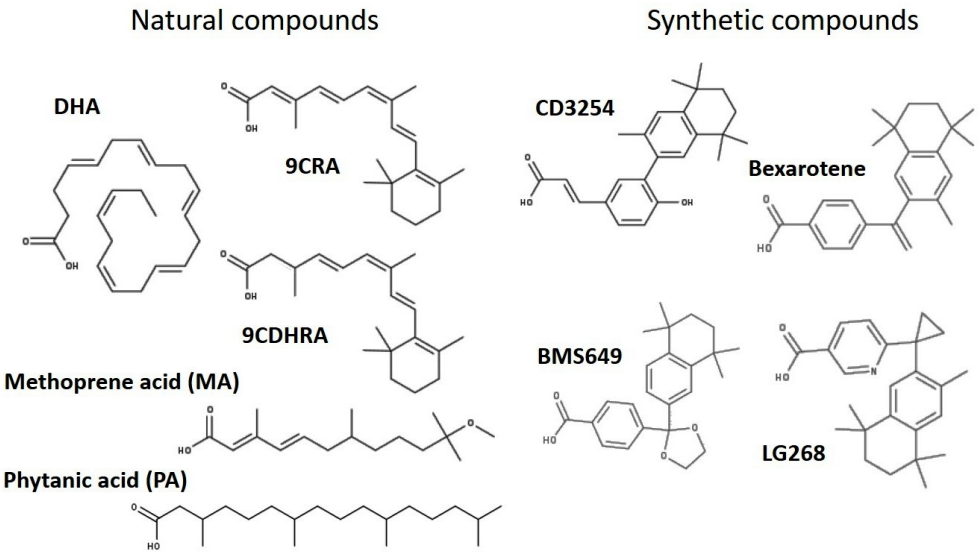
indicated compounds. Assays were performed in triplicate and data are expressed as mean (± s.e.m). (C) Melting temperatures (Tm) determined for 3mRXRα LBD in the absence of ligand (DMSO) or in the presence of the indicated compounds. Assays were performed in triplicate and data are expressed as mean (± s.e.m).

Figure 5: 2mRXRs and 3mRXRs retain their DNA-binding and dimerization capabilities. (A) Electrophoresis under non-denaturing conditions revealing the oligomeric states of wtRXRα, 2mRXRα, and 3mRXRα LBDs individually and in complex with TRα1 LBD. The white asterisks indicate RXRα–TRα heterodimers. Bovine serum albumin (BSA) was used as a molecular weight marker. (B) Size-exclusion chromatography-multi angle light scattering (SEC-MALS) analysis showing that wtRXRα, 2mRXRα, and 3mRXRα LBDs heterodimerize with TRα1 LBD in solution.

Figure 6: Structure of LG268-bound 2mRXRα. (A) Superimposition of LG268-bound 2mRXRα with LG268-bound wtRXRβ (PDB 1H9U). The ligands and amino acid side chains in 2mRXRα and wtRXRβ are displayed as orange and cyan sticks, respectively. (B) Superimposition of LG268-bound 2mRXRα (orange) with 9CRA-bound wtRXRα (cyan, PDB 1FBY). (C) Superimposition of LG268-bound 2mRXRα (orange) with CD3254-bound wtRXRα (cyan, PDB 3FUG). (D) Superimposition of LG268-bound 2mRXRα (orange) with 9CDHRA-bound wtRXRα (cyan, PDB 4ZSH). The red asterisk denotes steric incompatibility between Q437 and the ligands displayed in cyan.

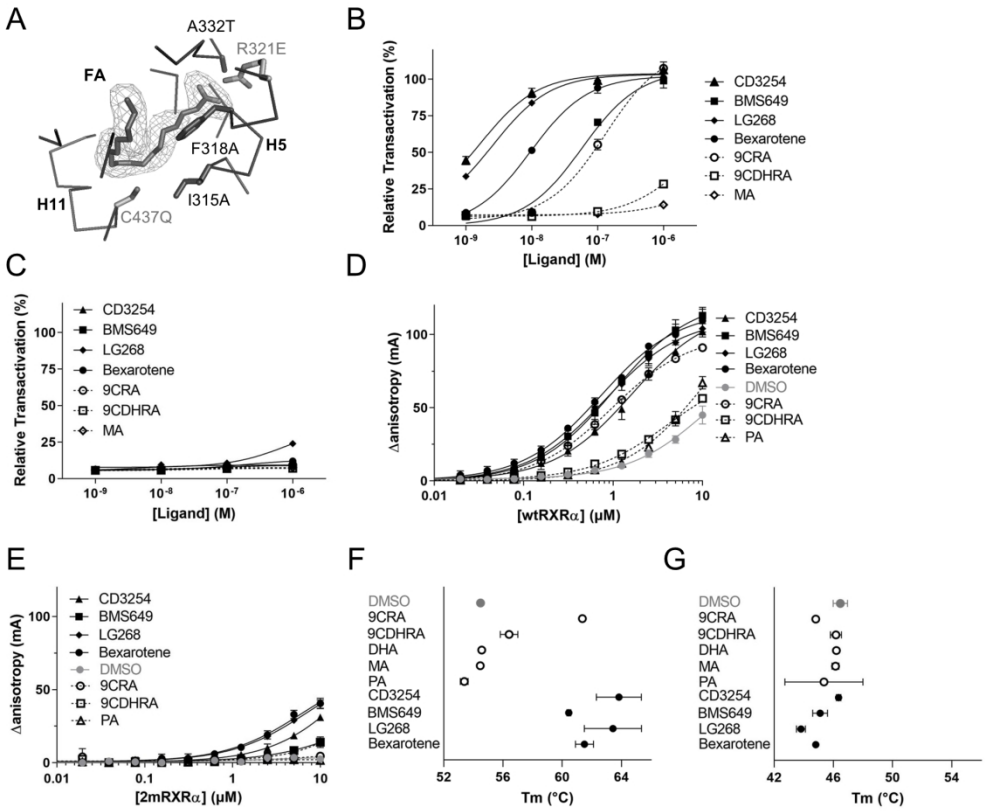
Figure 7: Structure of BMS649-bound 3mRXRα. (A) Superimposition of BMS649-bound 3mRXRα with BMS649-bound wtRXRα (PDB 1MVC). The ligands and amino acid side chains in 3mRXRα and wtRXRα are displayed as orange and cyan sticks, respectively. (B) Another view to show the different binding modes of BMS649 in 3mRXRα (orange) and wtRXRα (cyan). Note the shift towards H5, H7 (and H11, not shown), and the rotation of the benzoic acid ring to maintain van der Waals contacts with residues in helices H3 and H5 of 3mRXRα (red dotted lines). (C) Superimposition of the indicated ligands as it comes when superimposing the corresponding RXR complex structures. 9CRA and

667 CD3254 are markedly longer than LG268, BMS649 and bexarotene. Values in parentheses indicate the
668 longest inter-atomic distance in each ligand.



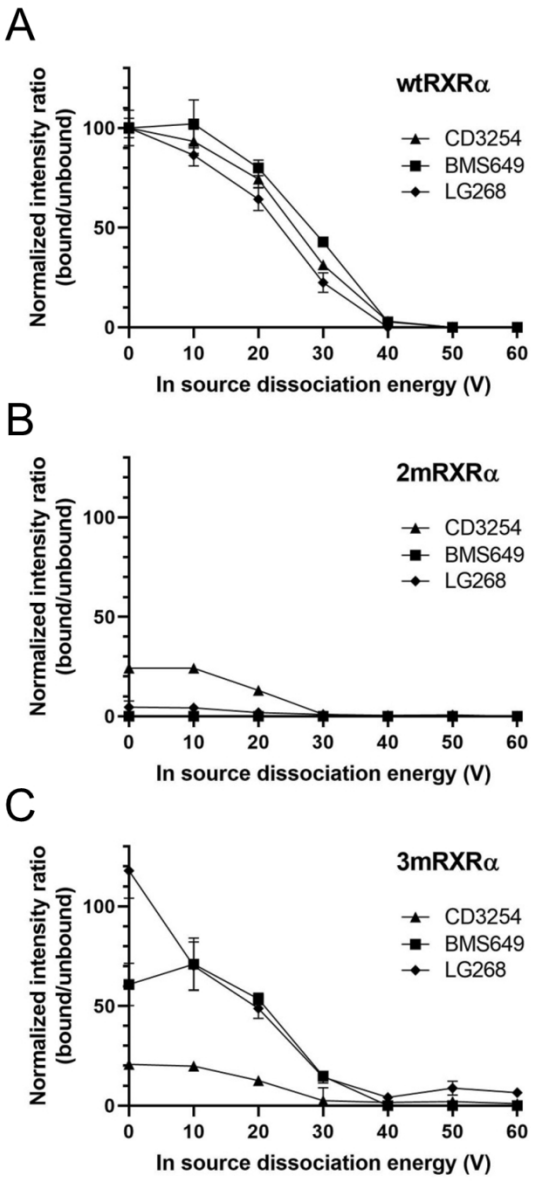
Chemical structures of natural and synthetic compounds used in this study.

226x129mm (150 x 150 DPI)



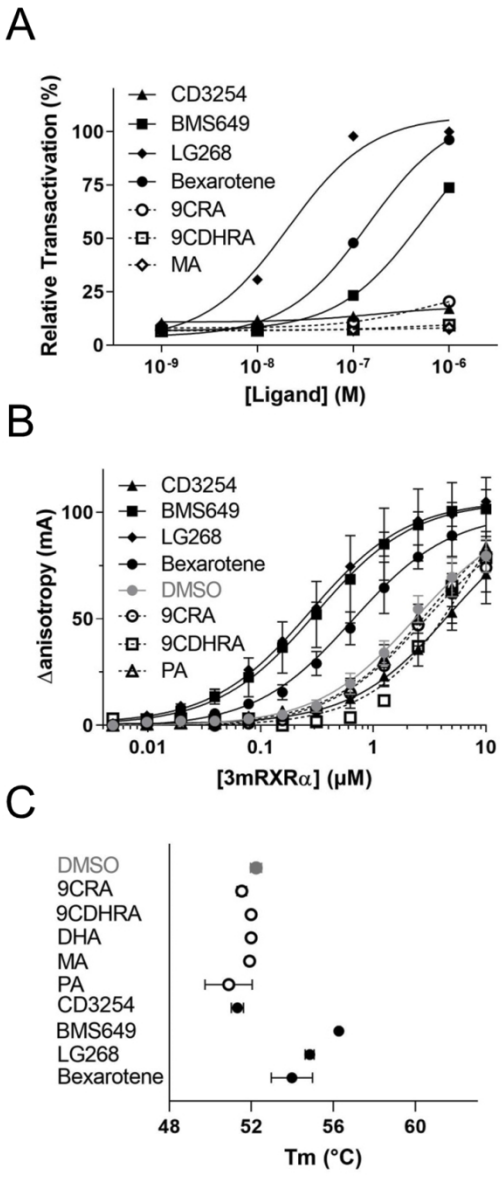
The ligand-binding function is impaired in 2mRXRs.

231x189mm (230 x 230 DPI)



Analysis of ligand binding by native mass spectrometry.

91x188mm (230 x 230 DPI)



3mRXRs respond exclusively to benzoic acid-containing synthetic compounds.

84x190mm (230 x 230 DPI)

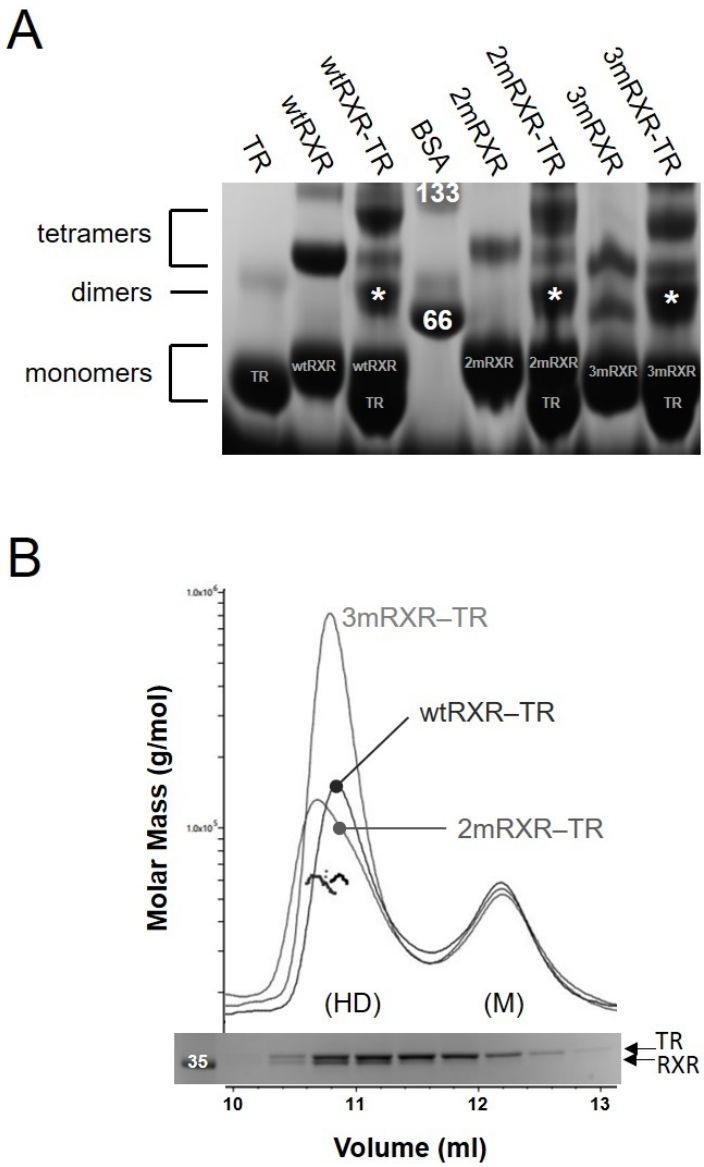
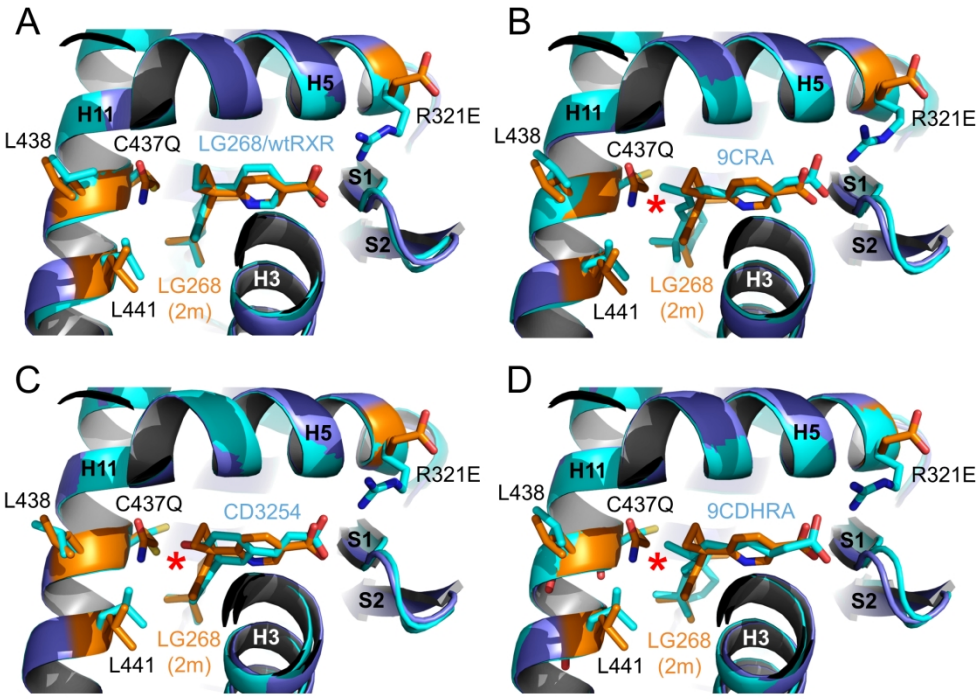


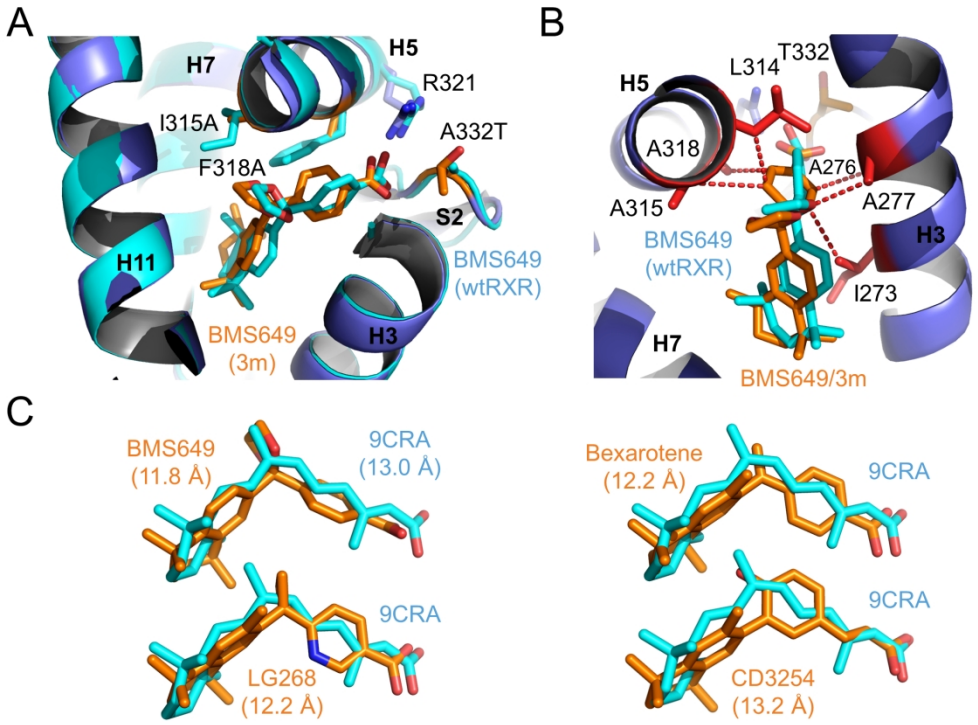
Figure 5

125x186mm (150 x 150 DPI)



Structure of LG268-bound 2mRXRα.

262x186mm (230 x 230 DPI)



Structure of BMS649-bound 3mRXRα.

243x182mm (230 x 230 DPI)

Supplementary material for

Design and in vitro characterization of RXR variants as tools to investigate the biological role of endogenous rexinoids

Albane le Maire[1], Martial Rey[2], Valérie Vivat[3,*], Laura Guée[1], Pauline Blanc[1], Christian Malosse[2], Julia Chamot-Rooke[2], Pierre Germain[1], and William Bourguet[1]

[1]CBS (Centre de Biologie Structurale), Univ Montpellier, CNRS, Inserm, Montpellier, France

[2]Institut Pasteur, Université de Paris, CNRS USR2000, Mass Spectrometry for Biology Unit, 75015 Paris, France

[3]IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Univ Strasbourg, CNRS, Inserm, Illkirch, France

***Current address:** Flare Therapeutics, 215 First Street, Suite 150, Cambridge, MA 02142, USA

Includes:

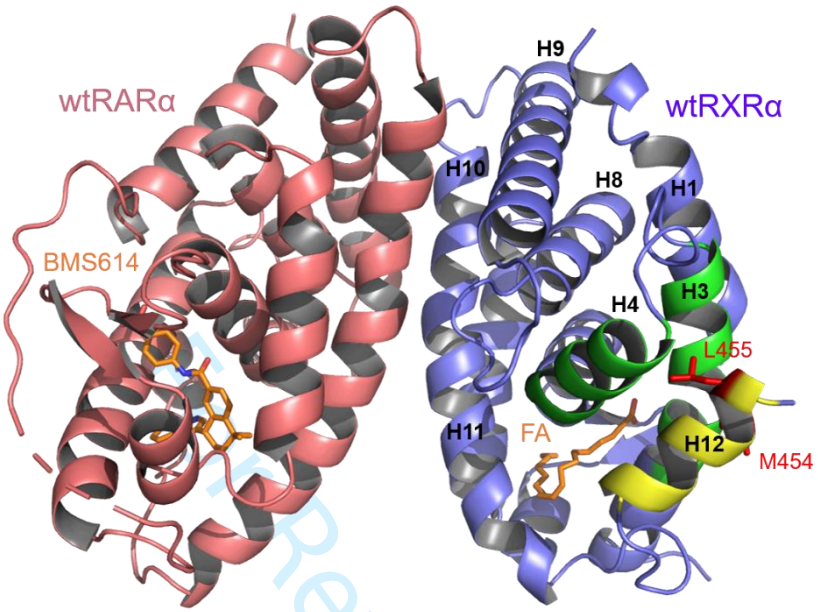
Supplementary Materials and Methods

Label-free thermal shift assay (nanoDSF)

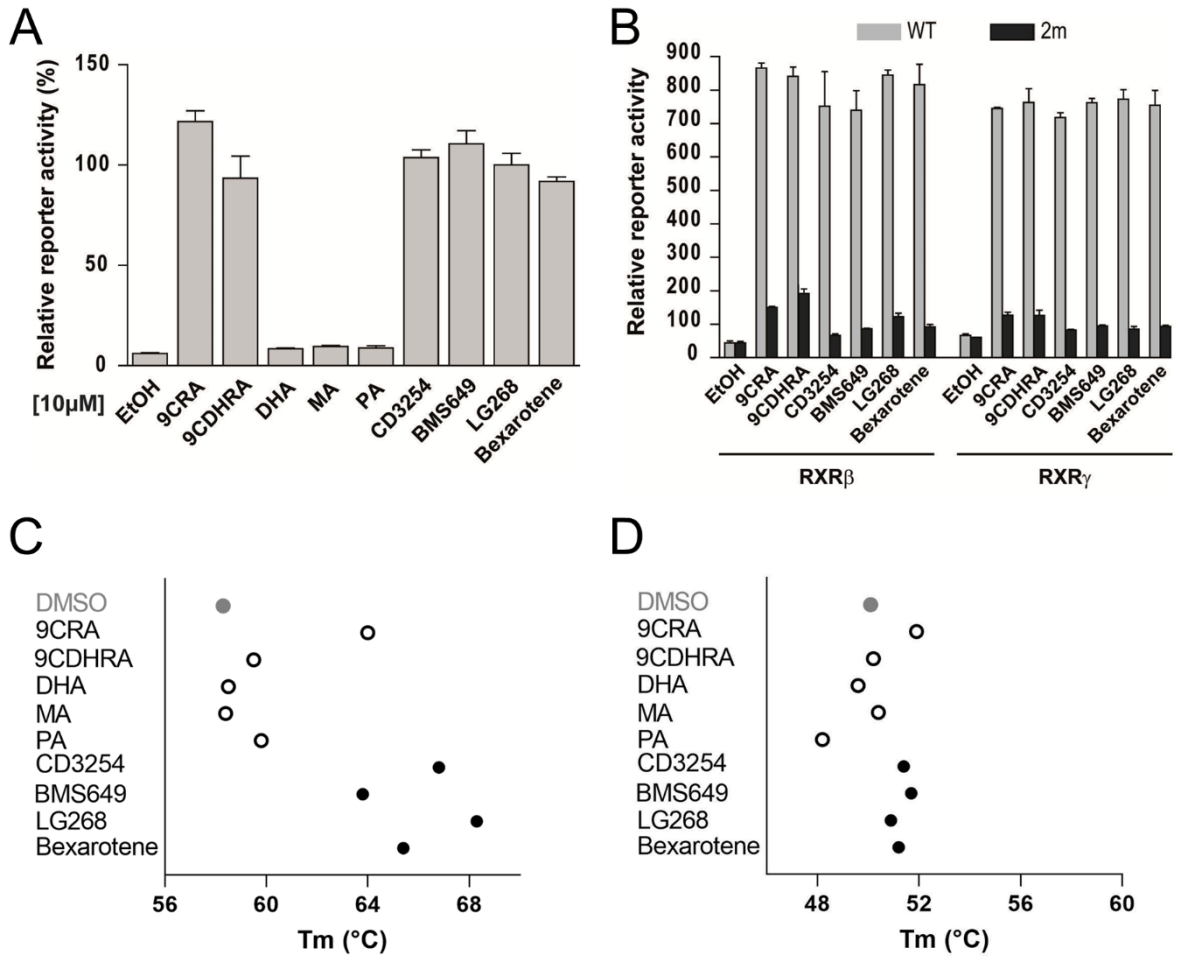
Label-free thermal shift assay experiments were performed with a Tycho NT.6 instrument (NanoTemper Technologies). The samples (20 µM of RXRα protein), in the presence of DMSO (condition without ligand) or in the presence of three molar excess of ligands, were heated in a glass capillary at a rate of 30 K/min (20°C to 95°C) in 50 mM Tris HCl pH 7.5, 150 mM NaCl, 5% glycerol, and the internal fluorescence at 330 and 350 nm was recorded. Data analysis, data smoothing, and calculation of derivatives was done using the internal evaluation features of the Tycho instrument.

Two-hybrid assays

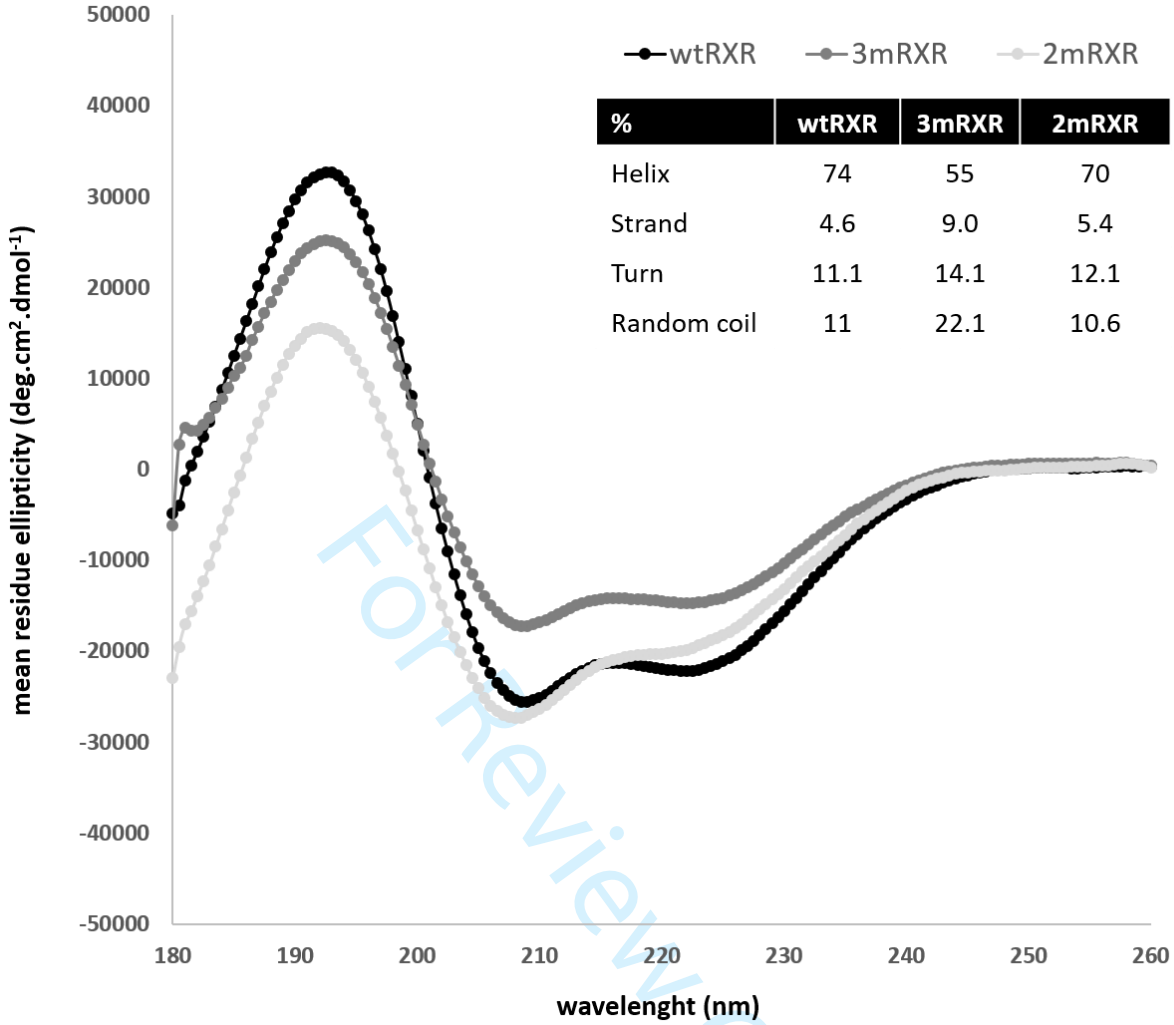
COS cells were cultured in DMEM with Glutamax and 10% (v/v) FCS and transfected using JetPei transfectant (Ozyme). After 24 h, the medium was changed to a medium containing the selective RAR agonist TTNPB or vehicle. Cells were lysed and assayed for reporter gene expression 48 h after transfection. The luciferase assay system was used according to the manufacturer's instruction (Promega). In each case results were normalized to co-expressed β–galactosidase. TTNPB was purchased from Sigma. RXRα mutants were generated into pSG5-RXRα by PCR-assisted site-directed mutagenesis with Deep Vent DNA polymerase (New England Biolabs). The construct was verified by DNA sequencing. The pSG5-based Gal-NCoR and VP16-RARα LBD (VP16-RAR) expression vectors, and the (17m)5x-βGlob-Luc reporter gene have been described (Germain, et al. 2002).



Supplementary Figure 1: Structure of wtRXRα-RARα LBD heterodimer. The RXRα (blue) and RARα (pink) subunits are bound to a fatty acid (FA) and the antagonist BMS614, respectively. Ligands are displayed as orange sticks. The corepressor binding groove made of helices H3 and H4 is highlighted in green. The RXR activation helix H12 occupying the corepressor binding site in the fatty acid-bound state is colored in yellow. M454 and L455 that have been mutated to delocalize H12 from the corepressor binding site are shown in red.

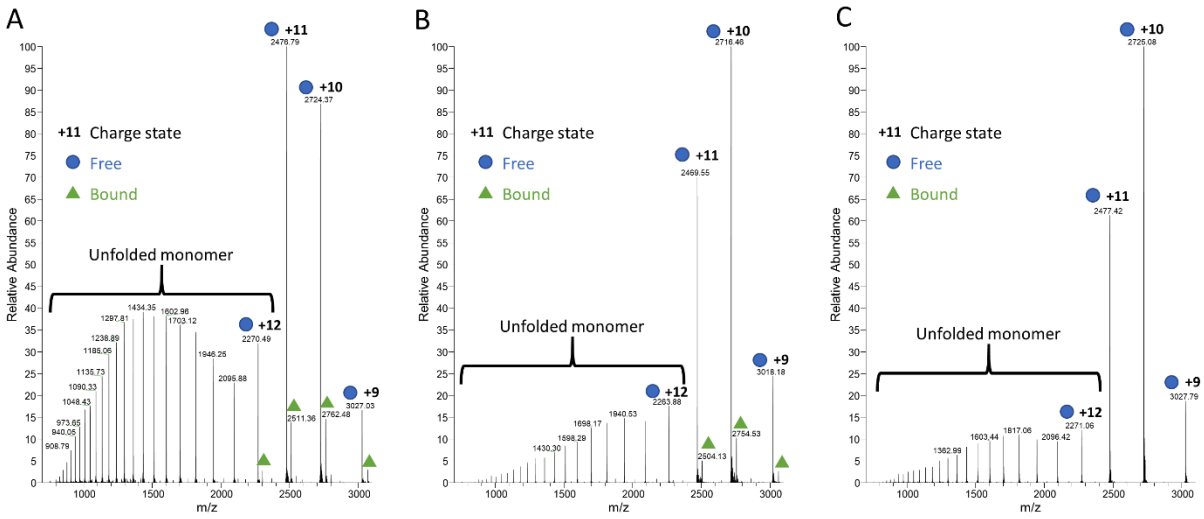


Supplementary Figure 2: The ligand-binding function is impaired in 2mRXRs. (A) COS cells were transiently co-transfected with the reporter gene (DR1)-tk-Luc and RXRα to assess the RXR agonist potential of compounds at 10 μM. 100% corresponds to reporter gene transcription induced in the presence of the full RXR agonist LG268. All error bars are expressed as s.e.m. (B) COS cells were transiently co-transfected with the reporter gene (DR1)-tk-Luc and either RXRβ (light grey) or 2mRXRβ (black) (left panel) or RXRγ (light grey) or 2mRXRγ (black) (right panel). All error bars are expressed as s.e.m. (C, D) Label-free thermal shift assay (nanoDSF). Melting temperatures (Tm) determined for wtRXRα (C) and 2mRXRα LBDs (D) in the absence of ligand (DMSO) or in the presence of the indicated compounds.

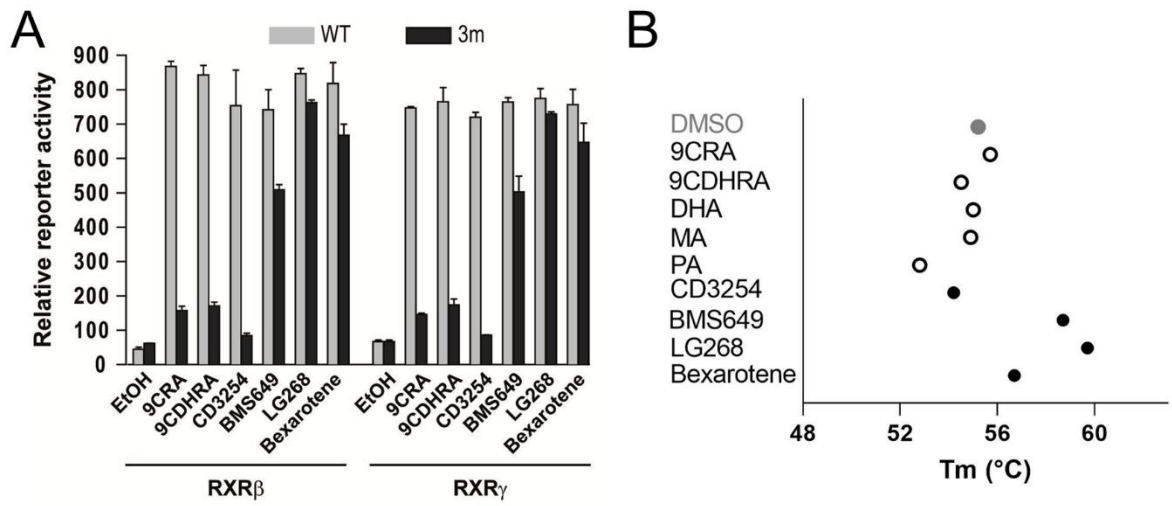


%	wtRXR	3mRXR	2mRXR
Helix	74	55	70
Strand	4.6	9.0	5.4
Turn	11.1	14.1	12.1
Random coil	11	22.1	10.6

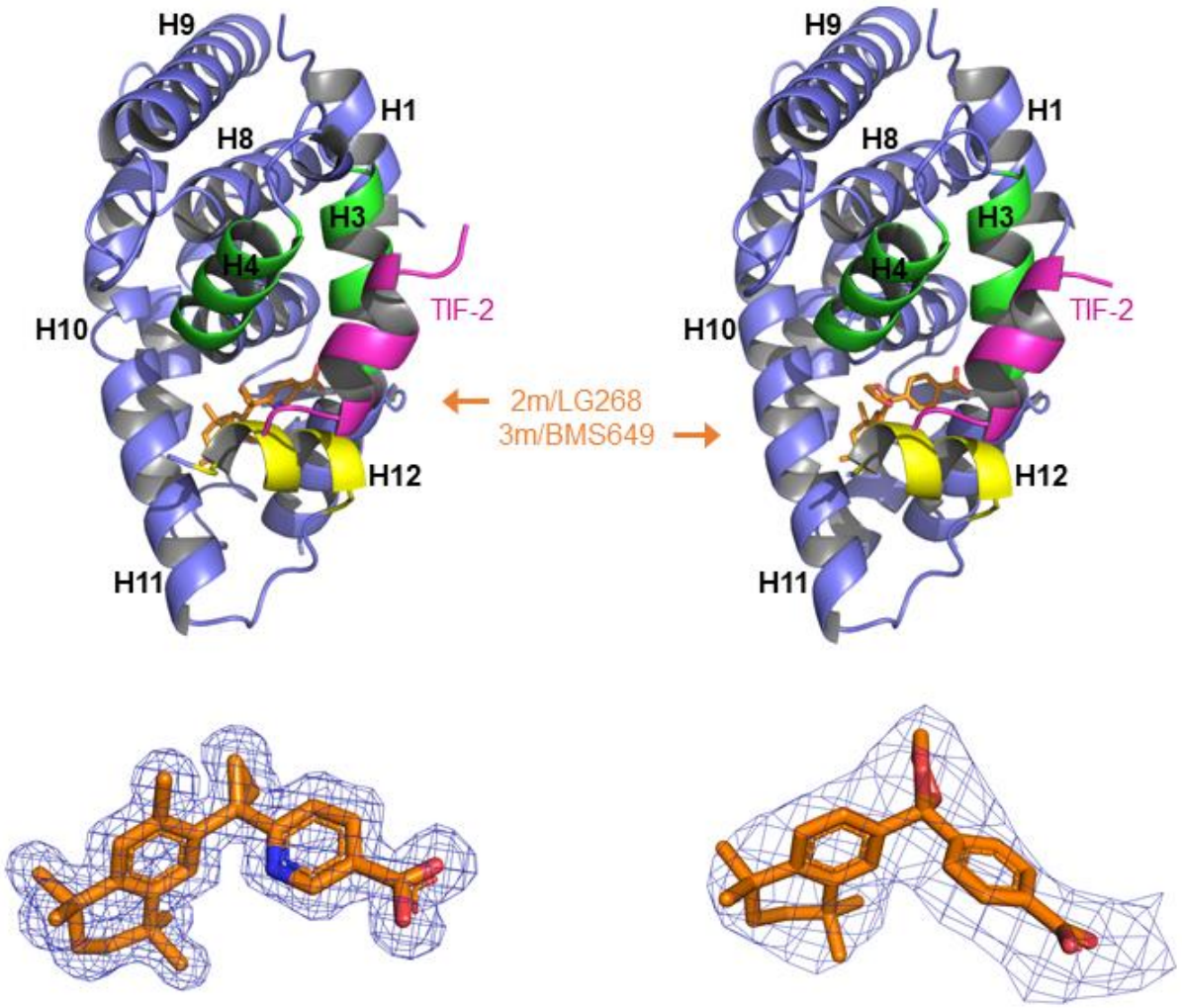
Supplementary Figure 3: Far-UV circular dichroism spectra of wt, 2m- and 3mRXRα at room temperature in purification buffer.



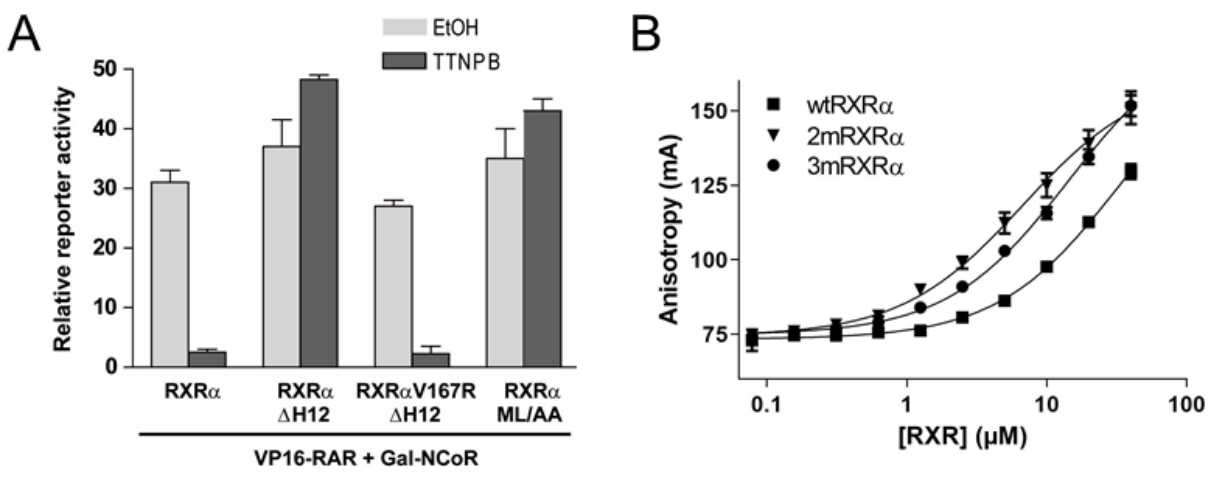
Supplementary Figure 4: Representative mass spectra of the (A) wtRXRα, (B) 3mRXRα and (C) 2mRXRα in complex with BMS649 with no IS-CID energy. Folded monomer signal is spread between 4 different charge states (+12, +11, +10 and +9). The intensities of the bound (green triangles) and free (blue circles) forms are measured in function of the IS-CID energies to generate the graph presented in the Figure 3.



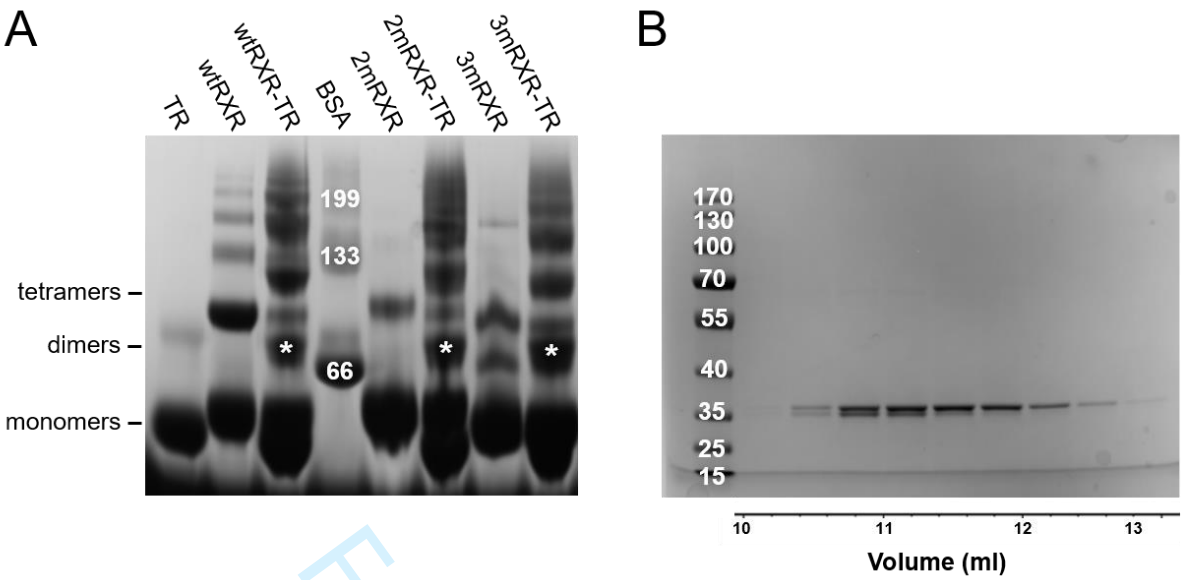
Supplementary Figure 5: 3mRXRs respond exclusively to benzoic acid-containing synthetic compounds. (A) COS cells were transiently co-transfected with the reporter (DR1)-tk-Luc and either RXRβ (light grey) or 3mRXRβ (black) (left panel) or RXRγ (light grey) or 3mRXRγ (black) (right panel). All error bars are expressed as s.e.m. (B) Label-free thermal shift assay (nanoDSF). Melting temperatures (Tm) determined for 3mRXRα LBD in the absence of ligand (DMSO) or in the presence of the indicated compounds.



Supplementary Figure 6: Structures of 2m- and 3mRXRα in complex with LG268 and BMS649. Ligands are displayed as orange sticks. The coregulator binding groove made of helices H3 and H4 is highlighted in green. The activation helix H12 in the transcriptionally active position is colored in yellow. The TIF-2 coactivator-derived peptide is displayed in magenta. The electron densities shown are unbiased omit Polder maps contoured at 2.0σ, with model bias reduction and exclusion of solvent density.



Supplementary Figure 7: Interaction of RXRα with the corepressor NCoR. (A) Interactions of RXRα-RARα heterodimers with the corepressor NCoR using mammalian two-hybrid assays with $(17m)_{5x}$-βGlob-Luc reporter in COS cells and with chimeras containing the GAL4 DNA-binding domain fused to the nuclear receptor interaction region of NCoR (Gal-NCoR) as bait and RARα LBD fused to the activation domain of VP16 (VP16-RAR) as prey in combination with RXRα, RXRαΔH12, RXRαV167RΔH12 or RXRαML/AA in the presence (light gray) or absence (dark grey) of the RAR selective agonist TTNPB (10nM). All error bars are expressed as s.e.m. In the presence of wtRXRα, addition of TTNPB is sufficient to dissociate NCoR from the heterodimer, indicating that RXRα plays little role in the association. Upon RXRα H12 deletion (RXRαΔH12), TTNPB fails to dissociate the complex, suggesting that a new interaction surface for NCoR is present in RXRαΔH12. When preventing the interaction between RXRαΔH12 and NCoR by introducing an additional mutation in the corepressor-binding site of RXRα (RXRαV167RΔH12) (Hu and Lazar 1999), the dissociation capacity of TTNPB is restored, confirming that the groove formed by RXR helices H3 and H4 can serve as a corepressor binding surface, provided that helix H12 is not masking it (**Supplementary Fig. 1**). In the same line, the double mutation M454A/L455A (RXRαML/AA) yields a similar interaction pattern than that observed with RXRαΔH12, indicating that the introduced mutations in helix H12 (**Supplementary Fig. 1**) cause the release of the RXRα hydrophobic groove and NCoR recruitment. (B) Titration of fluorescein-labeled NCoR peptide by wtRXRα, 2mRXRα or 3mRXRα LBDs. Assays were performed in triplicate and data are expressed as mean (± s.e.m). Altogether, these observations suggest that RXRα helix H12 may adopt a fatty acid-dependent conformation that prevents NCoR binding to RXRα, so that in a physiological context, transcriptional activation of the heterodimer is controlled by the partner receptor.



Supplementary Figure 8: 2mRXRs and 3mRXRs retain their DNA-binding and dimerization capabilities. (A) Full-size native gel from Figure 5A. (B) Full-size SDS-PAGE gel from Figure 5B. Molecular weight protein standards are indicated on each gel.

Supplementary Table 1: Data collection and refinement statistics

	3mRXRα	2mRXRα	RXRα-RARα
PDB code	**7PDT**	**7PDQ**	**7QAA****
Wavelength	0.97242	0.99998	0.918
Resolution range	42.83-3.30 (3.41-3.30)*	48.39-1.58 (1.64-1.58)*	19.47-2.76 (2.86-2.76)*
Space group	P 21 21 21	P 43 21 2	P 61 2 2
Unit cell	63.39 68.53 109.72 90 90 90	68.43 68.43 105.61 90 90 90	116.60 116.60 207.80 90 90 120
Total reflections	36709 (3841)	761965 (19995)	
Unique reflections	7562 (749)	34564 (2903)	
Multiplicity	4.9 (5.1)	22.0 (6.9)	
Completeness (%)	98.94 (99.47)	98.25 (83.90)	
Mean I/sigma (I)	10.77 (3.40)	32.76 (2.04)	
Wilson B-factor	83.50	19.68	
R-merge	0.122 (0.548)	0.0604 (0.668)	
R-meas	0.137 (0.610)	0.06176 (0.7216)	
R-pim	0.06108 (0.2639)	0.01247 (0.2585)	
CC1/2	0.994 (0.749)	1 (0.812)	
CC*	0.999 (0.926)	1 (0.947)	
Refl. used in refinement	7530 (750)	34557 (2902)	19728 (2041)
Refl. used for R-free	761 (81)	1708 (130)	1009 (96)
R-work	0.1946 (0.2670)	0.1713 (0.2599)	0.1947 (0.2221)
R-free	0.2098 (0.2786)	0.1967 (0.2954)	0.2450 (0.3012)
Non-hydrogen atoms	3491	2120	3654
macromolecules	3428	1756	3479
ligands	56	27	54
solvent	7	337	121
Protein residues	437	220	447
RMS (bonds)	0.003	0.016	0.005
RMS (angles)	0.66	1.47	0.99
Ramach. favored (%)	97.18	98.60	96.00
Ramach. allowed (%)	2.82	1.40	3.56
Ramach. outliers (%)	0.00	0.00	0.44
Rotamer outliers (%)	0.00	0.00	0.00
Clash score	10.14	6.14	10.71
Average B-factor	70.22	23.70	47.51
macromolecules	70.46	21.81	47.60
ligands	56.95	15.36	38.70
solvent	55.71	34.20	48.00
Number of TLS groups	12	5	0

*****Values in parentheses are for highest resolution shell

******Structure deposited at the PBD without data collection statistics (data collected 25 years ago and lost since)