SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
X-RAY CRYSTALLOGRAPHY
THERMODYNAMICS

Received
20 October 2014

Accepted
29 December 2014

Published
3 February 2015

Correspondence and requests for materials should be addressed to N.R. (rochel@igbmc.fr)

* These authors contributed equally to this work.

Structural Basis of Natural Promoter Recognition by the Retinoid X Nuclear Receptor

Judit Osz[1]*, Alastair G. McEwen[2]*, Pierre Poussin-Courmontagne[2], Emmanuel Moutier[3], Catherine Birck[2], Irwin Davidson[3], Dino Moras[1] & Natacha Rochel[1]

[1]Department of Integrative Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U964/Centre National de la Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 67404 Illkirch, France, [2]Structural Biology Platform, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U964/Centre National de la Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 67404 Illkirch, France, [3]Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U964/Centre National de la Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 67404 Illkirch, France.

Retinoid X receptors (RXRs) act as homodimers or heterodimerisation partners of class II nuclear receptors. RXR homo- and heterodimers bind direct repeats of the half-site (A/G)G(G/T)TCA separated by 1 nucleotide (DR1). We present a structural characterization of RXR-DNA binding domain (DBD) homodimers on several natural DR1s and an idealized symmetric DR1. Homodimers displayed asymmetric binding, with critical high-affinity interactions accounting for the 3′ positioning of RXR in heterodimers on DR1s. Differing half-site and spacer DNA sequence induce changes in RXR-DBD homodimer conformation notably in the dimerization interface such that natural DR1s are bound with higher affinity than an idealized symmetric DR1. Subtle changes in the consensus DR1 DNA sequence therefore specify binding affinity through altered RXR-DBD-DNA contacts and changes in DBD conformation suggesting a general model whereby preferential half-site recognition determines polarity of heterodimer binding to response elements.

Signaling by nuclear receptors (NRs) controls a multitude of physiological phenomena (embryogenesis, homeostasis, reproduction, cell growth, differentiation, and apoptosis)[1–3]. NRs act as key components of gene regulation through binding to hormone response elements, (HREs) in the regulatory sequences of their target genes. NRs share a common structural organization comprising a variable N-terminal domain (NTD) harboring a ligand-independent activation function, the conserved DNA binding domain (DBD) and the C-terminal ligand binding domain (LBD)[4]. The LBD is a key regulatory domain containing the ligand binding pocket, multiple interaction surfaces for homo and heterodimerisation and interactions with coregulators. NR dimerization involves strong interactions between the LBDs of the interacting partners and from the binding of the two DBDs to neighboring hexanucleotide DNA motifs [half-sites of consensus (A/G)G(G/T)TCA] that make up the HREs. Specificity results not only from the DNA sequence of the two half-sites, but also from the geometry, spacing and relative orientation of the half-sites in the HRE[5].

Among NR members, the Retinoid X nuclear Receptors (RXRs) possess the unique ability to act as homodimers or through heterodimerisation with other class II NRs in multiple signalling pathways critical for embryonic development, metabolic processes, differentiation and apoptosis[6–9]. RXRs comprise three isotypes, RXRα (NR2B1), RXRβ (NR2B2) and RXRγ (NR2B3), that are differentially expressed in tissues and whose expression profiles are altered in several diseases[10]. RXR binds direct repeats of the hexanucleotide half-site separated by 1 nucleotide (DR1), either as a homodimer or heterodimer with RAR or PPAR[11]. ChIP-seq and protein binding microarray (PBM) have revealed RXR binding even in the absence of a heterodimeric partner[12–14]. RXR heterodimers bind DR1 with a defined polarity with RXR bound to the 3′ half-site[15–18]. In homodimers, RXR has also been shown to bind preferentially to asymmetric DR1[19,20] with a preferential binding to the 3′ half-site[14]. The cellular environment and promoter context play critical roles in determining receptor specificity[21,22]. Few crystal structures of NR bound to natural REs have been reported[23–25]. However, the structure of the estrogen receptor bound to a non-consensus ERE in which a single change of one nucleotide leads to a side chain reorientation,

alternate base contacts and lower affinity has been described[23]. For the ecdysone nuclear receptor (EcR)-ultraspiracle protein (Usp) heterodimer, significant differences were observed between a natural inverted repeat IR1 compared to a consensus IR1[25]. No structural information is available on RXR on natural DR1s and little is known about the specificity of DNA recognition of natural HREs by the RXR dimer at the atomic level or how DNA sequence can discriminate between different RXR dimers.

To elucidate the molecular basis for DNA target specificity of the RXR dimer and to determine whether and how the DNA sequence and topological organization of the half-sites in the HRE exert allosteric control of RXR homodimer function, we performed a structure-function analysis of RXRα bound to natural DR1s from target genes. To analyze the effects of genomic variation in natural DR1 sequences on RXR binding and their structural impact we used isothermal titration calorimetry (ITC) to monitor the thermodynamics of RXR homodimer DBD binding to natural DR1 elements from the calcitonin receptor activity-modifying protein 2 (*Ramp2*), the NR subfamily 1, group D, member 1 (*Nr1d1*) and the glycerophosphodiester phosphodiesterase 1 (*Gde1*) genes[26], the known natural RXR HRE (RXRE) from the malic enzyme PPRE gene (*MEp*)[21] and consensus idealized RXREs. The crystal structures of the RXR DBD complexes with *Ramp2*, *Nr1d1, Gde1* and an idealized DR1 were solved. Our data reveal that the differences in half-site sequence of the natural DR1s affect RXR homodimer binding affinity and conformation hence defining the molecular determinants of RXR interactions with natural HREs.

Results

Binding affinities of natural DR1s for the RXR DBD homodimer. We previously identified several DR1 elements as bound by RAR-RXR in genome-wide analyses[26,27]. Amongst these are DR1s located around the *Ramp2*, *Nr1d1* and *Gde1* genes (Fig. 1a). Previous studies showed that the *Ramp2* and *Nr1d1* elements efficiently bind RAR-RXR, whereas the *Gde1* element, which has perfect consensus half-sites, but a C/G base pair as spacer, shows almost no binding[26]. Exchange of the C/G base pair by A/T (*Gde1*SpA*)*, as is present in the *Ramp2* DR1, resulted in strong binding indicating the critical nature of the spacer nucleotide (Ref. 26 and see Fig. 1a). To determine the variability in the DR1 sequence bound by RXR, we analyzed the sequence motifs from RXR Chip-Seq data[28] allowing one mismatch (Fig. 1b). In the derived consensus element, the second half-site shows higher conservation with, in particular, high base preference at positions 2, 4, 5 and 6 ($>$ 1 bit) and almost no C as spacer with G and A being highly represented.

To understand how sequence variations of the DR1 response elements modulate the affinity of protein-DNA interactions and the orientation of RXR dimer, we determined the binding affinities and thermodynamic parameters of DNA binding of natural DR1s by RXR DBD homodimers. We assessed the binding of highly purified recombinant RXR DBD to these elements by quantitative ITC. Fig. 1c show representative isotherms obtained from ITC measurements, and a detailed analysis of thermodynamic parameters corresponding to the averaged values of a least three independent experiments is presented (Table 1). RXR DBD does not dimerize until bound to the DNA where 2 DBDs bind to 1 DNA molecule in a cooperative manner with binding of the first monomer favoring binding of the second as previously described for other NR DBDs[29,30]. In agreement with previous studies reporting reduced binding to monomeric sites by RXR[31] or steroid receptors[32], the RXR-DBD monomer binds weakly DNA sequences containing a single half-site (GGGTCA or AGTTCA), thus confirming that the affinity of RXR is markedly enhanced through dimerization and cooperative binding. All DR1s, even the weakest, bind the RXR dimer cooperatively. The binding affinities observed for RXR homodimers vary up to 16 fold, from 10 nM for the DR1 from the malic enzyme *MEp* gene, a PPAR HRE that acts as a RXRE[21], to 160 nM for the consensus 'idealized' (id)DR1. The binding is driven by a favorable enthalpic contribution that results from the extensive electrostatic interactions, together with an entropic penalty resulting from a loss of conformational freedom compensated by entropy favorable release of water molecules at the DNA/protein and protein/protein interfaces (Table 1). Replacing the spacer C/G by A/T in *Gde1* (*Gde1SpA*), as in the *Ramp2* DR1, results in a 2 order of magnitude increase of its binding affinity. These quantitative data are in agreement with our semi-quantitative competition EMSA assays using transfected cell extracts ([26] and data not shown). These quantitative data reveal that differences in half-sites and spacer sequence strongly influence binding affinity and surprisingly show that RXR-DBD homodimers show the weakest affinity for the idDR1.

Crystal structures of the RXR-DBD homodimer bound to various DR1s. To analyze the effect of the DR1 sequences on RXR conformation, we examined the structures of RXR-DBD homodimers in complex with natural DR1s from the *Ramp2*, *Nr1d1* and *Gde1SpA* loci and with the idDR1 by X-Ray crystallography. Note that the same RXR DBD construct in identical buffer and DNA excess were used for the complexes and that all complexes crystallized in similar conditions. The crystal structures of the RXR-DBD-*Ramp2*, RXR-DBD-*Nr1d1,* RXR-DBD-*Gde1SpA* and RXR-DBD-idDR1 complexes were solved at 2.07Å, 2.00 Å, 2.35Å and 2.34 Å, respectively. The RXR-DBD-*Ramp2* and RXR-DBD-*Nr1d1* complexes crystallized in the C2 space group with 1 homodimer-DNA complex per asymmetric unit and the RXR-DBD-*Gde1SpA* complex in $P2_12_12_1$ with 2 homodimer-DNA complexes per asymmetric unit that are arranged in an anti-parallel fashion with a rotation of approximately 190° along the helical axis (Supplementary Fig. S1). For the *Gde1SpA* complex, the two first nucleotides before the first half-site are distorted and do not form canonical base pairing interactions as a consequence of crystal packing and the crystal contacts are formed by mainly non-canonical base interactions. The RXR-DBD-idDR1 crystallized in the $P2_12_12_1$ space group that differs from the previously published crystal structure[33] and contained 2 homodimer-DNA complexes per asymmetric unit (Supplementary Fig. S1).

All DR1s are bound asymmetrically by an RXR-DBD homodimer (Fig. 2a) and show regular B-form DNA structure with similar degrees of deformation (Fig. 2b) as shown in the minor groove width plot (Fig. 2c), however differences are observed between the different DR1s. The conformations of the DBD subunits are conserved in the different complexes with a root mean square deviation (RMSD) of the protein backbone Cα atoms of about 0.3 to 0.4 Å. Comparison with the published RXR-DBD homodimer structure on the idDR1 indicates a higher RMSD of 0.7Å.

The tertiary structure of the RXR-DBD is similar to those previously reported and is composed of an N-terminal β-hairpin, two α-helices followed by a single turn of 3_{10}-helix, and a C-terminal extension (Fig. 2a). The N-terminal α-helix (helix I) directly interacts with the DNA half-site in the major groove. Helix II is perpendicular to the N-terminal helix I and stabilizes the core of the DBD. For the 5′ subunit, a third short α-helix is observed for the Zn-II region. The DBD monomers lie in a head to tail orientation with non-equivalent protein-protein interactions from each monomer. Helix I of each RXR-DBD forms direct and water mediated base contacts that involve highly conserved residues (Fig. 2d-e for the *Ramp2* complex). Protein-DNA contacts are summarized schematically in Fig. 3a and Supplementary Fig. S2 for the idDR1, *Ramp2*, *Nr1d1* and *Gde1SpA* complexes.

Recognition of natural DR1s. Although, the overall fold of the RXR homodimer structures is similar and the key specific base contacts are conserved between the different structures, significant differences

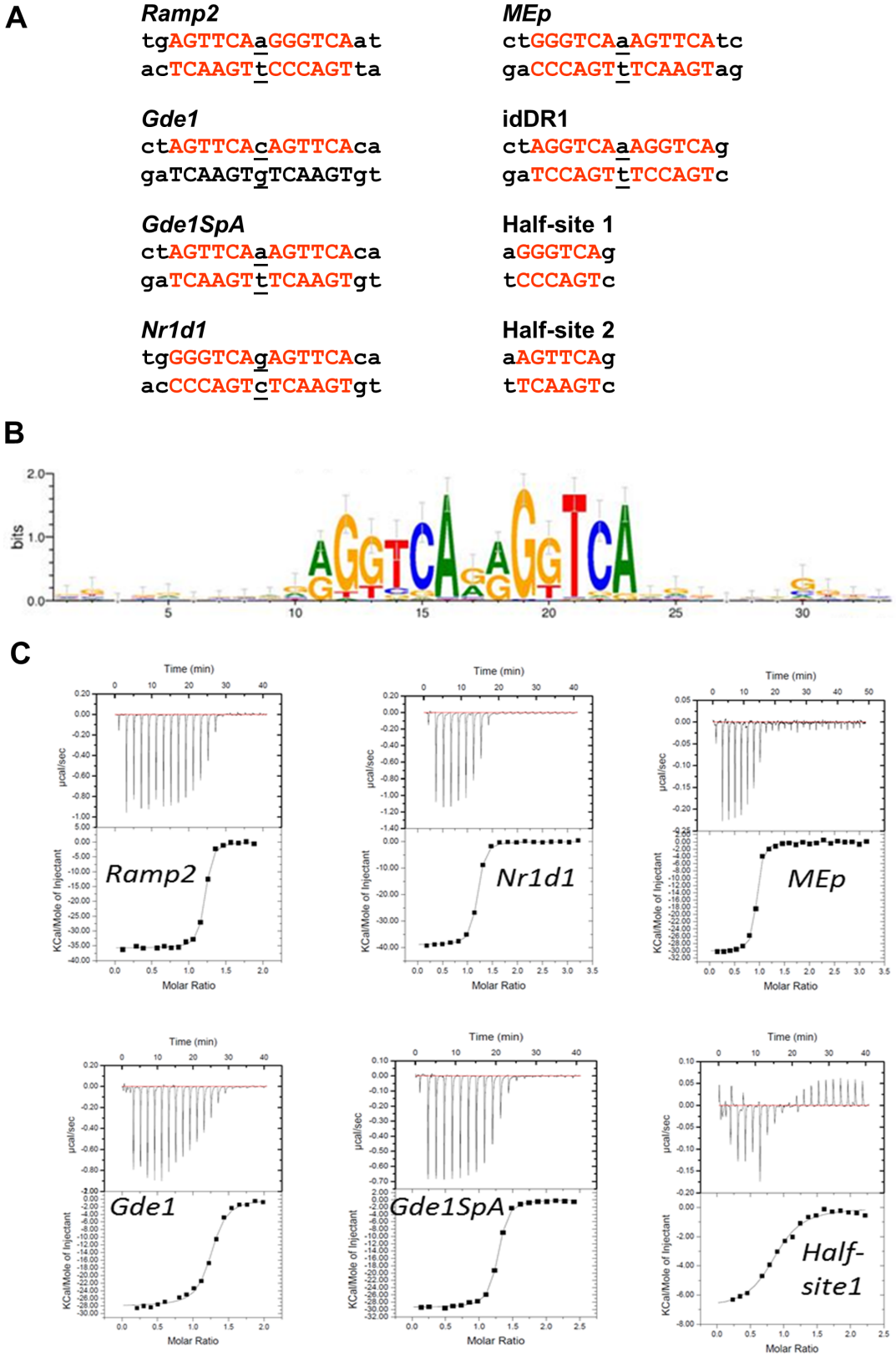

Figure 1 | (a) Natural DR1 response element double strand nucleotide (ds) sequences of DNA DR1 response elements. Hexanucleotide half-site motifs are shown in red. (b) RXR binding motif identified by RXR-ChIP sequencing from Ref. 28. (c) Quantification of the interaction between RXR and DR1s by ITC. Representative ITC isotherms for the binding of the DR1 duplex (*Ramp2*, *Nr1d1*, *MEp*, *Gde1*, *Gde1SpA* and half-site 1) to the RXR-DBD. The top panels show the raw ITC data expressed as the change in thermal power with respect to time over the period of titration. Lower panels: change in molar heat is expressed as a function of molar ratio of corresponding DR1 to dimer-equivalent RXR or half-site to monomer RXR. The solid lines in the lower panels represent the fit of data to a one-site model using the ORIGIN software. Standard free energies of binding and entropic contributions were obtained, respectively, as $\Delta G = -RT \ln(Ka)$ and $T\Delta S = \Delta H - \Delta G$, from the Ka and ΔH values derived from ITC curve fitting.

Table 1 | Quantification of the interaction between the RXR DBD and DNA by ITC. All data were obtained at 25°C (T = 298°K). Values represent the means values of 2–4 independent experiments and errors correspond to one standard deviation. N corresponds to the number of moles of dimer per mole of DNA except for the half-sites where N corresponds to the number of monomer per mole of DNA

DNA	K_d nM	N	ΔH_{obs} kcal × mol^{-1}	ΔS_{obs} cal × mol^{-1} × deg^{-1}
Ramp2-DR1	38 ± 15	1.1	−39 ± 3	−98
Gde1-DR1	140 ± 20	1.3	−25 ± 3	−51
Gde1SpA-DR1	44 ± 5	1.3	−29 ± 1	−66
Nr1d1-DR1	40 ± 5	1.1	−38 ± 1	−92
idDR1	155 ± 40 350[&]	1.2	−26 ± 1	−55
MEp-DR1	10 ± 2	1.1	−30 ± 0.3	−64
Half-site 1	980 ± 110	1.0	−11 ± 1	−9
Half-site 2	820 ± 150	1.0	−6 ± 1	6

[&]measured by fluorescence anisotropy in Ref. 33.

are observed between the idDR1 and the natural DR1s which display additional direct and water-mediated interactions and with equivalent amino acids generating different interactions. In the *Ramp2* and *Nr1d1* complexes, each half-site makes specific contacts through 5 of the 6 base pairs (Fig. 3a). For the *Gde1SpA* complex, less specific contacts are formed with only 3 base-pairs for the first half-site, but 5 base-pairs for the second half-site. The structure of RXR-DBD-idDR1 reported here revealed an increased number of specific interactions in each half-site compared to the previous RXR DBD idDR1 crystal structure[33], however fewer contacts are observed with the idDR1 compared to the natural DR1s (Fig. 3a). In the 3 complexes with natural DR1s, both RXR-DBDs also form extensive contacts with the phosphate backbone along 12 base pairs for *Ramp2*, 10 base pairs for *Nr1d1* and *Gde1SpA* and 8 base pairs for idDR1.

The second hexanucleotide half-site shows two variations between *Ramp2* (GGGTCA) and *Gde1SpA* or *Nr1d1* (AGTTCA). The first nucleotide differs (G in *Ramp2* and an A in *Gde1SpA* and *Nr1d1*), but a rearrangement of the Arg209 side chain maintains similar water-mediated interactions. The second variation in the second half-site is at the third nucleotide position which is a G/C pair in *Ramp2* and a T/A pair in *Gde1SpA* and *Nr1d1*. The introduction of a thymine residue at this position and the presence of a methyl group prevent a water-mediated interaction with Lys160 which adopts an alternative conformation to minimize its contact with the methyl group. Importantly however, several specific interactions are observed only in the complexes with the natural DR1s such as the interactions mediated by the first base pair of the 5′ half-site with Arg209 or the first base pair of the 3′ half-site with Lys156. In addition, the half-site spacer A/T base pair forms specific water mediated contacts with Arg209 of the T-box of the 3′ RXR-DBD in the *Ramp2* and *Gde1spA* complexes, but not in the idDR1 (Fig. 3b–c). The spacer in *Nr1d1* is a G/C pair that forms interactions similar to the A/T pair of *Ramp2*. However, the small size of the thymine base compared to guanine allows a water molecule to be trapped at the DNA surface and to mediate specific interactions in the minor groove. Modeling a replacement of the A/T spacer by a C/G pair indicates that the specific water mediated contact with Arg209 will not be formed thus lowering the binding affinity, in agreement with our quantitative ITC data. The T-box makes additional specific interactions in the minor groove with the first half-site involving Arg209 and Gln206 of the 5′ DBD and phosphate interactions for the 3′ DBD (Fig. 3a).

Differences between the 2 DBDs reflect the different modes of recognition between the 5′ and 3′ half-sites. Unexpectedly, in each case, the 3′ half-site contributes more interactions than the 5′ half-site. The comparison of the 2 DBDs of the *Ramp2* complex reveals several differences that reflect specific interactions with the half-site DNA and dimer interactions. The downstream DBD extends its interactions upstream of the 3′ half-site to reach the backbone sugar of the last nucleotide of the 5′ site (Fig. 3b) and contributes to enhanced RXR homodimerization. Analysis of the average temperature factor on the overall structure reveals a stabilization of the dimerization interface and of the 5′ bound RXR in the *Ramp2* complex compared to the idDR1 complex (Supplementary Fig. S3a) that hold true for the 2 dimers seen in the asymmetric unit.

Half-site DNA sequence specifies DBD orientation and dimerization. The DR1 sequences dictate specific interactions with the RXR-DBDs, but also impact on the RXR homodimer conformation as shown in Figure 4. Together with DNA curvature, the RXR-DBD adjusts its conformation to changes in the DR1 sequence and the tight interactions between the monomers leads to increased cooperativity. The complexes of RXR-DBD with *Ramp2* and *Nr1d1* show similar positioning (Fig. 4a). Interestingly due to weaker interactions for the idDR1, the RXR-DBD conformation and the relative position of the upstream DBD differs significantly between the structures of RXR-DBD-*Ramp2* and the RXR-DBD-idDR1 when the 3′ DBDs are superimposed (Fig. 4c). Helix I and Helix II of the 5′ DBD are shifted by 1 Å while the D-box region is shifted by 2 Å. The quality of the electron density maps for this loop in the *Ramp2* and idDR1 structures are shown in Supplementary Fig. S4. The two homodimers present in the asymmetric unit of the RXR-DBD-idDR1 show similar deviations compared to RXR-*Ramp2* complex, these conformational changes are thus independent of the crystal environment. This difference in 5′RXR positioning between RXR-DBD-*Ramp2* and RXR-DBD-idDR1 is also observed in the previously published RXR-idDR1 structure with even larger differences (Supplementary Fig. S5a).

The RXR-DBD-*Gde1SpA* complex shows an intermediate conformation (Fig. 4b). While the overall fold of the receptors and their placement on the half-sites are very similar, notable differences can be attributed to the dimerization contacts and specific DNA interactions (Fig. 4e–f). These conformational changes induced by the DR1 sequence should also impact on the position of the RXR hinge domain. Tighter RXR homodimer complexes are observed on the natural *Ramp2* or *Nr1d1* DR1s compared to the symmetric idDR1, but also compared to the crystal structures of RAR-RXR DBD (Fig. 4d) or full-length PPAR-RXR with idealized DR1 (Supplementary Fig. S5b).

The above changes in DBD orientation have important effects on DBD-dimerization. The dimerization interface faces the minor groove and is formed by the C-terminal T-box of the 3′ DBD with the Zn-II region of the 5′ DBD which forms a short α-helix. Dimerization in the RXR-DBD-*Ramp2*, RXR-DBD-*Nr1d1* and RXR-DBD-*Gde1SpA* complexes involves numerous Van der Waals interactions and hydrogen bonds. For the *Ramp2* complex, which shows the largest dimerization interface, numerous Van der Waals interactions (36 at a cutoff of 4Å) and H-bonds are observed between

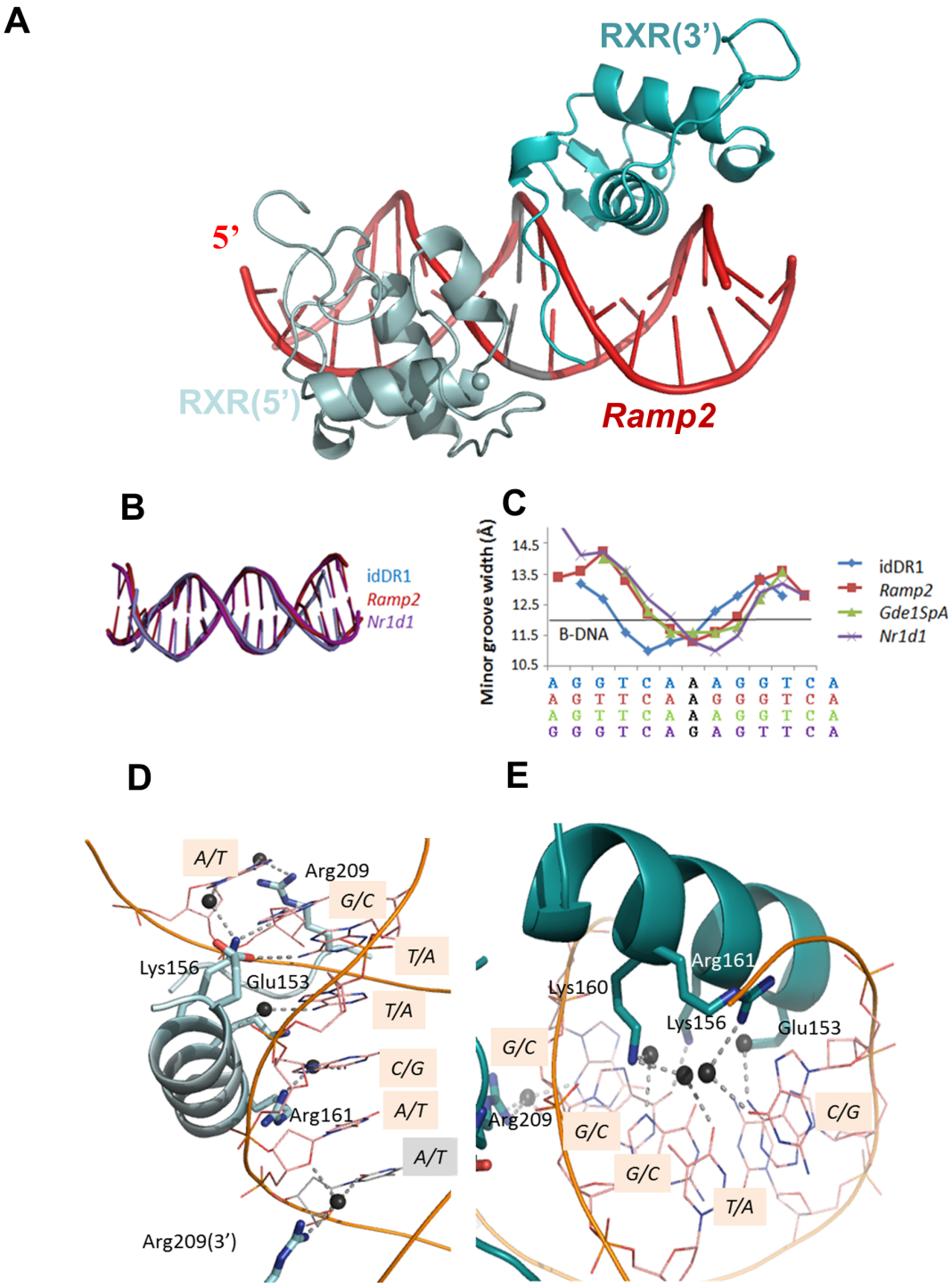

Figure 2 | (a) Overall structure of RXR DBD-*Ramp2*. The upstream RXR (in light cyan) and downstream RXR (in cyan) bound to their hexanucleotide motifs shown in red. The spheres indicate the Zn molecules. (b) Comparison of DNA bending of the DR1 elements. The ds oligonucleotides used in the crystallographic structures of RXR-DBD homodimers complexed with DR1, show similar deformation. (c) Plot of the minor groove widths of the DR1 ds oligonucleotides. The values were derived using the 3DNA software. The solid black line represents standard values for B-DNA. (d-e) RXR homodimers exhibit specific interactions and polarity on natural DR1s. *Ramp2* DNA sequence recognition by the upstream RXR subunit (d). View along the DNA-recognition helix (α1) of RXR showing residues Glu153, Lys156, Arg161 and Arg209 and their direct and water-mediated base contacts. Hydrogen-bonds and water molecules are shown as dotted blue lines and dark spheres, respectively. The interspacer nucleotide is highlighted in grey. The corresponding view of *Ramp2* DNA sequence recognition by the downstream RXR subunit (e).

Glu207 (3′ DBD) and Arg182 (5′ DBD) and between Gln210 (3′ DBD) and Arg186 (5′ DBD) (Fig. 4e). In contrast, the structure of RXR with the idDR1 shows only one H-bond and few hydrophobic interactions in the dimer interface (Fig. 4f). Consequently, a marked increase in the dimer buried surface area is observed in the RXR complexes with natural DR1s (340 Å^2 for *Ramp2*, 300 Å^2 for *Nr1d1*, 310 Å^2 for *Gde1SpA* and 190 Å^2 for idDR1 (140 Å^2 for the previous structure)) (Supplementary Fig. S6). Overall, increased

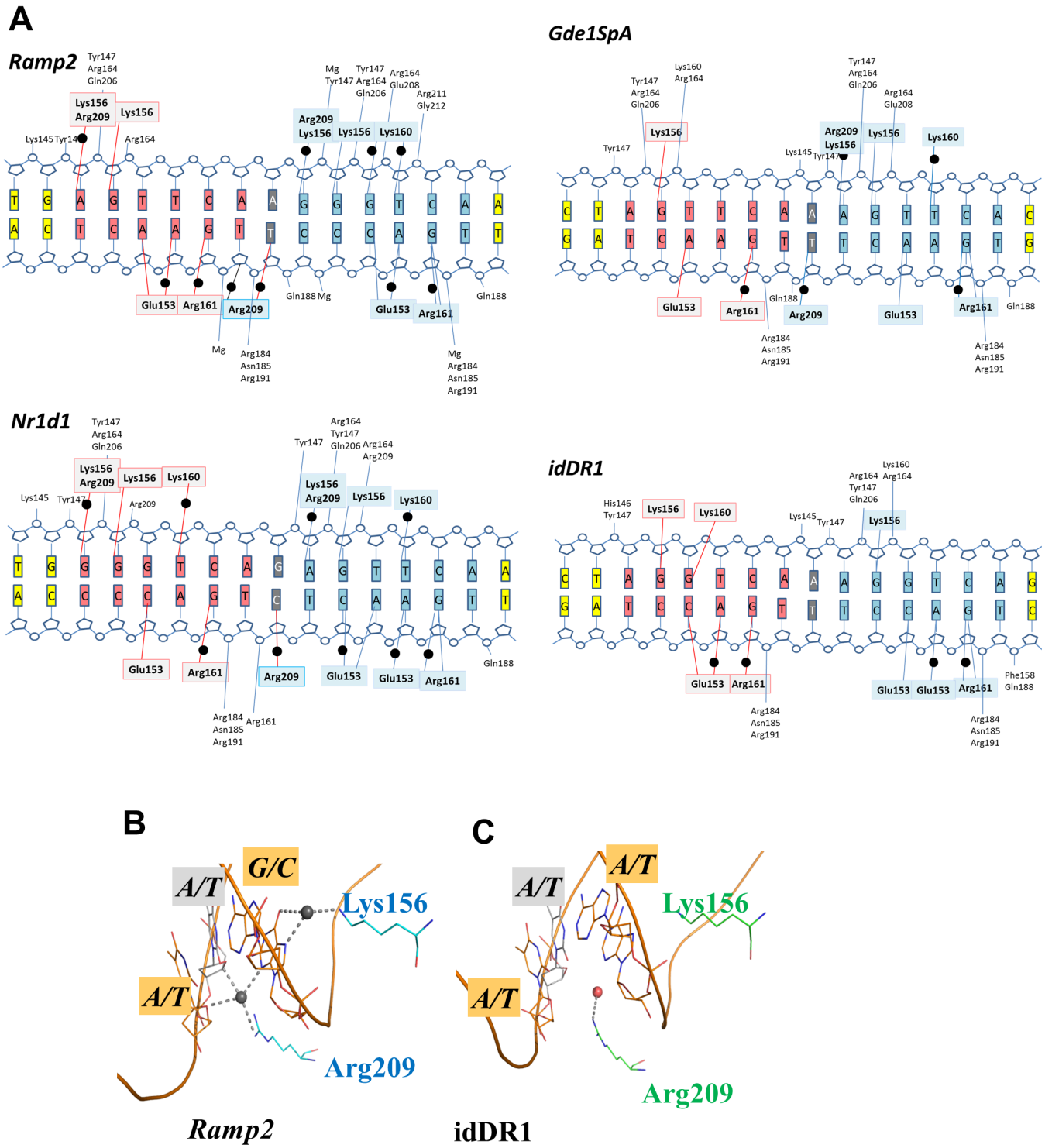

Figure 3 | Polarity of the bound RXR homodimers revealed by asymmetric DNA recognition of natural DR1s. The RXR-DBDs establish unique interactions to recognize natural asymmetric DR1s that are more numerous than in the complex with symmetric idDR1 as revealed by the schematic view of the protein/DNA contacts calculated with NUCPLOT with a 3.9 Å distance cutoff (a). Note that all crystal structures have comparable resolution (between 2 Å and 2.35 Å). For the three natural DR1s studied, the 3′ half-site is more tightly bound with more observed interactions with RXR. Bridging water molecules are shown as black circles. The residues forming hydrogen-bond interactions with the 5′ subunit and 3′subunit are highlighted in light grey and cyan, respectively. The first hexanucleotide is shown in salmon, the second one in blue and the interspacer nucleotide in grey. The gray circles indicate the DNA phosphates and the labeled residues their contacts with RXR. (b-c) Comparison of the interactions around the interspacer nucleotides of the natural *Ramp2* and the idealized DR1 complexes. The interspacer is highlighted in grey and the surrounding base pairs in orange. Only the RXR DBD complex with natural DR1 forms specific H-bonds with the interspacer nucleotide. For *Ramp2*, the downstream DBD extends its interactions upstream of the 3′ half-site to reach the backbone sugar of the last nucleotide of the 5′ site. H-bonds are shown by dashed lines and the water molecules as dark spheres for *Ramp2* and red sphere for the idDR1.

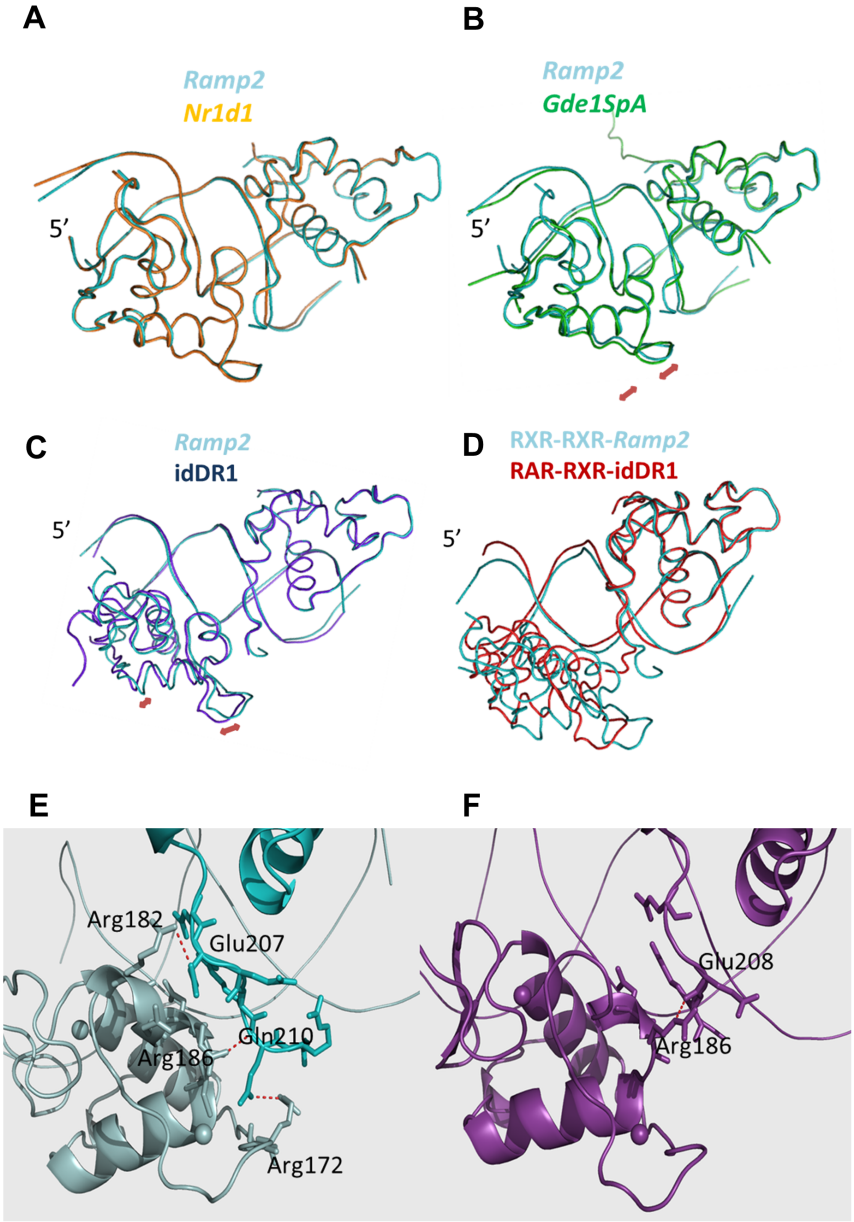

Figure 4 | Structural changes of RXR-DBDs induced by the half-site sequence. Significant differences in RXR-DBD positioning and in the dimerization interface are observed for the natural DR1s. (a–c) Superimposed crystal structures of RXR-DBD-*Ramp2* (cyan) with RXR-DBD-*Nr1d1* (a; orange), RXR-DBD-*Gde1SpA* (b; green) and RXR-DBD-idDR1 (c; blue). Superposition was performed on the 3′ bound RXR-DBD. The largest differences are observed for the idDR1 as indicated by red arrows. Differences are also observed for the complex with the natural *Gde1SpA* DR1. (d) Superimposed crystal structures of RXR DBD-*Ramp2* (cyan) with RAR-RXR DBD-idDR1 (PDB ID: 1XDK, red). Superposition was performed on the 3′ bound RXR-DBD. (e–f) Dimerization interface that involves the DNA minor groove, hydrogen-bonding between atoms of the 2 subunits (residues highlighted) and Van der Waals interactions for RXR-DBD-*Ramp2* (e) and RXR-DBD-idDR1 (f).

binding affinity for the different DR1 elements therefore correlates with the enhanced protein-protein and base specific interactions observed in the crystal structures revealing a higher specificity for the natural DR1s compared to the idDR1.

Discussion

Here we describe for the first time at the atomic level how variations in the consensus half-site sequence of DR1 elements regulate DNA binding by the RXR-DBD. All of the DR1s investigated in this study comprise variants of the recognized consensus RGKTCA half-site sequence, yet these elements show large differences of up to one order of magnitude in their ability to bind RXR-DBD homodimers *in vitro*. The natural DR1s used here were identified in chromatin immunoprecipitation and EMSA assays as binding RAR-RXR[26]. Our ITC and structural data shows that these DR1s are bound with high affinity by RXR-DBD homodimers. Our data reveal that precise receptor-DNA contacts can vary with modest changes in orientation and conformation of the receptor, in agreement with previous data[23,24,34].

The polarity and strength of the dimer is imposed not only by the half-site spacing, but also by the nature of the half-site sequences whose contacts with the DBD modulate DBD geometry and the interactions between the two DBDs. In all three complexes with the natural DR1s, the 3′ half-site shows a larger number of specific interactions and the 3′ RXR-DBD extends its interaction to the last nucleotide of the 5′ half-site. The tighter interactions with the 3′ half-site suggest that this site is occupied first, favoring the cooperative formation of the homodimer-DR1 complex. These observations are in agreement with the observed polarity of the RAR-RXR and PPAR-

RXR heterodimer-DR1 complexes, where RXR occupies the 3′ half-site of DR1[18]. Notably, the RXR-DBD contacts with the 3′ half-site are also stronger than those of the RAR-DBD or PPAR-DBD with the 5′ half-site seen in the corresponding heterodimer structures[15,16]. The tight interaction of the RXR-DBD with the 3′ half-site revealed by our structures therefore provides an explanation for this preferred polarity.

The differences in affinity can be explained not only by the differential recognition of the half-sites by the RXR-DBD, but we also highlight the important role played by the spacer base pair. Previous EMSA studies and the quantitative ITC performed here show that the C/G base pair in the *Gde1* element hinders binding of both full-length RAR-RXR and RXR-DBD homodimers. Examination of the crystal structure indicates that the RXR-DBD cannot form a specific hydrogen bond with the C/G pair that is seen with the A/T pair in the *Ramp2* and *Gde1SpA* complexes. Moreover, the *Nr1d1 and MEp* DR1s share identical half-sites, but the *Nr1d1* element has a significantly lower affinity, that may be explained by the presence of a G/C spacer that prevents specific interactions in the minor groove seen with the A/T spacer in *MEp*. It is interesting to note that analysis of RXR and indeed RAR ChIP-seq data reveal that G/C and A/T are strongly enriched in the consensus DR1 whereas C/G is not represented. Our structures thus provide a molecular explanation for this preference for the spacer base pair. Similarly, we previously found that the sequence 5′-AG-3′ is strongly preferred as spacer in DR2 elements[26] suggesting that critical contacts between the RAR-RXR-DBDs and the spacer also play an important role in DR2 recognition.

As the RXR DBDs do not dimerise in absence of DNA binding, our data suggest a model in which the RXR-DBD first binds to the 3′ half-site of the DR1 response elements. The conformation imparted to this DBD by the sequence through the DBD-DNA contacts then promotes binding of the second DBD to the 5′ half-site. This mode of recognition is however specific to the isolated DBDs, as full-length RXRs recognize the DR1 as preformed heterodimers. Nevertheless, tighter binding of the RXR-DBD to the 3′ half-site likely determines the orientation of the heterodimers on the DR1. In this respect it is interesting to note that the consensus DR1 sequence derived from PPAR ChIP-seq experiments displays a higher conservation in the 3′ half-site consistent with the idea tight RXR binding to this site is critical in determining heterodimer orientation[12,35]. However, the extended DNA contacts that PPAR makes 5′ to the first half-site likely also contribute to the polarity. The example of RXR may reflect a more general mode of recognition in which preferential binding of a DBD to one of the two half-sites promotes heterodimer orientation on other types of DR and IR elements. Further structural studies will be required to determine whether it is preferential recognition by the RXR-DBD that is critical or whether it is the DBD of the heterodimerisation partner that plays the determining role.

The structural data rationalized the differences observed in the thermodynamic parameters with additional H-Bonds accounting for favorable enthalpy and for entropy, a balance between favorable entropy of dehydration and less favorable entropy of the bridging water molecules to fulfill specific interactions. Increased DBD-DNA contacts made in a sequence- dependent manner with each half-site in turn strongly influence the orientation of the two DBDs and the DBD-DBD interactions at the dimerization interface and correlate with the highest overall DNA binding affinity. In this way, half-site sequence controls the cooperativity that results from the protein-protein contacts that the DBDs form within the spacer minor groove. The structure of RXR-DBD-*Ramp2* illustrates the structural adaptability of the DBD to accommodate its association with specific response elements and to optimize cooperative binding. Notable differences are observed mainly in the second zinc motif and the D-box that form the dimerization interface. This region has been shown to be significantly different between the DNA-free state seen in the NMR structure and in complex with DNA[31]. The differing affinities of natural DR1s therefore result from differential DBD-DNA interactions and their effects on DBD-DBD interactions. The DNA sequence-dependent regulation of RXR-DBD-DBD interactions on the DR1 described here are reminiscent of those described for GR where the GR DBD bound to a series of natural GREs was shown to exhibit distinct conformation in the loop connecting H1 and the D-box, identified as the lever arm[24]. These subtle structural changes upon DNA binding have been shown to allosterically affect the recruitment of coactivators and GR transcriptional activity[24,36]. Allosteric control of coactivator recruitment by the DNA sequence has also been observed in case of the RAR-RXR heterodimer[37] and the RXR homodimer[21]. More recent studies have revealed that the dimerization interactions and positive or negative cooperativity are major determinants in transcriptional activation or repression by GR[32,36].

The DR1 configuration is also a promiscuous HRE for the RAR-RXR and PPAR-RXR heterodimers. As for RXR homodimers, the RXR subunit in the heterodimer uses its T-box to mediate DBD dimer interactions. It is probable that interactions of the RXR-DBD with its heterodimerization partners will also be allosterically regulated by the half-site DNA sequence analogous to what is observed for the RXR-DBD homodimers. Our study therefore provides a molecular model for how half-site sequence in general and the sequence of the 3′ half-site in particular of DR1s may exert an allosteric regulation on DBD homo and heterodimers through changes in contacts at the dimerization interface. This allosteric regulation of DBD positioning may propagate to other domains of the receptor and hence contribute to the fine-tuning of transcription.

Methods

Constructs, expression and purification. The HsRXRα-DBD (130-212) was expressed in fusion with Thioredoxine and hexahistidine tags. Fusion proteins were removed by thrombin proteolysis. The cleaved protein was then purified by gel filtration. The oligonucleotide strands were purchased from SIGMA and annealed as described previously[38], added in a 1.2 fold excess to the dimers, and the complex was gel-filtrated on a Superdex S200.

Crystallization and structure determination. The crystallization experiments were carried out by sitting drop vapor diffusion at 290 K using a Cartesian nanoliter dispensing robot and mixing equal volumes (0.1 µl) of protein-DNA complex and reservoir solution. RXRα-DBD, RARα−DBD and DNA were mixed in an equimolar ratio in 25mM Tris pH 8, 50mM NaCl, 5mM MgCl2, 1mM TCEP and concentrated to a final concentration of 9.6 mg ml^{-1} for the *Ramp2* DR1 complex, 9.9 mg ml^{-1} for the *Nr1d1* complex, 14.6 mg ml^{-1} for the *Gde1SpA* complex and 10.0 mg ml^{-1} for the *idDR1* complex. Although RAR and RXR DBDs were mixed together with the DNA in the crystallization experiments, only crystals of RXR homodimer-DNA were obtained. Crystals of *Ramp2* complex were grown in 20% PEG 3350, 0.2 M NH_4Cl, 0.1 M $MgCl_2$, 0.1 M MES pH 6, crystals of *Nr1d1* in 20% PEG 3350, 0.2 M KNO_3, crystals of *Gde1SpA* complex appeared in 20% PEG 3350, 0.2 M NH_4Cl and crystals of the idDR1 complex appeared in 20 % PEG 3350, 0.3M NH_4Cl, 0.1M $MgCl_2$, 0.1M sodium citrate pH 5.0. The crystals of the *Ramp2* complex and the *Gde1SpA* complex were transferred to artificial mother liquor containing 20% PEG 400 and flash cooled in liquid nitrogen. The crystals of the *Nr1d1* complex and the idDR1 complex were transferred to artificial mother liquor containing 35 % PEG 3350 before flash cooling.

For the *Ramp2* DR1 complex, data were collected at the zinc edge (1.2825Å) on a Quantum 315r CCD detector (ADSC) at the ID23-1 beamline of the ESRF. A total of 180° of data were collected using 0.5° rotation and 0.5s exposure per image (50% attenuated beam). For the *Nr1d1* DR1 complex, data were also collected at the zinc edge (1.2833Å) on a Quantum 315r CCD detector at the BM30A beamline of the ESRF. A total of 180° of data were collected using 1° rotation and 4s exposure per image (unattenuated beam). For the *Gde1SpA* DR1 complex and the *idDR1* complex, data were collected at 0.873Å on a MX-225 CCD detector (Marresearch) on the ID23-2 beamline of the ESRF. For the *Gde1SpA* complex a total of 180° of data were collected using 0.3° rotation and 5s exposure per image (unattenuated beam). For the *idDR1* complex a total of 111° of data were collected using 0.2° rotation and 0.91s exposure per image (unattenuated beam). The data for the *Ramp2* and *Gde1SpA* complexes were indexed, integrated, and scaled using HKL2000[39]. The data for the *Nr1d1* and the *idDR1* complexes were indexed and integrated using XDS[40] and scaled using AIMLESS[41–43]. The crystals of the *Ramp2* complex belonged to space group C2 with unit cell parameters a = 113.2Å, b = 44.0Å, c = 63.3Å, β = 106.016°. The structure was solved by molecular replacement in PHASER[44] using the structures of the RAR and RXR DBDs bound to 6 base pairs of DNA[15] as search models. The asymmetric unit contains one copy of the homodimer of RXR bound to the *Ramp2*

DR1 response element, with a corresponding Matthews' coefficient[45] of 2.45 Å^3/Da and a solvent content of 55.3%. The crystals of the *Nr1d1* complex also belonged to space group C2, with slightly smaller unit cell dimensions a = 103.3Å, b = 44.3Å, c = 63.9Å, β = 98.95°. The structure was solved in PHASER as above and the asymmetric unit contains one copy of the homodimer of RXR bound to the *Nr1d1* DR1 response element, with a corresponding Matthews' coefficient of 2.33 Å^3/Da and a solvent content of 52.9%. The crystals of the *Gde1SpA* complex belonged to space group $P2_12_12_1$ with unit cell parameters a = 53.1Å, b = 69.5Å, c = 139.1Å. The structure was solved by molecular replacement in MOLREP[46] using the RAR- and RXR-DBDs from the structure 1DSZ as probes. The asymmetric unit contains two copies of the homodimer of RXR bound to the *Gde1SpA* DR1 response element, with a corresponding Matthews' coefficient of 2.04 Å^3/Da and a solvent content of 46.1%. The crystals of the *idDR1* complex belonged to space group $P2_12_12_1$, with unit cell dimensions a = 37.63Å, b = 65.35Å, c = 209.12Å. The structure was solved in PHASER as above and the asymmetric unit contains two copies of the homodimer of RXR bound to the idDR1 response element, with a corresponding Matthews' coefficient of 2.33 Å^3/Da and a solvent content of 52.9%. Refinement of all structures was performed using PHENIX[47] and BUSTER[48] followed by iterative model building in COOT[49]. The quality of the refined model was assessed using MOLPROBITY[50] and PROCHECK[51]. Data collection and refinement statistics are given in Supplementary Table S1. Structural figures were prepared using PyMOL (www.pymol.org/).

Isothermal titration calorimetry experiments. ITC measurements were performed at 25°C on a MicroCal ITC200 (MicroCal). Purified proteins and DNA were dialyzed extensively against the buffer 25 mM Hepes pH 8.0, 50 mM sodium chloride and 1 mM TCEP for the DBDs. In a typical experiment 2 µl aliquots of DNA at 80 to 150 µM were injected into a 10 µM RXR dimer solution (200 µl sample cell). The c values (c = K*M*n) were in the optimal limits ($10 \leq c \leq 500$). The delay between injections was 120 to 180 s to permit the signal to return to baseline before the next injection. To extract various thermodynamic parameters, the binding isotherms were iteratively fit to an one-site model by non-linear least squares regression analysis using the software Origin 7.0 (OriginLab) as described[52]. Standard free energies of binding and entropic contributions were obtained, respectively, as $\Delta G = -RT\ln(Ka)$ and $T\Delta S = \Delta H - \Delta G$, from the Ka and ΔH values derived from ITC curve fitting.

Accession numbers. The coordinates and structure factors are deposited in the Protein Data Bank under the accession codes 4CN2 (RXR-DBD-*Ramp2*), 4CN3 (RXR-DBD-*Gde1SpA*), 4CN5 (RXR-DBD-*Nr1d1*) and 4CN7 (RXR-DBD-idDR1).

1. Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–9 (1995).
2. Altucci, L. & Gronemeyer, H. Nuclear receptors in cell life and death. *Trends Endocrinol Metab* **12**, 460–8 (2001).
3. Aagaard, M. M., Siersbaek, R. & Mandrup, S. Molecular basis for gene-specific transactivation by nuclear receptors. *Biochim Biophys Acta* **1812**, 824–35 (2011).
4. Huang, P., Chandra, V. & Rastinejad, F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annu Rev Physiol* **72**, 247–72 (2010).
5. Claessens, F. & Gewirth, D. T. DNA recognition by nuclear receptors. *Essays Biochem* **40**, 59–72 (2004).
6. Brelivet, Y., Kammerer, S., Rochel, N., Poch, O. & Moras, D. Signature of the oligomeric behaviour of nuclear receptors at the sequence and structural level. *EMBO Rep* **5**, 423–9 (2004).
7. Zhang, X. K. *et al.* Homodimer formation of retinoid X receptor induced by 9-cis retinoic acid. *Nature* **358**, 587–91 (1992).
8. Mangelsdorf, D. J. *et al.* A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. *Cell* **66**, 555–61 (1991).
9. Han, K., Moon, I. & Lim, H. J. All-trans- and 9-cis-retinoic acids activate the human cyclooxynase-2 gene: a role for DR1 as RARE or RXRE. *Mol Biol Rep* **38**, 833–40 (2010).
10. Germain, P. *et al.* International Union of Pharmacology. LX. Retinoic acid receptors. *Pharmacol Rev* **58**, 712–25 (2006).
11. Germain, P. *et al.* International Union of Pharmacology. LXIII. Retinoid X receptors. *Pharmacol Rev* **58**, 760–72 (2006).
12. Boergesen, M. *et al.* Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor alpha in mouse liver reveals extensive sharing of binding sites. *Mol Cell Biol* **32**, 852–67 (2012).
13. Szeles, L. *et al.* Research resource: transcriptome profiling of genes regulated by RXR and its permissive and nonpermissive partners in differentiating monocyte-derived dendritic cells. *Mol Endocrinol* **24**, 2218–31 (2010).
14. Fang, B., Mane-Padros, D., Bolotin, E., Jiang, T. & Sladek, F. M. Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. *Nucleic Acids Res* **40**, 5343–56 (2012).
15. Rastinejad, F., Wagner, T., Zhao, Q. & Khorasanizadeh, S. Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. *EMBO J* **19**, 1045–54 (2000).
16. Chandra, V. *et al.* Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. *Nature* **456**, 350–6 (2008).
17. Kurokawa, R. *et al.* Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. *Nature* **371**, 528–31 (1994).
18. Kurokawa, R. *et al.* Polarity-specific activities of retinoic acid receptors determined by a co-repressor. *Nature* **377**, 451–4 (1995).
19. Yang, Y. Z., Subauste, J. S. & Koenig, R. J. Retinoid X receptor alpha binds with the highest affinity to an imperfect direct repeat response element. *Endocrinology* **136**, 2896–903 (1995).
20. Castelein, H., Janssen, A., Declercq, P. E. & Baes, M. Sequence requirements for high affinity retinoid X receptor-alpha homodimer binding. *Mol Cell Endocrinol* **119**, 11–20 (1996).
21. IJpenberg, A. *et al.* In vivo activation of PPAR target genes by RXR homodimers. *EMBO J* **23**, 2083–91 (2004).
22. Han, K. *et al.* Utilization of DR1 as true RARE in regulating the Ssm, a novel retinoic acid-target gene in the mouse testis. *J Endocrinol* **192**, 539–51 (2007).
23. Schwabe, J. W., Chapman, L. & Rhodes, D. The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation. *Structure* **3**, 201–13 (1995).
24. Meijsing, S. H. *et al.* DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **324**, 407–10 (2009).
25. Jakob, M. *et al.* Novel DNA-binding element within the C-terminal extension of the nuclear receptor DNA-binding domain. *Nucleic Acids Res* **35**, 2705–18 (2007).
26. Delacroix, L. *et al.* Cell-specific interaction of retinoic acid receptors with target genes in mouse embryonic fibroblasts and embryonic stem cells. *Mol Cell Biol* **30**, 231–44 (2010).
27. Moutier, E. *et al.* Retinoic acid receptors recognize the mouse genome through binding elements with diverse spacing and topology. *J Biol Chem* **287**, 26328–41 (2012).
28. Mendoza-Parra, M. A., Walia, M., Sankar, M. & Gronemeyer, H. Dissecting the retinoid-induced differentiation of F9 embryonal stem cells by integrative genomics. *Mol Syst Biol* **7**, 538 (2011).
29. Zechel, C. *et al.* The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. *EMBO J* **13**, 1425–33 (1994).
30. Perlmann, T., Rangarajan, P. N., Umesono, K. & Evans, R. M. Determinants for selective RAR and TR recognition of direct repeat HREs. *Genes Dev* **7**, 1411–22 (1993).
31. Holmbeck, S. M., Dyson, H. J. & Wright, P. E. DNA-induced conformational changes are the basis for cooperative dimerization by the DNA binding domain of the retinoid X receptor. *J Mol Biol* **284**, 533–9 (1998).
32. Hudson, W. H., Youn, C. & Ortlund, E. A. The structural basis of direct glucocorticoid-mediated transrepression. *Nat Struct Mol Biol* **20**, 53–8 (2013).
33. Zhao, Q. *et al.* Structural basis of RXR-DNA interactions. *J Mol Biol* **296**, 509–20 (2000).
34. Gewirth, D. T. & Sigler, P. B. The basis for half-site specificity explored through a non-cognate steroid receptor-DNA complex. *Nat Struct Biol* **2**, 386–94 (1995).
35. Nielsen, R. *et al.* Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev* **22**, 2953–67 (2008).
36. Watson, L. C. *et al.* The glucocorticoid receptor dimer interface allosterically transmits sequence-specific DNA signals. *Nat Struct Mol Biol* **20**, 876–83 (2013).
37. Mouchon, A., Delmotte, M. H., Formstecher, P. & Lefebvre, P. Allosteric regulation of the discriminative responsiveness of retinoic acid receptor to natural and synthetic ligands by retinoid X receptor and DNA. *Mol Cell Biol* **19**, 3073–85 (1999).
38. Juntunen, K., Rochel, N., Moras, D. & Vihko, P. Large-scale expression and purification of the human vitamin D receptor and its ligand-binding domain for structural studies. *Biochem J* **344 Pt 2**, 297–303 (1999).
39. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* **276**, 307–326 (1997).
40. Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125–32 (2010).
41. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72–82 (2006).
42. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* **67**, 282–92 (2011).
43. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–3 (1994).
44. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458–64 (2005).
45. Matthews, B. W. Solvent content of protein crystals. *J Mol Biol* **33**, 491–7 (1968).
46. Vagin, A. & Teplyakov, A. A translation-function approach for heavy-atom location in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **54**, 400–2 (1998).
47. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–21 (2010).
48. Smart, O. S. *et al.* Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **68**, 368–80 (2012).
49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–32 (2004).

50. Chen, V. B. *et al*. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21 (2010).
51. Laskowski, R. A., Moss, D. S. & Thornton, J. M. Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049–67 (1993).
52. Deegan, B. J., Seldeen, K. L., McDonald, C. B., Bhat, V. & Farooq, A. Binding of the ERalpha nuclear receptor to DNA is coupled to proton uptake. *Biochemistry* **49**, 5978–88 (2010).

Acknowledgments

We thank Virginie Chavant for technical help at the beginning of the project, Tao Ye for the ChIP-seq analysis and the staff of the beamline at the ESRF and SOLEIL for the experimental assistance during data collection. The project was supported by the Centre National pour la Recherche Scientifique (CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM), the Agence Nationale de Recherche (ANR-11-BSV8-023), the Fondation pour la Recherche sur le Cancer (ARC), the Fondation pour la Recherche Médicale (FRM), Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) and supported by national member subscriptions, the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01).

Author contributions

N.R. designed the research. J.O., A.G.M., P.P.C., E.M. performed experiments. J.O., A.G.M., C.B., I.D., D.M. and N.R. analyzed data. J.O., I.D., D.M. and N.R. wrote the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Osz, J. *et al*. Structural Basis of Natural Promoter Recognition by Retinoid X Nuclear Receptor. *Sci. Rep.* **5**, 8216; DOI:10.1038/srep08216 (2015).