Interplay of Protein Disorder in Retinoic Acid Receptor Heterodimer and Its Corepressor Regulates Gene Expression

Tiago N Cordeiro, Nathalie Sibille, Pierre Germain, Philippe Barthe, Abdelhay Boulahtouf, Frédéric Allemand, Rémy Bailly, Valérie Vivat, Christine Ebel, Alessandro Barducci, et al.

Interplay of protein disorder in retinoic acid receptor heterodimer and its corepressor regulates gene expression

Tiago N. Cordeiro[1,2], Nathalie Sibille[1], Pierre Germain[1], Philippe Barthe[1], Abdelhay Boulahtouf[3], Fréderic Allemand[1], Rémy Bailly[1], Valérie Vivat[4], Christine Ebel[5], Alessandro Barducci[1], William Bourguet[1], Albane le Maire[1,6,*], Pau Bernadó[1,7,*]

[1]Centre de Biochimie Structurale (CBS). CNRS, INSERM, Université de Montpellier. 29, rue de Navacelles. 34090 Montpellier, France

[2]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2790-157 Oeiras, Portugal

[3]IRCM, INSERM, ICM, Univ Montpellier, 34298 Montpellier, France

[4]NovAliX, 67400 Illkirch, France

[5]IBS, Univ. Grenoble Alpes, CEA, CNRS, 38000 Grenoble, France

[6]Brazilian Biosciences National Laboratory (LNBio), Brazilian Center for Research in Energy and Materials (CNPEM), 13083-970 Campinas, Sao Paulo, Brazil.

[7] Lead contact

*Correspondence: albane.lemaire@cbs.cnrs.fr; pau.bernado@cbs.cnrs.fr

21 **Summary**

22 In its unliganded form, the retinoic acid receptor (RAR) in heterodimer with the

23 retinoid X receptor (RXR) exerts a strong repressive activity facilitated by the

24 recruitment of transcriptional corepressors in the promoter region of target genes. By

25 integrating complementary structural, biophysical, and computational information, we

26 demonstrate that intrinsic disorder is a required feature for the precise regulation of

27 RAR activity. We show that structural dynamics of RAR and RXR H12 regions is an

28 essential mechanism for RAR regulation. Unexpectedly, we found that, while mainly

29 disordered, the corepressor N-CoR presents evolutionary conserved structured regions

30 involved in transient intramolecular contacts. In the presence of RXR/RAR, N-CoR

31 exploits its multivalency to form a cooperative multi-site complex that displays an

32 equilibrium between different conformational states, that can be tuned by cognate

33 ligands and receptor mutations. This equilibrium is key to preserve the repressive

34 basal state while allowing the conversion to a transcriptionally active form.

35

36 **Introduction**

37   Nuclear Receptors (NRs) are transcription factors that have a direct role in

38 regulating the expression of ligand-responsive genes. This regulatory capacity of NRs

39 occurs through their ability to recognize specific sequences in the promoters of their

40 target genes and their relationships with the RNA polymerase II holocomplex as well

41 as the chromatin environment that surrounds these genes (Roeder, 1998). Like many

42 other members of the NR family, retinoic acid receptors (NR1B1 (RARα), NR1B2

43 (RARβ) and NR1B3 (RARγ)) form heterodimeric complexes with the retinoid X

44 receptors (NR2B1 (RXRα), NR2B2 (RXRβ) and NR2B3 (RXRγ)) and function as

45 ligand (retinoic acid)-regulated transcription factors (Gronemeyer et al., 2004; Perissi

46 and Rosenfeld, 2005). In fact, RXR/RAR heterodimers may act either as repressors or

47 activators of gene transcription depending on their ligation status that in turn

48 determines the ability of these DNA-bound receptors to recruit coregulators (either

49 corepressors or coactivators) to target gene promoters (Perissi and Rosenfeld, 2005).

50 Coactivator recruitment is usually ligand-dependent, whereas corepressors interact in

51 most cases with unliganded (apo) receptors.

52 RARs and RXRs have a conserved modular structure with an N-terminal

53 activation function (AF-1), a central DNA-binding domain (DBD) and a C-terminal

54 ligand-binding domain (LBD) (Germain et al., 2006). The multifunctional LBD is

55 responsible for ligand binding and dimerization and contains a ligand-dependent

56 activation function (AF-2), which corresponds to coregulator interaction surfaces that

57 can be modulated by natural (e.g. retinoic acid) or pharmacological ligands

58 (Gronemeyer et al., 2004). Over the years, different classes of synthetic ligands have

59 been generated to induce or repress gene transcription. While agonists enhance the

60 recruitment of coactivators and destabilize interaction with corepressors, thus

61 inducing the transcription of target genes, inverse agonists do the opposite, decreasing

62 the basal transcriptional activity of apo-receptors (Germain et al., 2002; le Maire et

63 al., 2012). Neutral antagonists inhibit both interactions and block the receptor in an

64 inactive conformation.

65 In the absence of ligand or in the presence of inverse agonists, RARα (named

66 RAR hereafter) exhibits strong repressive activity that is brought about by the

67 recruitment of corepressors (Glass and Rosenfeld, 2000; McKenna et al., 1999;

68 Germain et al., 2002; Germain et al., 2009; le Maire et al., 2010). The two main

69 corepressors, Nuclear receptor CoRepressor (N-CoR/NCoR1/RIP13) (Hörlein et al.,

70 1995) and the Silencing Mediator of Retinoic acid receptor and Thyroid hormone

3

71 receptors (SMRT/NCoR2/TRAC) (Chen and Evans, 1995), have been shown to reside

72 in, or recruit, high molecular weight complexes that display histone deacetylase

73 activity (Heinzel et al., 1997; Nagy et al., 1997). Deacetylated histones are associated

74 with silent regions of the genome, and it is generally accepted that histone acetylation

75 and deacetylation shuffle nucleosomal targets between a relaxed and condensed

76 chromatin configuration, the former being requisite for transcriptional activation.

77 From a mechanistic point of view, ligand binding to receptors induces a

78 rearrangement of the C-terminal region of the LBD, the so-called helix H12, leading

79 to corepressor dissociation and coactivator recruitment. Coactivators, such as those of

80 the TIF-2/SRC-1/RAC3 (p160) family, mediate the interaction of coactivator

81 complexes with NRs. CBP, p300, P/CAF, and some p160 coactivators themselves are

82 reported to act as histone acetyltransferases (HATs) (Glass and Rosenfeld, 2000;

83 Lonard and O'Malley, 2007). They are capable of acetylating specific residues in the

84 N-terminal tails of different histones, a process that is believed to play an important

85 role in the opening of chromatin during transcription activation (Chen and Evans,

86 1995; Imhof et al., 1997).

87 Like many proteins in signaling pathways, coregulators are mainly disordered

88 proteins that act as a platform where multiple proteins attach to perform activities

89 linked to gene transcription (Csizmok et al., 2016; Hegyi et al., 2007). These

90 interactions are mediated by Short Linear Motifs (SLiMs) that are embedded in the

91 sequence and that enable the simultaneous recognition of multiple partners (Van Roey

92 et al., 2014). Regarding the interaction with NRs, two major conserved

93 corepressor/NR recognition motifs (CoRNR box 1-2) LxxI/HIxxxI/L have been

94 identified in SMRT and N-CoR (Hu and Lazar, 1999; Nagy et al., 1999; Perissi et al.,

95 1999). The boxes, which are close in corepressor sequences, define the NR interaction

4

96 domain (NID). The co-existence of two NR interaction motifs in the NID of

97 corepressors (multivalency) raises the question of the actual mode of binding to

98 RXR/RAR heterodimer, namely, whether binding occurs to a single LBD (assymetric

99 model) or simultaneously to both LBDs using the two helical motifs (deck model).

100 Important information regarding the molecular mechanisms that regulate the

101 alternative interactions of RAR LBD, with either class of cofactors has been decoded

102 by crystallographic studies using short coregulator-derived peptides (Bourguet et al.,

103 2000; le Maire et al., 2010; Pogenberg et al., 2005). Whereas H12 is primarily

104 involved in the interaction with coactivators, the discovery of a specific interface

105 between RAR and a fragment (CoRNR1) of N-CoR revealed that a secondary-

106 structure transition affecting H11 plays a master role in corepressor association and

107 release. The constitutive interaction of RAR with corepressors involves: (i) the

108 formation of an antiparallel β-sheet between β-strand S3 of the receptor and β-strand

109 β1 of the corepressors, and (ii) the binding of the four-turn α-helix α1 of the

110 corepressors to the coregulator groove of RAR. Agonist binding induces the S3 to

111 H11 secondary-structure switch due to the stabilization of the H11 conformation, as

112 observed in the crystal structure of RAR in complex with the synthetic agonist

113 AM580 (le Maire et al 2010). On the contrary, in the structure of RAR in complex

114 with the inverse agonist BMS493 (le Maire et al., 2010), the RAR–corepressor

115 interaction is strengthened by the stabilization of the β-sheet S3–β1 interface.

116 However, multi-protein complexes containing intrinsically disordered

117 segments, such as the NID of N-CoR, have an extraordinary structural heterogeneity,

118 which poses significant technical challenges for their structural characterization.

119 Indeed, X-ray crystallography is best suited for studies of rigid folded domains and

120 tightly bound complexes. The application of solution Nuclear Magnetic Resonance

5

121 (NMR), which has been extensively used to study disordered proteins (Dyson and

122 Wright, 2004; Jensen et al., 2013), encounters severe limitations when characterizing

123 large biomolecular complexes. Lower resolution methods can synergistically

124 complement these high-resolution techniques. In particular, small angle X-ray

125 scattering (SAXS) offers a source of structural and dynamic information for highly

126 flexible biomolecules (Bernadó and Svergun, 2012; Cordeiro et al., 2017a). Hybrid

127 approaches, which integrate information from these different techniques into

128 computational tools are the most promising strategy for the structural characterization

129 of highly dynamic proteins and complexes in solution. Here, we report on the

130 structural and dynamic details of the RXR/RAR heterodimer, the disordered NCoR

131 NID, and the highly flexible complex that they form, by integrating solution

132 techniques and computational methods.

133 Our study reveals that, in addition of the two CoRNR motifs, a conserved

134 region of the NID is partially structured and forms transient intramolecular contacts.

135 Furthermore, we show that NID binds to RXR/RAR through both CoRNR motifs in a

136 highly cooperative manner inducing an equilibrium between several conformational

137 states. Perturbations on the individual CoRNR/NR affinities using RAR ligands and

138 point mutations have a global effect on the structure, dynamics and thermodynamics

139 of the complex. Moreover, we demonstrate that although both receptors contribute to

140 the interaction with N-CoR NID, RAR plays a dominant role over RXR. Thus, we

141 report novel insights into the structural basis of the recruitment of corepressors by

142 RXR/RAR heterodimer, emphasizing the interplay of N-CoR and H12 helix disorder-

143 to-order transitions in the regulation of NR-mediated gene transcription.

144

145

6

Results

Disorder in RXR/RAR H12 helices is modulated by ligands and mutations. The most remarkable observation of the multiple crystallographic structures of RXR and RAR reported so far is the conformational variability of their C-terminal helices H12 depending on the ligands and cofactors bound (Bourguet et al., 1995; Bourguet et al., 2000; Chandra et al., 2017; Germain et al., 2002; Germain et al., 2009; le Maire et al., 2010; Pogenberg et al., 2005; Renaud et al., 1995; Sato et al., 2010). Nuclear Magnetic Resonance (NMR) and fluorescence anisotropy have also shown the plasticity of these helices in solution depending on the ligation state (Lu et al., 2006; Nahoum et al., 2007). Here, we have used SAXS to validate these previous observations and to analyze other non-crystallized conditions.

SAXS data indicate that unliganded RXR/RAR heterodimer is a globular particle in solution with a radius of gyration, R_g of 26.6 ± 0.4 Å and a maximum intramolecular distance, D_{max} of 89.0 ± 3.0 Å (Table S1; Fig. 1A). Molecular weight estimation suggests that the particle is a heterodimer, in line with Sedimentation Velocity Analytical Ultracentrifugation (SV-AUC) experiments, s_{20w} = 3.78 ± 0.13 S and f/f_{min} = 1.30 ± 0.05 (Fig. S2). The smooth asymmetrical pair-wise distance distribution, $P(r)$, suggests the presence of moderate flexibility in the RXR/RAR heterodimer (Fig. 1A). Crystallographic structures of RXR/RAR (PDB entries 1DKF (Bourguet et al., 2000), 1XDK (Pogenberg et al., 2005), and 3A9E (Sato et al., 2010) (Table S2)), all obtained in the presence of agonist or antagonist of RXR and RAR, were not consistent with measured SAXS data (χ_i^2 = 2.55, 2.34 and 1.83, respectively). We hypothesized that the observed discrepancy can be attributed to two reasons: firstly, the missing flexible regions, including the N- and C-termini and the loops connecting helices 2 and 3 in both RXR and RAR (Fig. S1B), and secondly, the

7

171 positions of RXR and RAR H12 helices that are in antagonist or agonist-bound

172 conformations in the X-ray structures. To validate this hypothesis and based on

173 several crystallographic structures, we built three ensemble models for the RXR/RAR

174 heterodimer in which both H12 helices were maintained disordered, in agonist or in

175 antagonist conformation (see Fig. S1 and methods section for details). SAXS curves

176 derived from these ensemble models were compared with the experimental one. As

177 observed in Figure 1 and Figure S1, the average scattering profile computed from the

178 ensemble model with disordered H12 fragments yielded an excellent agreement to the

179 experimental curve ($\chi^2 = 0.77$), in accordance with the high propensity of RXR and

180 RAR H12 sequences to be disordered, in particular in their C-terminal part (Fig.

181 S1A). The other two alternative ensembles with ordered H12 regions displayed a

182 small but significant decrease in the agreement to the experimental curve, with χ^2 of

183 1.37 and 2.44 for the antagonist and agonist positions, respectively (Fig. S1). These

184 results indicate that in the unliganded form, H12 helices of both RXR and RAR are

185 disordered, and highlight the sensitivity of SAXS measurements and analysis to

186 minute structural changes in the heterodimer.

187 We have exploited this sensitivity to monitor the structural changes in

188 RXR/RAR induced by the binding of two selective RAR ligands, BMS493 (RAR

189 inverse agonist) and Am580 (RAR agonist). The presence of these two ligands

190 induces subtle but noticeable differences in the resulting curves (Fig. 1, Fig. S1 and

191 Table S1). In the presence of BMS493, the observed R_g, 26.5 ± 0.3 Å, is similar to the

192 one measured for the unliganded RXR/RAR heterodimer. Conversely, the presence of

193 Am580 induces a compaction of the particle, with a R_g of 25.6 ± 0.2 Å. We have used

194 the ensemble models based on the available X-ray structures to understand the

195 structural bases of these differences. Concretely, two ensembles of the heterodimer

8

196 were built in which RXR H12 was maintained disordered whereas RAR H12 was

197 assumed disordered or placed in agonist position (see methods section). The average

198 curves from these ensembles were linearly combined with that of the disordered H12

199 RAR to optimally describe the measured SAXS curves. SAXS curve of RXR/RAR in

200 the presence of BMS493 is nicely described ($\chi^2 = 0.87$) with models consisting in a

201 major contribution (85%) of a fully disordered RAR H12 helix (Fig. 1). Conversely,

202 in the presence of Am580, SAXS data are in agreement with the 100% of RAR H12

203 folded in the agonist position ($\chi^2 = 0.88$) (Fig. 1).

204 Similarly, we have analyzed the conformational changes of RAR H12 helix in

205 the RARI396E point mutant in the context of the heterodimer (RXR/RARI396E).

206 This point mutation, which was designed based on the crystal structure of the

207 complex between RARα and a corepressor peptide, is expected to destabilize the

208 RAR β-strand S3 conformation and favor a helical conformation in H11, thus

209 mimicking agonist-induced conformational change (le Maire et al., 2010).

210 Interestingly, the SAXS analysis indicates that the main heterodimeric species of this

211 mutant has the RAR H12 in a compact agonistic disposition although a certain

212 population (~20%) of disordered H12 was also observed ($\chi^2 = 1.11$, Fig. 1D). Not

213 surprisingly, the *P(r)* functions of the mutant RXR/RARI396E and Am580-bound

214 heterodimer are similar, whereas the unliganded RXR/RAR resembles more the

215 BMS493-bound RXR/RAR (Fig. 1A).

216 In summary, our structural analysis shows that H12 helices of unliganded

217 RXR/RAR heterodimer are mainly disordered in solution, and that the flexibility of

218 RAR H12 can be modulated by the presence of specific ligands or mutation. These

219 observations substantiate the pivotal role of H12 region as a modulator of RAR

220 activity.

9

221 **N-CoR$_{NID}$ is a disordered protein with evolutionarily conserved and partially**

222 **structured elements.** We produced and characterized a large fragment of the nuclear

223 receptor corepressor N-CoR (Hörlein et al., 1995) spanning from residue Gln2059 to

224 Glu2325. This fragment, N-CoR$_{NID}$ from now on, corresponds to the nuclear

225 interaction domain (NID) of N-CoR and encompasses the two nuclear receptor

226 binding motifs involved in the interaction of the protein with RXR/RAR heterodimer,

227 CoRNR1 (from 2065 to 2088) and CoRNR2 (from 2269 to 2291) (Fig. 2). The

228 biophysical characterization of N-CoR$_{NID}$ unambiguously indicates that the protein

229 behaves as an Intrinsically Disordered Protein (IDP) (Table S3, Fig. S2 and Fig. 2).

230 Concretely, N-CoR$_{NID}$ displays a reduced ^{1}H dispersion in NMR spectra (Fig. 2A),

231 and the Kratky plot does not present a clear maximum (Fig. 2B). Moreover, the s_{20w}

232 and frictional ratio measured by SV-AUC, 2.37 ± 0.19 S and 1.53 ± 0.13, are not

233 compatible with a globular protein of this size (≈29 kDa) (Fig. S2). Interestingly, far-

234 UV Circular Dichroism (CD) measured on N-CoR$_{NID}$ presents features that suggest

235 the presence of helical regions (Fig. 2C). Secondary structure is manifested by a shift

236 in the negative maximum at 205 nm, rather than at 198 nm for a pure random coil

237 profile, the negative shoulder near 220 nm, which is more pronounced than that

238 observed for fully disordered proteins, and the positive signal at 190 nm (Bienkiewicz

239 et al., 2002).

240 We performed the NMR study of N-CoR$_{NID}$ to identify structural features at

241 the residue level (Fig. 2A). The resonances of backbone nuclei of N-CoR$_{NID}$ were

242 assigned using standard triple resonance spectra at high spectrometer field. Out of the

243 241 expected ^{1}H-^{15}N HSQC backbone correlation peaks, only 183 could be

244 unambiguously assigned (Fig. S3). The non-identified correlations were either non-

245 visible peaks or had extremely low intensities precluding assignment. When mapping

246   the missing peaks on the amino acid sequence of N-CoR$_{NID}$, they clustered in three

247   non-consecutive regions of the protein (Fig. 2E and Fig. S3B). Interestingly, two of

248   these clusters were centered in the consensus NR binding domains CoRNR1 and

249   CoRNR2 and extended towards both flanking regions (Fig. 2E). Furthermore, a third

250   region, spanning from 2204 to 2234, also displayed absence or systematic decrease of

251   NMR intensities. Within this third region, which will be named Intermediate Region

252   (IR) from now on, no correlation peaks could be assigned for the residues 2204-2217

253   and 2222-2223. We performed ^{1}H-^{15}N HSQC experiments at different temperatures

254   (283, 288 and 293 K) but the number of peaks in the spectra did not change. Based on

255   these observations, we attributed the absence or decrease of NMR intensities in these

256   three regions to the formation of transient secondary structural elements that

257   experience chemical exchange processes in the *μs-ms* time-scale inducing severe

258   broadening of the signals. The presence of partially structured regions was

259   substantiated by the analysis of N-CoR$_{NID}$ sequence using several disorder prediction

260   servers (Fig. 2F). All predictors applied coincided in identifying CoRNR1 and

261   CoRNR2 as helices and their respective flanking regions as partially structured, both

262   motifs named as ID1 and ID2, respectively, from now on. Interestingly, disorder

263   predictors also identify the IR as partially structured, although to a lesser extent

264   (Buchan et al., 2013).

265   A sequence conservation bioinformatics analysis of N-CoR$_{NID}$ fragments from

266   multiple eukaryotic organisms indicates that ID1 and the large C-terminal region

267   (LCR, 2190 to 2295), which encompasses IR and ID2, are evolutionarily conserved

268   (Fig. 2G). On the contrary, the N-terminal region of the N-CoR$_{NID}$, with the exception

269   of ID1, is poorly conserved as typically observed in IDPs (Ota and Fukuchi, 2017).

270   Interestingly, the LCR presents a sequence composition that is closer to globular

11

271 proteins according to the charge hydropathy plot (Uversky and Gillespie, 2000), a

272 behaviour that is different from the N-terminus (Fig. 2D).

273 **N-CoR$_{NID}$ presents intramolecular transient contacts between conserved, co-**

274 **evolved and partially structured IR and ID2 regions.** The presence of local

275 compaction and long-range contacts in the partially structured and conserved C-

276 terminal region of N-CoR$_{NID}$ was explored using SAXS, Paramagnetic Relaxation

277 Enhancement (PRE) NMR experiments, which report on distance-depending induced

278 relaxation on NMR active nuclei, and Molecular Dynamics simulations (MD). In

279 addition to the native Cys2074, which sits in the middle of the ID1 motif, Ser2213

280 (preceding the IR) and Ser2288 (succeeding ID2) were mutated to cysteines. Note that

281 for these two mutants, the native Cys2074 was mutated into serine (C2074S) to have

282 only one cysteine at a time in the entire sequence of N-CoR$_{NID}$. After incorporating a

283 PROXYL stable radical on each of the cysteine residues independently, PRE-ratios

284 were measured for the three samples. When the paramagnetic moiety was introduced

285 in native Cys2074, no distal effects were observed, indicating that the N-terminal

286 region of N-CoR$_{NID}$ does not present long-range interactions with the rest of the

287 protein (Fig. 3A). Conversely, PRE data measured in the two point mutants, S2213C

288 and S2288C, provide a different picture of N-CoR$_{NID}$. For both single-cysteine

289 variants, a substantial decrease in intensity is observed for ^{1}H-^{15}N HSQC peaks from

290 the region between the IR and ID2 indicating the presence of extensive long-range

291 contacts (Fig. 3A). Interestingly, PREs measured in this region display a bell-shape

292 with stronger PRE effects in the proximity of both partially structured regions.

293 Moreover, PRE profiles for both mutants in the LCR are very similar suggesting a

294 direct interaction between the IR and the ID2 motifs that also affects the connecting

295 region (Fig. 3A). The compactness of N-CoR$_{NID}$ observed by PRE was substantiated

296 by SAXS. A simple random coil model built with Flexible-Meccano (Bernado et al.,

297 2005) could not reproduce the experimental SAXS curve (Fig. 3B). Indeed, the

298 theoretical ensemble turned out to be more extended than N-CoR$_{NID}$ in solution, with

299 R_g of 50.4 Å and 47.2 ± 1.2 Å for the theoretical and experimental SAXS curves,

300 respectively.

301 PRE and SAXS data were used to further characterize the structural

302 compaction observed in N-CoR$_{NID}$. Large ensembles of conformations were built with

303 Flexible-Meccano to which multiple explicit dispositions of the PROXYL moiety

304 were attached to the native or engineered cysteine residues. Subsequently, the

305 theoretical averaged PRE ratios were computed as previously described (Salmon et

306 al., 2010), and compared with the three experimental PRE profiles (see methods for

307 details). Not surprisingly, the random coil model only presented contacts in the

308 vicinity of the paramagnetic sites, and therefore did not reproduce the experimental

309 profiles especially in the C-terminal region (Fig. 3A, blue lines). In order to interpret

310 the long-range contacts observed in the LCR, a structurally biased model was built

311 based on previous experimental and bioinformatics observations. Concretely,

312 conformations with at least two contacts of ≤ 15 Å between residues from distal

313 regions of the LCR presenting partial structuration (average disorder below 0.5, Fig.

314 2F), evolutionary conservation (Bits ≥ 1.6, Fig. 2G), or co-evolution (Fig. 3D), were

315 selected from a large ensemble of random coil conformations. This filtered ensemble

316 resulted too compact, and it was further refined using the SAXS curve to yield a sub-

317 ensemble compatible with the scattering profile (Fig. 3B, C), which was subsequently

318 used to compute the PRE values for the three PROXYL-tagged N-CoR$_{NID}$ constructs

319 and compared with the experimental ones. The resulting theoretical PRE profiles

320 displayed an excellent agreement with the experimental ones (Fig. 3A, red lines). Not

13

321 surprisingly, a systematic decrease of the PRE values in the LCR compared to the N-

322 terminal region was observed. More interestingly, PRE fluctuations in the connecting

323 region between the IR and the ID2 region were nicely reproduced for the S2213C and

324 S2288C mutants.

325   In order to further investigate the formation of these long-range contacts, we

326 performed MD simulations of N-CoR$_{NID}$ based on a coarse-grained model that was

327 specifically designed for IDPs (Smith et al., 2014) and takes into account sequence-

328 specific electrostatic and hydrophobic interactions. The resulting conformational

329 ensemble, which was not biased by experimental data, revealed that N-CoR$_{NID}$

330 sequence is poised to form transient yet noticeable interactions in the C-terminal

331 region. Importantly, the contact matrix derived from coarse-grained MD is in good

332 agreement with the PRE-derived ensemble (Fig. 3E). These results substantiate the

333 presence of transient intramolecular contacts between distal regions within the C-

334 terminal region of N-CoR$_{NID}$. Moreover, these regions present relatively stable

335 secondary structural elements and a high level of evolutionary conservation.

336 **RXR/RAR heterodimer interacts with the corepressor mainly through RAR and**

337 **in a cooperative manner.** After characterizing the individual partners, we

338 investigated the complex between RXR/RAR and N-CoR$_{NID}$. The affinities of

339 RXR/RAR heterodimer for NCoR peptides encompassing either CoRNR1 or

340 CoRNR2 motifs and for the N-CoR$_{NID}$ fragment were measured by fluorescence

341 anisotropy and thermophoresis, respectively. As shown in Figure 4, formation of the

342 RXR/RAR heterodimer does not modify the binding capacity of individual RXR and

343 RAR monomers for isolated CoRNR peptides. In fact, the RXR/RAR heterodimer

344 binds the two isolated CoRNR peptides with nearly the same affinity, 1.68 and 1.47

345 μM for CoRNR1 and CoRNR2, respectively (Fig. 4B). These affinity values are very

14

346 similar to those measured previously (le Maire et al., 2010) for the unliganded RAR

347 monomer (1.40 and 1.55 μM for peptides encompassing CoRNR1 and CoRNR2,

348 respectively) (Fig. 4A). The slight preference observed for CoRNR1 is induced by the

349 ability of this motif to form a β-sheet interface with S3 of RARα, which is further

350 stabilized in the presence of the inverse agonist BMS493 with an affinity of 0.17 μM.

351 Conversely, in its monomeric form, RXR presents a moderate affinity (22.5 μM) for

352 CoRNR2, and no measurable interaction with CoRNR1 (Fig. 4A). As expected, the

353 presence of the RAR-selective agonist Am580 or the RARI396E mutation cause a

354 noticeable decrease in the affinity of RXR/RAR for both N-CoR peptides with a

355 stronger effect on CoRNR1 (Fig. 4B). Conversely, the RAR inverse agonist BMS493

356 efficiently increases the binding affinity of RXR/RAR for CoRNR1, but has not much

357 effect on the interaction with CoRNR2 (Fig. 4A-B). These observations suggest a

358 sequential and directional mechanism in which, in the context of N-CoR$_{NID}$,

359 comprising the two CoRNR motifs, the main anchoring point would involve RAR on

360 the heterodimer side and CoRNR1 on the N-CoR side. This primary contact would

361 then enable a second lower-affinity interaction between RXR and CoRNR2.

362 Microscale thermophoresis measurements further substantiated the

363 cooperativity and directionality of the interaction. The affinity of RXR/RAR

364 heterodimer for N-CoR$_{NID}$ was found to be much higher than for individual peptides

365 with a value of 0.21 ± 0.09 μM, to be compared to 1.68 and 1.47 μM for CoRNR1

366 and CoRNR2, respectively (Fig. 4B-C). The inverse agonist BMS493 or deletion of

367 RXR helix H12 (RXRΔH12), which are known to enhance the interaction of RAR

368 with CoRNR1 (Germain et al., 2009; le Maire et al., 2010) (Fig. 3A) and that of RXR

369 for CoRNR2 (Hu and Lazar, 1999), respectively, were shown to increase significantly

370 the overall affinity of the heterodimer for N-CoR$_{NID}$ (Fig. 4C). In contrast, the RAR

15

371 agonist Am580, or the RAR mutation I396E induced a strong decrease of the

372 affinities with Kd values of 8.6 ± 3.70 μM or 4.42 ± 1.86 μM, respectively, due to the

373 weakening of the interaction of RAR with CoRNR1 in both cases. The observation

374 that the perturbation of individual anchoring points has severe consequences on the

375 affinity of RXR/RAR for N-CoR$_{NID}$ demonstrates the cooperativity and directionality

376 of the complex.

377 **N-CoR$_{NID}$ forms a transient multi-site complex with RXR/RAR.** Affinity

378 measurements indicate that N-CoR$_{NID}$ binds cooperatively to RXR/RAR heterodimer.

379 To unveil the structural bases of this cooperativity, we combined SAXS and NMR to

380 study the N-CoR$_{NID}$:RXR/RAR heterotrimeric complex.

381 SAXS data indicates that the complex is monodisperse with a R_g and D_{max}

382 values of 48.4 ± 1.1 Å and 194 ± 10 Å, and with a stoichiometry of 1:1:1 according to

383 SV-AUC and mass spectrometry data (Table S3, Fig. S2 and Fig. S4). Moreover, the

384 Kratky representation with elevated baseline at high s, and asymmetric P(r) function

385 suggest that N-CoR$_{NID}$:RXR/RAR has a significant degree of flexibility (Fig. S5).

386 The observation of cooperative effects prompted towards the possibility that

387 the two binding regions of N-CoR$_{NID}$, CoRNR1 and CoRNR2, could simultaneously

388 interact with RXR/RAR forming the so-called deck model (le Maire and Bourguet,

389 2014). The latter scenario is in contrast with the asymmetric model, where a single

390 corepressor binding side would interact with the RXR/RAR heterodimer as observed

391 for the complex between RXR/RAR and a fragment of the coactivator Med1 (Rochel

392 et al., 2011). Ensembles of conformations were built for both the deck and the

393 asymmetric binding modes (see SI Appendix for details), integrating previous

394 knowledge of the system. This includes the directionality of the complex derived from

395 affinity measurements, and two crystallographic structures of RAR and RXR with

16

396 corepressor peptides that were used to define the atomic details of the interacting sites

397 (Table S2) (le Maire et al., 2010; Zhang et al., 2011). H12 helices from both RXR and

398 RAR were considered disordered based on our SAXS analysis (see above). Finally,

399 the non-interacting regions of N-CoR$_{NID}$ were assumed to adopt random coil

400 conformations based on the NMR data (see below). Theoretical SAXS curves were

401 computed from large ensembles of structures for both interacting modes yielding

402 notably different profiles (Fig. S6). These curves were optimally combined to derive

403 the relative population of both interacting modes by minimizing the agreement to the

404 experimental curve. An excellent agreement ($\chi_i^2 = 0.88$) to the experimental SAXS

405 curve of N-CoR$_{NID}$:RXR/RAR was obtained when a 85/15 ratio for the

406 asymmetric/deck models was used (Fig. 5A, B). This result indicates that both

407 binding modes coexist in solution although the asymmetric model is the major species

408 in the absence of ligand. The similarity of the scattering curves between the two

409 alternative asymmetric models (N-CoR$_{NID}$ bound to RAR through ID1 or N-CoR$_{NID}$

410 bound to RXR through ID2) precludes the identification of the major asymmetric

411 species.

412 In order to understand the interaction at the residue level, we performed NMR

413 experiments by adding an equimolar amount of RXR/RAR to a ^{15}N-isotopically

414 labeled N-CoR$_{NID}$ sample. The NMR ^{1}H-^{15}N HSQC spectrum of N-CoR$_{NID}$ presents

415 correlation peaks that are superimposable to these obtained in the absence of

416 heterodimer, and no correlation peaks shifted upon the addition of the heterodimer

417 (Fig. S7). This observation indicates that the conformational properties of N-CoR$_{NID}$

418 are equivalent in the free and in the bound states, and that N-CoR$_{NID}$ remains mostly

419 disordered when bound to the RXR/RAR heterodimer. However, in the presence of

420 RXR/RAR, a systematic decrease in peak intensities is observed for N-CoR$_{NID}$ when

17

421   compared to the free form (Fig. S7 and Fig. 2E). We attribute this observation to the

422   perturbation of the hydrodynamic properties of N-CoR$_{NID}$ upon binding to the

423   heterodimer that senses the presence of a large globular particle and increases the

424   apparent correlation time. Interestingly, the intensity decrease is not homogeneously

425   distributed along the protein, and two regions of the protein can be clearly

426   distinguished. The N-terminal region connecting ID1 and the IR presents moderate

427   intensity decreases that are slightly larger in the proximity of both partially structured

428   regions. Conversely, the region connecting the IR with ID2 displays a more important

429   intensity reduction. The transient binding of ID2 to RXR as demonstrated by SAXS

430   and the change in the dynamic regime for the intra-molecular interactions could

431   explain the enhanced intensity reduction in the C-terminus. To explore this

432   phenomenon, we performed PRE experiments for the complexes formed by the

433   previously described N-CoR$_{NID}$ cysteine mutants and the heterodimer RXR/RAR. The

434   incorporation of radical moieties on the N-CoR$_{NID}$ S2213C and S2288C point mutants

435   induced a sequence-dependent decrease on the PRE ratios of residues placed at their

436   flanking regions, a phenomenon which is typically observed in IDPs (Fig. 5D, E).

437   Interestingly, for both mutants a small but systematic increase in the PRE ratios was

438   found for the N-terminal residues when compared with the PRE data of the free forms

439   (Fig. 3A). This is most probably caused by the reduced conformational exploration in

440   N-CoR$_{NID}$ when bound to the heterodimer, which limits sporadic contacts of the C-

441   terminal region with the N-terminus. Compared with the free form of N-CoR$_{NID}$, the

442   presence of the heterodimer causes a PRE-ratio increase in the region connecting the

443   IR and ID2, and partially suppresses the above-described bell-shape of PRE ratios.

444   This increase is not homogeneous and the region around 2250 presents strong PRE

445   effects. These observations indicate that the population of conformations experiencing

18

446 long-range contacts in the C-terminus is diminished in the bound form. This last

447 observation is coherent with the existence of a minor population of the deck model

448 that conformationally restricts N-CoR$_{NID}$, and partially hampers intra-molecular

449 contacts involving ID2.

450 **Cognate ligands and mutations modulate the conformational equilibrium in the**

451 **complex.** Thermophoresis experiments demonstrated that the affinity of N-CoR$_{NID}$ for

452 RXR/RAR can be finely tuned by the addition of cognate ligands or by mutations at

453 the recognition sites of both RXR and RAR. We applied SAXS and NMR to explore

454 the structural bases of this affinity modulation.

455   The addition of the RAR inverse agonist BMS493 reinforced the interaction of

456 RXR/RAR with N-CoR$_{NID}$ through CoRNR1 (Fig. 4B and 4C). The analysis of the

457 SAXS curve of the ternary complex in the presence of BMS493 indicates that the

458 overall size of the particle is slightly reduced with respect to the unliganded form of

459 the complex, R_g = 47.5 ± 1.0Å (Table S3). The characterization of that curve in terms

460 of atomistic ensemble models indicated that the asymmetric model is still the major

461 species but the population of the doubly-bound deck model increases up to 35% (Fig.

462 6A and 6B). A more extreme situation was observed when the complex was formed

463 with RXRΔH12/RAR that strongly reinforces the overall affinity with N-CoR$_{NID}$ by

464 increasing the interaction of RXR with ID2 (Fig. 4B and 4C). This RXR H12 deletion

465 renders the RXR hydrophobic groove more accessible and has been reported to

466 significantly increase the interaction of RXR with corepressor(Hu and Lazar, 1999;

467 Schulman et al., 1996; Zhang et al., 2011). In that situation, the SAXS curve, which

468 presented a smaller R_g of 42.0 ± 2 Å, could be explained with the only presence of the

469 deck model (Table S3, Fig. S5 and Fig. 6A and 6B). The ^{15}N N-CoR$_{NID}$ NMR

470 intensities measured for both complexes, upon addition of RXR/RAR in the presence

471 of BMS493 and of RXRΔH12/RAR, displayed a systematic decrease with respect to

472 the unliganded N-CoR$_{NID}$:RXR/RAR (Fig. 6C). We attributed this observation to the

473 increase of the overall correlation time sensed by N-CoR$_{NID}$ when the population of

474 the deck model increases. This is especially significant in the region close to ID2 that

475 fits well with the higher interaction of ID2 with RXRΔH12. Compared with the

476 unliganded (Fig. S7), the region connecting the IR and ID2 in the N-

477 CoR$_{NID}$:RXRΔH12/RAR displays a less prominent decrease of intensities, indicating

478 that some of the structural and/or dynamical properties that cause the decrease in the

479 native complex are abolished by the formation of a doubly-bound state.

480   We also performed SAXS experiments of the complex N-CoR$_{NID}$:RXR/RAR

481 in the presence of the RAR agonist Am580, and of the complex N-

482 CoR$_{NID}$:RXR/RARI396E. Affinity measurements indicated that both conditions

483 diminished the interaction of RXR/RAR with N-CoR$_{NID}$ through CoRNR1 (Fig. 4B

484 and 4C). The analysis of the SAXS data using explicit models indicated that the

485 decrease of the interaction has severe structural consequences for both complexes. On

486 one hand, the SAXS curve measured on N-CoR$_{NID}$:RXR/RAR in the presence of

487 Am580 can only be described ($\chi^2 = 0.81$) if large populations, 51%, of unbound N-

488 CoR$_{NID}$ and RXR/RAR are invoked (Fig. 6A and 6B), showing that Am580 partially

489 breaks the complex by diminishing the interaction of CoRNR1 with RAR. Similarly,

490 the SAXS curve measured for the complex N-CoR$_{NID}$:RXR/RARI396E is optimally

491 described using a combination of unbound species, asymmetric complex and a low

492 percentage of deck complex. ^{15}N N-CoR$_{NID}$ NMR intensities measured in these two

493 conditions were also coherent with SAXS observations (Fig. 6C). The general

494 increase in these intensities is in line with the presence of a weaker complex with

495 lower dragging forces and smaller apparent correlation time.

20

Mammalian two-hybrid experiments confirm the interaction in cellular context.

496

497 Mammalian two-hybrid experiments were performed to validate in a cellular context

498 the cooperativity and the directionality of the interaction between N-CoR$_{NID}$ and

499 RXR/RAR heterodimer, deduced from the measurements of *in vitro* binding constants

500 and from the SAXS/NMR modelling. Two-hybrid analyses were performed in COS

501 cells with chimeras containing GAL4 DNA-binding domain fused to a N-CoR

502 fragment encompassing ID1 and ID2 (from 1629 to 2453 termed Gal-NCoR) or either

503 CoRNR1 or CoRNR2 of N-CoR (Gal-CoRNR1 and Gal-CoRNR2, respectively), and

504 the LBD of RXRα or of RARα fused to the activation domain of VP16 (termed VP16-

505 RXR or VP16-RAR, respectively).

506 First, we confirmed that wt RXR/RAR heterodimer interact with both

507 CoRNR1 and CoRNR2 (Fig. 7A), with a slightly better effectiveness for CoRNR1.

508 However, the interaction of RXR/RAR heterodimer with the longer N-CoR fragment

509 (equivalent to N-CoR$_{NID}$) is significantly stronger than the individual CoRNR regions.

510 Therefore, in a cellular environment the two CoRNR motifs in N-CoR bind to the

511 RXR/RAR heterodimer in a cooperative manner and may induce a stronger affinity.

512 In addition, the selective RAR agonist TTNPB is able to release CoRNR1 and the N-

513 CoR fragment from RXR/RAR heterodimer (Fig. 7A). Importantly, a significant

514 interaction of CoRNR2 is measured even in the presence of TTNPB which only

515 targets RAR likely reflecting the interaction of RXR with the CoRNR2 motif (Fig.

516 7A). The directionality of the interaction between N-CoR$_{NID}$ and RXR/RAR

517 heterodimer was further confirmed with two-hybrid experiment with the N-CoR

518 fragment (Fig. 7B). Whereas RXR alone interacts very weakly with N-CoR, in

519 agreement with previous observations of our group (Germain et al., 2002) and with

520 our fluorescence anisotropy data on RXR monomer and CoRNR peptides (Fig. 4),

21

521 addition of RARα allows to measure a significant interaction between the RXR/RAR

522 heterodimer and Gal-NCoR. Interestingly, the addition of the RAR-selective agonist

523 TTNPB was sufficient to decrease this interaction (Fig. 7B). Moreover, RAR alone

524 can efficiently recruit N-CoR (Fig. 7C). All these observations confirm that RAR is

525 indispensable for the corepressor recruitment by RXR/RAR heterodimer. On the other

526 side, the role of RXR for the recruitment of N-CoR was subsequently addressed in a

527 cellular environment (Fig. 7C). Relative to RAR alone, addition of RXR yielded a

528 slight but significant increase of the N-CoR interaction confirming an active role of

529 RXR in the heterodimeric form (Fig. 7C), and thus the cooperativity in the

530 interaction. This effect was more pronounced for the RXRΔH12 deletion mutant (Fig.

531 7C), in agreement with MST experiments and SAXS modelling. We reasoned that the

532 RXR H12 deletion may generate a new interaction surface for N-CoR. Contrary to the

533 wt heterodimer, TTNPB does not reduce N-CoR interaction with RXRΔH12/RAR

534 heterodimer (Fig. 7C). This observation indicates that RXRΔH12/RAR retains the

535 ability to interact with N-CoR, through the RXR hydrophobic groove. Two-hybrid

536 experiments with RXRΔH12/RAR heterodimer and Gal-CoRNR1 and Gal-CoRNR2

537 clearly demonstrated that CoRNR2 interacts with the more accessible groove of RXR

538 as TTNPB was unable to reduce CoRNR2 association, and that CoRNR1 interacts

539 with RAR as its binding is impaired by addition of TTNPB (Fig. 7D), confirming the

540 directionality in the interaction of N-CoR with RXRΔH12/RAR.

541

542 **Discussion**

543 In this study we have integrated multiple structural, biophysical and cell-

544 biology techniques to decipher the molecular bases of transcriptional repression of the

545 retinoic acid nuclear receptor by the corepressor N-CoR. Our results demonstrate that

22

546   the molecular mechanism relies on the interplay of the flexible elements found in

547   RXR/RAR heterodimer and the intrinsically disordered N-CoR.

548   We have exploited the high sensitivity of SAXS to probe minute structural

549   changes that perturbe the shape of the RXR/RAR LBDs heterodimer in the absence or

550   in the presence of various RAR ligands and point mutations. The main structural

551   changes of the heterodimer reside in the conformation and position of helices H12 of

552   each monomer. Our analyses demonstrate that helices H12 of RAR and RXR are

553   better described as disordered in the absence of ligand, in line with previous NMR

554   and amide hydrogen/deuterium (H/D) exchange experiments for a number of NRs

555   including RXR (Johnson et al., 2000; Kallenberger et al., 2003; Lu et al., 2006).

556   These results invalidate the initial hypothesis of allosteric inhibition of RXR by a

557   subset of nuclear receptors, such as RAR and THR. It was suggested that, in

558   unliganded RXR/RAR and RXR/THR heterodimers, H12 of RXR docks to the

559   coregulator-interaction site of the partner (Westin et al., 1998; Zhang et al., 1999).

560   The lack of structural order of RAR helix H12 was also deduced from SAXS

561   data measured on RXR/RAR heterodimer in the presence of RAR inverse agonist

562   (BMS493), in agreement with the absence of density for RAR H12 in the crystal

563   structure of the complex formed by RARα-LBD, the corepressor N-CoRNR1 and the

564   inverse agonist (le Maire et al., 2010). These two forms of RXR/RAR heterodimers,

565   unliganded and in complex with the inverse agonist, bind corepressors, confirming

566   that the disordered helices H12 are not directly involved in corepressor binding.

567   Nevertheless, the disordered and highly flexible RXR and RAR helices H12 may

568   sterically screen the interaction with corepressors. Indeed, RXRΔH12 and RARΔH12

569   deletion mutants recruit corepressors more efficiently than wt RXR and RAR (le

570   Maire et al., 2010; Zhang et al., 1999), probably due to a higher accessibility of the

23

571 LBD interaction region for the corepressors. Conversely, addition of RAR agonist

572 (Am580) induces a compaction of the heterodimer that fits with RAR helix H12

573 adopting the agonist position, as observed in crystal and solution structures of RAR in

574 the presence of agonists (Bourguet et al., 2000; Egea et al., 2001; Klaholz et al., 2000;

575 le Maire et al., 2010). Our SAXS analysis of the mutant RARI396E, for which no

576 structural data had been reported before, and whose mutation was suggested to induce

577 the S3 strand to H11 helix transition (le Maire et al., 2010), also provokes a main

578 rearrangement of RAR helix H12 that is compatible with an agonist conformation.

579 The transition from a disordered state to an ordered and active conformation of RAR

580 helix H12 is induced by the binding of an agonist that efficiently triggers corepressor

581 release. Importantly, this release is dependent of the proportion of agonist

582 conformation of this helix. With this analysis we confirm that the disorder to order

583 transition of helix H12 conformation is a key mechanism of corepressor release and

584 coactivator recruitment by RXR/RAR, contrary to the region corresponding to

585 H11/S3 in RAR which is the master regulator of corepressor association to RAR (le

586 Maire et al., 2010).

587 Our extensive experimental and computational analysis of N-CoR$_{NID}$

588 demonstrates that it is a disordered protein presenting local and long-range structural

589 phenomena that are directly linked to its function. With the exception of the CoRNR1

590 motif, the N-terminal region presents the prototypical spectroscopic features of a

591 random coil. Moreover, bioinformatics analyses on this region indicate a poor

592 evolutionary conservation and the absence of co-evolutionary interactions with other

593 parts of the protein has often been observed in disordered regions without direct

594 functional roles (Ota and Fukuchi, 2017). Conversely, the C-terminal region including

595 the CoRNR2 motif presents multiple structural features and is evolutionary

24

596　conserved. In solution, CoRNR1 and CoRNR2 are preformed molecular recognition

597　elements (MOREs) (Mohan et al., 2006) that mediate the interaction with the

598　heterodimer. The lack of NMR information for these regions is probably due to

599　intermediate exchange regime processes linked to the transient formation of

600　secondary structures. Very interestingly, NMR has unveiled a new region, IR, placed

601　between the two NR interaction domains that is partially ordered and highly

602　conserved in eukaryotes. These two observations suggest a relevant functional role for

603　IR. Indeed, N-CoR$_{NID}$ experiences transient long-range tertiary interactions between

604　the IR and ID2 that we have identified and structurally characterized using PREs. The

605　accurate description of experimental PREs was achieved when the long-range

606　interaction between co-evolved residues in partially structured regions of the C-

607　terminal region of N-CoR$_{NID}$ was used. This long range contact is partially impaired

608　when N-CoR interacts with NRs as demonstrated by our NMR analysis of N-CoR$_{NID}$

609　in complex with the RXR/RAR heterodimer. The formation of the complex would

610　liberate the IR that would become available to interact with other proteins of the

611　repressive macromolecular complex.

612　　　A number of structural and biophysical studies have already revealed the

613　complex interactions between NR heterodimers and coactivators (Chandra et al.,

614　2017; Osz et al., 2012; Pavlin et al., 2014; Pogenberg et al., 2005; Rochel et al., 2011;

615　de Vera et al., 2017; Zheng et al., 2017). However, the interaction of NRs with

616　corepressors that hampers gene transcription in the basal state are poorly understood.

617　Our study, which included solution-state structural methods along with biochemical,

618　biophysical and computational approches, reveals for the first time the atomistic

619　details in terms of ensembles of the interaction between RXR/RAR heterodimer and

620　N-CoR. Our study shows a cooperative interaction of both CoRNR1 and CoRNR2

621 motifs with RXR/RAR as the binding affinity of N-CoR$_{NID}$ is stronger than the

622 individual interactions. In addition, the interaction has a defined directionality: N-

623 CoR$_{NID}$ is recruited primarily to RAR through CoRNR1 enabling CoRNR2 to

624 subsequently bind to RXR for which it has a moderate affinity. Importantly, the other

625 configuration would not produce a cooperative binding as CoRNR1 does not interact

626 with RXR.

627   Taken toghether, our data suggest the mechanistic model that is depicted in

628 Figure 8. In the absence of ligand, it exists an equilibrium between a major population

629 of assymetric binding of N-CoR$_{NID}$ to RXR/RAR and a minor population of doubly

630 bound N-CoR$_{NID}$ (deck binding mode) in which both CoRNR boxes simultaneously

631 interact with the heterodimer, accounting for the cooperativity of the interaction. The

632 major species of the assymetric binding mode can be reasonably assigned to N-

633 CoR$_{NID}$ binding to RAR through the CoRNR1 motif, which presents the strongest

634 local affinity. Our observations through mammalian two hybrid experiments

635 substantiate with this model, also suggesting that the structural phenomena probed *in*

636 *vitro* also occurs in cells. The addition of RAR ligands or point mutations in either

637 RXR and RAR modify the strenght of the individual interactions and, consequently

638 change the cooperativity and the equilibrium between the assymetric and the deck

639 binding modes. Concretely, the addition of a RAR inverse agonist, which strenghtens

640 the interaction of CoRNR1 with RAR, or the use of a heterodimer with the truncated

641 RXR H12 (RXRΔH12), which increases the interaction of CoRNR2 with RXR,

642 produce a notable increase of the overall affinity of the complex. This increase is a

643 consequence of an enhancement of the cooperativity by the equilibrium displacement

644 towards the deck model as observed in the SAXS and NMR analyses. Conversely,

645 when weakening RAR binding site using a RAR agonist or the RARI396E mutant,

646 there is a descrease in the overall affinity of RXR/RAR for N-CoR$_{NID}$, as a

647 consequence of the reduction of the population of the deck complex in favour of the

648 assymetric binding mode that eventually dissociates. Therefore, the equilibrium

649 between the two interaction modes has relevant consequences for gene transcription,

650 the assymetric mode facilitating coregulator swapping from corepressor to

651 coactivator. Thus, despite the high overall affinity of the corepressor complex in the

652 deck binding mode, the low affinity of the individual CoRNR motifs for the LBDs

653 permits effective coregulator binding versatility upon environmental perturbations.

654 The multivalency of coregulator proteins has important consequences in the

655 thermodynamics of the interaction and the kinetics of the transition between repressed

656 and active states. Multivalency is specially relevant for NR regulation as coactivators

657 and corepressors have been described to contain several LBD interaction boxes with

658 slightly different individual affinities. As a consequence, the number of coexisting

659 assembly states increases dramatically, complexifying the regulation mechanism. In

660 this context, the integrative approach applied in the present study will offer the

661 opportunity to disentangle this complexity, structurally characterize the individual

662 states and address the thermodynamics of NR gene transcription regulation.

663

681

682 Author contributions

683 T.C., W.B., P.G., N.S., A.lM. and Pa.B. designed research. T.C., Ph.B., F.A., V.V.,

684 C.E., P.G., N.S., Al.B. R.B., and A.lM. performed experiments. Ab.B. provided

685 reagents. T.C., W.B., N.S., A.lM. and Pa.B. analyzed data and wrote the manuscript.

686

687 The authors declare no conflict of interest.

688

689 Main figure titles and legends

690 **Figure 1: Modulation of RAR H12 disorder.** (A) Pairwise distance distribution

691 functions, *P(r),* computed from experimental SAXS curves of wt RXR/RAR

692 heterodimer (blue), in the presence of the ligands BMS493 (orange) and Am580 (light

693 green), and for the RXR/RARI396E mutant (dark green). The inset is a heat-map

694 showing their pairwise Kolmogorov-Smirnov-based similarities (KS). The scale is

695 relative from low (red) to high (blue) KS values. (B) Ensemble models (5

28

696   conformations) of RXR/RAR with RXR H12 disordered (residues 441-467) and

697   either RAR H12 in disordered (left, residues 394-421) or agonistic (right, residues

698   408-414) conformation. (C) SAXS intensity profiles (gray dots) in logarithmic scale

699   as a function of the momentum transfer $s = (4\pi \sin \theta)/\lambda$, where 2θ is the scattering

700   angle, compared with the optimal combination of conformational ensembles for the

701   RXR/RAR (blue), upon the addition of BMS493 (orange), Am580 (light green), and

702   for the RXR/RARI396E (dark green). Bottom, point-by-point residuals of the fitting

703   with the same color code. (D) Evolution of χ^2 as a function of the relative contribution

704   of the ensembles corresponding to RXR/RAR with disordered H12 and RAR H12 in

705   an agonistic conformation. Well-defined minima are observed for all scenarios.

706   **Figure 2**: **Structural characterization of N-CoR$_{NID}$.** (A) ^{1}H-^{15}N HSQC of N-

707   CoR$_{NID}$ displaying the reduced backbone amide ^{1}H frequency dispersion typical from

708   IDPs. (B) Kratky representation of the SAXS profile measured for N-CoR$_{NID}$

709   indicating its lack of compactness. (C) Far UV Circular Dichroism of N-CoR$_{NID}$. (D)

710   Charge Hydropathy (Uverky's plot) of N-CoR$_{NID}$. Mean net charge versus mean

711   hydropathy plot for the set of 275 folded (grey circles) and 91 natively unfolded

712   proteins (white circles). The N-terminus of N-CoR$_{NID}$ (blue) is clearly in the cluster of

713   disordered proteins, whereas C-terminus of N-CoR$_{NID}$ (red) is on the of folded

714   proteins side. (E) Normalized ^{1}H-^{15}N HSQC's intensities of ^{15}N -labeled N-CoR$_{NID}$

715   along the primary sequence. The colored bands represent the CoRNR1 and CoRNR2

716   motifs in green and pink, respectively. The intermediate region (IR) is colored in grey.

717   Secondary structure elements of CoRNR boxes are plotted as a red cylinder and blue

718   arrow for α-helix and β-strand, respectively. The PSIPRED predicted secondary

719   structure elements are represented in black and dashed line for coil. (F) Average

720   disorder prediction (solid grey line) and its standard deviation computed using

721   different computational tools along the N-CoR$_{NID}$ primary sequence. (G) Sequence

722   conservation profile computed from a Hidden Markov Model multiple sequence

723   alignment using Skylign. The overall height indicates the conservation per position.

724   **Figure 3: Long-Range interactions in NCoR$_{NID}$.** (A) Paramagnetic Relaxation

725   Enhancement data of N-CoR$_{NID}$ measured on WT C2074 (top), S2213C (middle) and

726   S2288C (bottom) N-CoR$_{NID}$ mutants. Experimental values (black bars) calculated

727   from the ratio of intensity (PRE Ratio = I_{para}/I_{dia}) of ^{1}H-^{15}N HSQC spectra measured

728   on the paramagnetic (I_{para}) samples in the position indicated with a yellow star, and on

729   the diamagnetic samples (I_{dia}). Averaged theoretical PRE-ratio profiles computed

730   from 10,000 N-CoR$_{NID}$ structures modelled as fully random coil (solid blue line) or

731   assuming long-range interactions (solid red line) between predicted co-evolved,

732   conserved and structured residues within the C-terminal region (LRC). (B)

733   Experimental SAXS curve of N-CoR$_{NID}$ (gray open circles) overlapped with curves

734   calculated form the random coil pool (blue) and EOM selected sub-ensemble with

735   long-range contacts (red). The bottom panel shows the point-by-point residuals of the

736   fitting for both conformational ensembles. (C) Radius of gyration, R_g, distributions

737   obtained for the random coil ensemble (blue) and the EOM selected sub-ensemble

738   (red). (D) Chord-plot of co-evolved residues predicted using BISAnalyser where the

739   lines connect residues identified as co-evolving pairs. (E) Cα-Cα contact map with

740   respect to a random coil model in logarithmic scale computed from the PRE/SAXS

741   refined ensemble (top triangle) and from the coarse-grained MD trajectory (bottom

742   triangle). Color scale quantifies the intensity of the contact where red indicates and

743   average close contact and yellow indicates an average distance equivalent to the

744   random coil model. CoRNR1, IR, and CoRNR2 are displayed in the diagonal to

745   facilitate interpretation of the map.

30

746 **Figure 4: RXR/RAR heterodimer interacts with N-CoR$_{NID}$ mainly through RAR**.

747 (A) Affinities of CoRNR motifs of N-CoR for wt RAR, mutant RARI396E and RXR

748 measured by fluorescence anisotropy, in the absence of ligand or the presence of

749 Am580 (RAR agonist) or BMS493 (RAR inverse agonist). Data on RAR and

750 RARI396E were previously published in [23] and are added here for clarity. (B)

751 Affinities of CoRNR motifs of N-CoR for wt RXR/RAR and mutant RXR/RARI396E

752 heterodimers measured by fluorescence anisotropy in same ligation states as in A. (C)

753 Affinities of N-CoR$_{NID}$ fragment for wt RXR/RAR (in same ligation states as in A),

754 mutant RXR/RARI396E and RXRΔH12/RAR heterodimers measured by Microscale

755 Thermophoresis. Affinity measurements were done by triplicates.

756 **Figure 5**: **Structural bases of the cooperative binding of N-CoR$_{NID}$ with the**

757 **RXR/RAR heterodimer.** (A) Ensemble models of the asymmetric and deck forms of

758 the N-CoR$_{NID}$:RXR/RAR. Three conformations for each of the ensembles are shown.

759 (B) Optimization of the relative population of the asymmetric and deck forms of the

760 N-CoR$_{NID}$:RXR/RAR. χ^2 values (black dots) for different relative populations of the

761 two forms of the complex (C) Scattering intensity, $I(s)$, as a function of the

762 momentum transfer, s, for the SAXS curve measured for N-CoR$_{NID}$:RXR/RAR (grey

763 empty dots) and the theoretical ones for the deck model (pink line) and the

764 combination of the asymmetric and deck models (88:12) (blue line). Residuals are

765 displayed at the bottom of the figure to substantiate the quality of the agreement. (D)

766 ^{1}H-^{15}N HSQC peak intensity differences for N-CoR$_{NID}$ in its free state and in the

767 presence of an equimolar amount of RXR/RAR. (E) PRE values for S2213C (top) and

768 S2288C (bottom) N-CoR$_{NID}$ mutants (positions of the paramagnetic tags are indicated

769 with yellow stars) in the presence of the RXR/RAR heterodimer (black bars).

770 Theoretical PRE values obtained for a structurally biased model of free N-CoR$_{NID}$

771 (equivalent to Fig. 3A) are displayed to highlight the differences in the long-range

772 intramolecular interaction network between the free and bound states of the protein.

773 **Figure 6**: **Modeling of N-CoR$_{NID}$:RXR/RAR complexes**. (A) Logarithmic

774 representation of SAXS intensity, *I*, versus momentum transfer *s* (open circles) of N-

775 CoR$_{NID}$:RXR/RAR in the presence of inverse agonist BMS493 (orange) or agonist

776 Am580 (light green), and N-CoR$_{NID}$: RXRΔH12/RAR (red) and N-

777 CoR$_{NID}$:RXR/RARI39E (green) variants. Simulated scattering patterns (solid black

778 lines) are the best linear combination of theoretical scattering profiles of the species in

779 panel B weighted by their relative population. χ^2 values indicate the goodness-of-fit.

780 The bottom panels show the point-by-point residuals of the fitting for each scenario.

781 (B) Schematic cartoon representation for all-atom ensembles generated for the

782 different coexisting species in solution. RXR, RAR and N-CoR$_{NID}$ are colored in

783 orange, green and grey, respectively. ID1 (green) and ID2 (red) NR binding motifs

784 are represented as boxes in the N-CoR$_{NID}$ cartoon. The relative population of the

785 different species are inside boxes with the same color code as the associated

786 experimental SAXS. (C) ^{1}H-^{15}N HSQC intensities along N-CoR$_{NID}$ in complex with

787 wt RXR-RAR upon addition of RAR inverse agonist BMS493 (orange), or RAR

788 agonist Am580 (light green); or in complex with RXRΔH12-RAR (red) or

789 RXR/RARI396E (green). The colored bars represent the CoRNR1 and CoRNR2

790 motifs in green and pink, respectively.

791 **Figure 7: Interactions of RXR/RAR heterodimers with N-CoR using mammalian**

792 **two-hybrid assays with (17m)$_{5x}$-G-Luc reporter in COS cells**. (A) Transient

793 transfections with VP16-RAR and RXR, and with Gal-NCoR or Gal-CoRNR1 or Gal-

794 CoRNR2. Both CoRNR1 and CoRNR2 are able to interact with the RXR/RAR

795 heterodimer. (B) Transient transfections with Gal-NCoR and with VP16-RXR alone

796 or in combination with RAR. RXR only minimally interacts with N-CoR and RAR is

797 required for N-CoR-heterodimer complex formation. (C) Transient transfection with

798 Gal-NCoR and with VP16-RAR alone (dark grey) or in combination with RXR

799 (white) or RXRΔH12 (light gray). (D) Transient transfections with VP16-RAR and

800 RXRΔH12, and with Gal-CoRNR1 or Gal-CoRNR2. N-CoR interacts with the

801 RXRΔH12/RAR heterodimer through an interaction between CoRNR2 and the RXR

802 groove. In all assays TTNPB, an RAR-selective agonist activating RARα as

803 efficiently as Am580, and BMS493 were used at 10nM. Each transfection was carried

804 out in duplicate and repeated three to six times each.

805 **Figure 8: Schematic illustrations summarizing the structural and**

806 **thermodynamic data presented in this study.** The central part represents the free

807 partners, RXR/RAR and N-CoR$_{NID}$, and the three forms of the assembly. Size of

808 arrows connecting these states indicates the thermodynamic preferences observed in

809 our study. Equilibrium composition can be modified upon addition of ligands or the

810 presence of mutations on the RXR/RAR heterodimer. These effects are depicted as

811 gradient colored triangles on the top (affecting the RAR binding) and the right

812 (affecting the RXR binding) sides of the figure. Asymmetric form bound through the

813 CoRNR2 and the dissociated RXR/RAR can evolve towards the formation of

814 complex with a coactivator upon addition of an RAR agonist. This transition from a

815 repressive to an activated state of the NR is depicted on the rigth part of the figure.

816

817 **STAR Methods**

818 **Key Resources Table**

819 **Contact for Reagent and Resource Sharing**

33

820   Further information and requesets for resources and reagents should be directed to

821   and will be fulfilled by the Lead Contact, Pau Bernado (pau.bernado@cbs.cnrs.fr).

822   **Method Details**

823   **Ligands and peptides**

824   The ligands Am580, BMS493 and TTNPB were purchased from Tocris Bioscience.

825   The fluorescent peptides CoRNR1 (fluorescein-RLITLADHICQIITQDFAR) and

826   CoRNR2   (fluorescein-DPASNLGLEDIIRKALMGSFD)   from   N-CoR   were

827   purchased from EZbiolab.

828   **Constructs, expression and purification of proteins and complexes**

829   The heterodimer RXR/RAR (mRXRα/hRARα, Uniprot codes P28700 and P10276,

830   respectively) was prepared as previously described in (Pogenberg et al., 2005). The

831   heterodimer RXRΔH12/RAR (mRXRαΔH12/hRARα) comprises a shorter form of

832   RXR in which the region N227 to D444 including helix H12 is deleted. The

833   heterodimer RXR/RARI396E (mRXRα/hRARαI396E) comprises the I396E mutation

834   of RARα (Le Maire et al., 2010). The N-CoR$_{NID}$ protein corresponds to the sequence

835   from Gln2059 to Glu2325 of mouse N-CoR (Uniprot code Q60974) and was prepared

836   as described in (Harrus et al., 2018). For the preparation of the complexes, the

837   purified heterodimers (RXR/RAR or RXRΔH12/RAR) were mixed with a two-fold

838   molar excess of purified N-CoR$_{NID}$ fragment and incubated overnight at 4°C. Then,

839   the mix was loaded on a S200 superdex gel filtration column in a buffer consisting in

840   50mM Tris HCl pH7.5, 150mM NaCl, 2mM DTT and fractions corresponding to the

841   ternary complexes were pooled. The purity of the complexes was checked on a SDS-

842   page gel.

843   **Analytical ultracentrifugation sedimentation velocity experiments (AUC-SV)**

844 Samples of N-CoR$_{NID}$, RXR/RAR, N-CoR$_{NID}$:RXR/RAR, N-CoR$_{NID}$:RXR/RAR in

845 the presence of BMS493, and N-CoR$_{NID}$:RXRΔH12/RAR at about 1 g/L in 50 mM

846 Tris pH 7.4, 150 mM NaCl, 2 mM DTT, were investigated in SV experiments. SV

847 experiments were conducted in an XLI analytical ultracentrifuge (Beckman, Palo

848 Alto, CA) using an ANTi-50 rotor and detection at 280 and 295 nm, at 42 000 revs.

849 per minutes (rpm) (130 000g) and 12°C overnight, using double channel center pieces

850 (Nanolytics, Germany) of 12 mm optical path length (loaded volume: 400μL, the

851 reference channel being filled with the solvent) equipped with sapphire windows. For

852 each sample, because the total absorbance at 280 nm exceeded 1.2, the set of SV

853 profiles obtained at 295 nm was analyzed, using analysis in terms of a continuous

854 distribution, $c(s)$, of sedimentation coefficients, s, and of one non interacting species,

855 providing independent estimated of s and the molar mass, M, of the Sedfit software

856 (Schuck, 2000), version 14.1 (freely available at:

857 http://www.analyticalultracentrifugation.com). The related figures were made with

858 the program Gussi (Brautigam et al., 2016) (freely available at:

859 http://biophysics.swmed.edu/MBR/software.html). The values of s were corrected for

860 experimental conditions (12°C, solvent density, $\rho° = 1.007$ g/mL, and viscosity, $\eta°$

861 $= 1.235$ g/mL, estimated using the program Sednterp (freely available at:

862 http://sednterp.unh.edu/) to values, s_{20w}, at 20°C in pure water ($\rho° = 0.99823$ g/mL,

863 and viscosity, $\eta° = 1.002$ g/mL), and interpreted, through the Svedberg equation $s = M$

864 $(1 - \rho° \bar{v}) / (N_A 6\pi \eta° R_H)$, where M is the molecular mass, $\bar{v}$ is the partial specific

865 volume (0.720 mL/g for N-CoR$_{NID}$, 0.743 mL/g for RXR/RAR complexes, 0.735

866 mL/g for N-CoR$_{NID}$:RXR/RAR complexes, calculated using the program Sedfit), and

867 N_A is Avogadro's number. R_H is the hydrodynamic radius. $R_H = f/f_{min} R_{min}$, with f/f_{min}

868 the frictional ratio and R_{min} the radius of the anhydrous volume. The s_{20w}-values were

35

869 analyzed in term of f/f_{min} considering the the molar masses calculated from different

870 association states. Stoichiometries with f/f_{min} in the range 1.25 to 1.5 were considered

871 as acceptable, since corresponding to globular compact to reasonably

872 anisotropic/elongated or extended shapes. The results of the non-interacting species

873 analysis provided molar masses about 20% lower than the derived stoichiometry,

874 which can reasonably be related to the low signal/noise level of the data.

875 Mass spectrometry (RPLC-UV/MS)

876 Relative abundance of RAR, RXR and N-CoR$_{NID}$ in purified complex was assessed

877 using reverse phase liquid chromatography (Waters Alliance 2695) coupled to UV

878 (diode array detector) and MS (Waters LCT operated in the positive ion mode)

879 detection. Protein sample was diluted in buffer A (0.1% TFA / 0.1% FA /water)

880 before 600 pmol of complex were injected onto reverse phase Xbridge BEH300 C4

881 column (3.5µm; 50 × 2.1mm; T 65°C; flow rate 0.2 ml/min). Eluting gradient was

882 ramped from 2 to 90% buffer B (0.08% TFA / 0.1% FA / 70% ACN / 30% IprOH) in

883 40 min. Eluting species were detected by UV and mass spectrometry. Relative

884 abundance of RAR, RXR and N-CoR$_{NID}$ was calculated from protein theoretical

885 extinction coefficient and UV peak area (λ = 280 nm) according to equation:

886 $Abs_{280\,nm\,(P)} = \varepsilon_{280\,nm\,(P)} \times L \times [P]$ where "Abs" and "ε" correspond to protein

887 (P) absorbance and extension coefficient ($M^{-1}\,cm^{-1}$), respectively, at 280 nm. "L" is

888 the light path length (cm) and [P] corresponds to protein concentration (M).

889 Steady-state fluorescence anisotropy

890 We performed assays using a Safire microplate reader (TECAN) with the excitation

891 wavelength set at 470 nm and emission measured at 530 nm for fluorescein-tagged

892 peptides. The buffer solution for assays was 20 mM Tris-HCl, pH 7.5, 150 mM NaCl,

893 1 mM EDTA, 5 mM DTT and 10% (v/v) glycerol. We initiated the measurements at

36

894 the highest concentration of protein (20 μM) and diluted the protein sample

895 successively two-fold with the buffer solution. For each point of the titration curve,

896 we mixed the protein sample with 4 nM of fluorescent peptide and 40 μM of ligand

897 (two molar equivalents). We fitted binding data using a sigmoidal dose-response

898 model (GraphPad Prism, GraphPad Software). The reported data are the average of at

899 least three independent experiments.

MicroScale Thermophoresis (MST)

901 A fluorescent probe (Atto647N maleimide, Sigma) was attached to the thiol group of

902 the sole cysteine of the N-CoR$_{NID}$ fragment, according to the manufacturer's

903 instructions. This cysteine (Cys 2074) is in the CoRNR1 motif but is not involved in

904 the interaction between RAR and N-CoR as it points to the outside in the crystal

905 structure of the complex (Le Maire et al., 2010). Heterodimers (RXR/RAR,

906 RXRΔH12/RAR and RXR/RARI396E) were prepared as a twofold serial dilution in

907 MST buffer (NanoTemper Technologies GmbH) and added to an equal volume of 80

908 nM labelled N-CoR$_{NID}$, in MST buffer. After 10 min incubation time, the complex

909 was filled into Monolith NT.115 Premium Coated Capillaries (NanoTemper

910 Technologies GmbH) and thermophoresis was measured using a Monolith NT.115

911 Microscale Thermophoresis device (NanoTemper Technologies GmbH) at an ambient

912 temperature of 22°C, with 5 s/30 s/ 5 s laser off/on/off times, respectively. Instrument

913 parameters were adjusted with 40% red LED power and 20% IR-laser power. Data

914 from three independent measurements were analyzed (NT Analysis software last

915 version, NanoTemper Technologies GmbH) using the signal from Thermophoresis +

916 T-Jump.

Structural bioinformatics on N-CoR$_{NID}$

918   Local secondary structural features of N-CoR$_{NID}$ were predicted using the support

919   vector machine-based PSIPRED3.2 server (Buchan et al., 2013). Disorder propensity

920   of N-CoR$_{NID}$ was assessed with IUPRED (Dosztányi et al., 2005), PrDOS (Ishida and

921   Kinoshita, 2007), PONDR-FIT (Xue et al., 2010) and DISOPRED3 (Jones and

922   Cozzetto, 2015). The GREMLIN software (Kamisetty et al., 2013) was fed with the

923   N-CoR$_{NID}$ sequence to capture evolutionary conserved sequences by searching a pre-

924   clustered Uniprot database using the HHblits algorithm (Remmert et al., 2012). A

925   Hidden Markov Model (HMM) multiple sequence alignment (MSA) was generated

926   excluding hits with $\geq$ 75 % gaps. A final set of 61 sequences was used for detecting

927   conservation and co-evolution signatures at residue-level. Conservation profile was

928   extracted from the HMM MSA using Skylign to compute the probability of observing

929   for a given position a specific residue above the background (Wheeler et al., 2014).

930   Co-evolution analysis within the MSA was performed using BIS2 (Oteri et al., 2017),

931   which bypass the statistics requirements to estimate the background noise and the

932   relevance of the co-evolution signal. BIS2 is a combinatorial method that allows

933   identifying co-evolving blocks, instead of only residues, in a small number of

934   homologous sequences.

935   **Small-angle X-ray scattering**

936   Small-angle X-ray scattering data were measured at the BM29 beam-line from the

937   ESRF (Grenoble, France) using the X-ray wavelength of 0.99Å and a sample-to-

938   detector distance of ~2.9 m (Pernot et al., 2013). Several datasets were collected at

939   multiple concentrations (Table S1 and S3) in 50mM Tris HCl pH7.5, 150mM NaCl,

940   2mM TCEP. Prior to data collection, the isolated proteins and the ternary complexes

941   were supplemented with 2mM TCEP and concentrated. When necessary, ligands

942   (BMS493 and Am580) were added to the complex in a three molar excess before

943 concentration. Repetitive measurements allowed to detect and to correct for radiation

944 damage and scattering patterns of the buffer solutions were recorded before and after

945 the measurements of each protein sample. Final curves for each concentration were

946 derived after subtracting the averaged buffer scattering from the protein sample

947 patterns using PRIMUS. SAXS curves of different concentrations were merged to

948 avoid interparticle interactions using standard protocols with ATSAS software

949 (Franke et al., 2017).

950 The R_g values in Table S1 and Table S3 were estimated by applying the Guinier

951 approximation in the range $s < 1.3/R_g$ for globular proteins and $s < 0.8/R_g$ for free N-

952 CoR$_{NID}$. Pair-wise distance distribution functions, $P(r)$, were obtained by indirect

953 Fourier Transform of the scattering intensities with GNOM (Svergun et al., 1988). We

954 used CRYSOL (Svergun et al., 1995) to compute the theoretical SAXS profiles from

955 conformational ensembles of free N-CoR$_{NID}$, RXR/RAR, and N-CoR$_{NID}$:RXR/RAR

956 complexes. All theoretical curves were obtained with 101 points, and a maximum

957 scattering vector of 0.5 Å^{-1} using 25 harmonics. The resulting ensemble-based

958 CRYSOL SAXS curves, or their weighted combination, were directly compared to

959 the corresponding experimental SAXS data with OLIGOMER (Konarev et al., 2003).

960 **Nuclear Magnetic Resonance (NMR)**

961 NMR data were recorded using 140 μM U-[^{15}N] N-CoR$_{NID}$ and 600 μM U-[^{13}C/^{15}N]

962 N-CoR$_{NID}$ labeled protein in 50mM Bis Tris buffer at pH 6.7, 150mM NaCl, 1mM

963 EDTA and 2mM TCEP with 7% (v/v) ^{2}H$_2$O. NMR assignment experiments were

964 recorded at 293 K on a 950 MHz Bruker Avance III spectrometer equipped with a

965 cryogenic triple-resonance probe (IR-RMN Gif/Yvette). ^{1}H chemical shifts were

966 referenced directly, and ^{15}N chemical shifts indirectly (Markley et al., 1998), to added

967 2,2-dimethyl-2- silapentane-5-sulfonate (DSS, methyl ^{1}H signal at 0.00 ppm). A set of

968 3D Best-TROSY HNCO, HN(CA)CO, HN(CO)CA, HNCA, iHNCA,

969 HN(CO)CACB, HNCACB, iHNCACB (Lescop et al., 2007), and H-N-N experiments

970 HN(CA)NH, HN(COCA)NH. NMR interaction and PRE experiments (Barrett et al.,

971 2013) were recorded at 293 K on a 700 MHz Bruker Avance III spectrometer

972 equipped with a cryogenic triple-resonance probe. Spectra were processed with

973 nmrpipe and analyzed using nmrview (Johnson and Blevins, 1994). NMR ^{1}H-^{15}N

974 HSQC's have been recorded for ^{15}N labelled N-CoR$_{NID}$, and for ^{15}N-N-

975 CoR$_{NID}$:RXR/RAR (1:1), ^{15}N-N-CoR$_{NID:}$RXRΔH12/RAR (1:1), ^{15}N-N-CoR$_{NID:}$

976 RXR/RARI396E (1:1.2). ^{1}H-^{15}N HSQCs were also measured for equimolar

977 complexes of ^{15}N-N-CoR$_{NID}$:RXR/RAR in the presence of 1.2 molar excess of RAR

978 agonist (Am580) or RAR inverse agonist (BMS493). For PRE measurements, ^{1}H-^{15}N

979 HSQC spectra of diamagnetic and paramagnetic samples of ^{15}N-N-CoR$_{NID}$:RXR/RAR

980 (1:1.3) were acquired with a recycling delay between scans of 2 s to ensure that

981 magnetization recovery level is identical for both states and using the same

982 concentration and number of scans for both samples. The paramagnetic contribution

983 to the relaxation rate was determined as the ratio of peak intensities in the

984 paramagnetic and diamagnetic states.

985 **Site-directed spin-labeling for paramagnetic studies**

986 In addition to the wt cysteine residue C2074, two additional single-cysteine variants

987 of N-CoR$_{NID}$ (S2213C and S2288C) were engineered by site-directed mutagenesis

988 after substitution of the native cysteine by a serine (C2074S), using QuikChange

989 Lightning Site-Directed mutagenesis kit. All constructs were verified by DNA

990 sequencing. In order to conjugate the paramagnetic tag, a 10-fold excess of 3-(2-

991 Iodoacetamido)-2,2,5,5-tetramethyl-1-pyrrolidinyloxy (IA-PROXYL) spin label was

992 immediately added to reducing agent-free samples, just after elution from PD-10

40

993 desalting columns, and left reacting for 3 h in the dark at room temperature. Excess of

994 unreacted tag was removed by passing the reaction mixture twice through a PD-10

995 column, and exchanged to 50mM Bis Tris buffer at pH 6.7, 150mM NaCl, 1mM

996 EDTA. Reference diamagnetic samples were obtained after adding 5-fold of fresh

997 ascorbic acid to the same sample used to acquire the paramagnetic spectra.

998 **Structural modeling of RXR/RAR heterodimer**

999 Multiple structural all-atoms models of the RXR/RAR were built in order to describe

1000 SAXS curves of the heterodimer in different experimental conditions. The fully

1001 disordered H12 scenario was built using the 1DKF structure as a template and by

1002 adding the disordered N- and C-termini (including the H12) extensions as well as the

1003 connecting linker between helices H1 and H3. Previous crystallographic structures

1004 were used to model the other scenarios where one (RAR) or both (RAR and RXR)

1005 H12 regions were placed folded and bound to the rest of the LDB in either an agonist

1006 or antagonist conformation. Concretely, the structure 1DKF (Bourguet et al., 2000)

1007 was used as a template for the antagonist conformation, and the structures 1XDK

1008 (Pogenberg et al., 2005) and 3KMR (Le Maire et al., 2010) were used as templates for

1009 the agonist conformation (Fig. S1 and Table S2). Ensembles of disordered fragments

1010 were built using *Flexible-Meccano* (FM) (Bernado et al., 2005; Ozenne et al., 2012)

1011 and attached to the X-ray templates using *in-house* scripts (Cordeiro et al., 2017b).

1012 Briefly, for each disordered segment built, side-chains were added using SCCOMP

1013 (Eyal et al., 2004) and then pre-processed with Rosetta 3.5 *fixbb* module (Kuhlman et

1014 al., 2003) to alleviate steric clashes. Each structure was then refined in explicitly

1015 solvent using Gromacs 5.0.2 (Hess et al., 2008). All ensembles of RXR/RAR

1016 heterodimer comprised 5,000 conformers.

1017 **Structural ensembles of free N-CoR$_{NID}$**

41

A random coil ensemble model of N-CoR$_{NID}$ containing 10,000 conformations was calculated also employing FM (Bernado et al., 2005; Ozenne et al., 2012) followed by side-chains modelling and refinement as described above. A second structurally biased ensemble derived from the evolutionary conservation and a co-evolutionary analysis was computed in the following way. A starting random coil ensemble of 200,000 conformations of N-CoR$_{NID}$ was filtered selecting those structures presenting at least two contacts (≤ 15 Å) between residues in distal blocks of conserved (with Bits ≥ 1.6), structured (average disorder propensity below 0.5), and/or co-evolving residues (Fig. 2 and Fig. 3) within the C-terminal region. The filtered conformations ($\approx 35,000$) resulted in a compact ensemble that did not describe the SAXS data. We further refined it by performing 200 independent Ensemble Optimization Method (EOM) runs (Bernado et al., 2007), and those conformations that were not selected in any of them were discarded. The SAXS-refined ensembles was more extended than that obtained by distance restraints. The final ensemble of 10,000 conformations was use to interpret the PRE data. Briefly, for each conformer, the theoretical intramolecular ^{1}H-^{15}N-PRE-rates were computed using the Solomon-Bloembergen approximation (Iwahara et al., 2004; Salmon et al., 2010). Dynamics of the paramagnetic PROXYL moiety was accounted for by using multiple conformational states derived from a 100 ns Molecular Dynamics (MD) simulation of a Gly-Gly-**Cys**-Gly-Gly peptide in water at 298K, where the PROXYL moiety was attached to the central cysteine residue (Beck et al., 2008; Polyhach et al., 2011). From this library, the native and engineered single-cysteine residues were *in-silico* labelled with multiple sterically allowed spin-label dispositions. This ensemble representation enabled the estimation of the order parameters that account for the motion of the dipolar proton-electron interaction vector (Iwahara et al., 2004; Salmon et al., 2010).

42

1043   The correlation time of N-CoR$_{NID}$ was estimated to be 5.79 ns using HYCUD (Parigi

1044   et al., 2014). With this strategy, the PRE-rates for C2074, S2213C and S2288C were

1045   independently calculated for each N-CoR$_{NID}$ conformer, averaged over the complete

1046   ensembles of the three N-CoR$_{NID}$ forms, and compared with the experimental ones

1047   without any optimization process (Salmon et al., 2010).

1048   **Molecular Dynamics Simulation**

1049   Simulations were based on the one-bead-per-residue coarse-grained model proposed

1050   by (Smith et al., 2014) for intrinsically disordered proteins. N-CoR$_{NID}$ fragments with

1051   high α-helix propensity (2065-2088, 2214-2234, 2269-2291) were restrained to

1052   helical conformation *via* an elastic network with a force constant of 500 kJ.mol^{-1}.nm^{-2}

1053   whereas the original bonded potential for disordered chains was used for the

1054   remaining regions. Electrostatic interactions between charged residues were

1055   represented with a Debye Huckel energy functions ($\lambda = 0.9$ nm and $\varepsilon ES = 1.485$).

1056   Excluded volume and attractive hydrophobic interactions were modeled combining a

1057   purely repulsive Weeks-Chandler-Andersen (WCA) potential with the attractive part

1058   of a Lennard Jones potential ($\sigma = 0.58$ nm). The parameter αCG, which determines

1059   the strength of the hydrophobic interactions, was specifically tuned for N-CoR$_{NID}$

1060   system by minimizing the χ^2 between experimental and simulated SAXS curves. The

1061   latter were calculated from coarse-grained trajectories using CRYSOL (Svergun et al.,

1062   1995) upon reconstruction of the full atomistic structure via BBQ (Gront et al., 2007)

1063   and SCWRL4 (Krivov et al., 2009). Notably, this procedure revealed that an optimal

1064   agreement with SAXS data ($\chi^2 = 0.87$) was obtained for $\alpha CG = 0.5$, which is

1065   remarkably similar to the result obtained in the original publication for a set of diverse

1066   disordered proteins by comparison with FRET data. All simulations were performed

1067   with the GROMACS 5.4 molecular dynamics package (Abraham et al., 2015). An

43

1068   extended configuration was generated and used as initial structure for the simulation

1069   inside a 78733 nm^3 cubic box with periodic boundary conditions. The equations of

1070   motion were integrated every 20 fs during stochastic dynamics with a damping

1071   coefficient set to 25 amu.ps^{-1}. A plain cutoff of 5.0 nm was used for both Lennard-

1072   Jones and electrostatic interactions. All of the simulations were performed in the NVT

1073   ensemble and the temperature was kept constant to a reference value of 300 K. During

1074   the production, five replicas of 1 μs each were run for a total simulation length of 5

1075   μs.

Structural Modelling of N-CoR$_{NID}$:RXR/RAR complexes

1077   The structure of full-length N-CoR$_{NID}$ bound to RXR/RAR was modeled exploiting

1078   the existing crystal structure of the RAR/NCoR CoRNR1 complex (3KMZ) (Le Maire

1079   et al., 2010) and that of the RXR in complex with SMRT CoRNR2 complex (3R29)

1080   (Zhang et al., 1999), which has a 75% of sequence identity with NCOR CoRNR2.

1081   Using S-CoRNR2 as a template we created a homology model for bound CoRNR2. In

1082   the complexes, the conformations of bound CoRNR1 and/or CoRNR2 were

1083   maintained as in the crystal structures. For singly bound complexes (asymmetric),

1084   with N-CoR$_{NID}$ bound to RAR or RXR through CoRNR1 or CoRNR2 respectively,

1085   disordered statistical coil N- and C-terminal extensions were built onto the structured

1086   peptide employing the above-described work-flow to model ensembles of flexible

1087   fragments (Cordeiro et al., 2017a). The doubly bound complex was created in two

1088   steps. Firstly, we use FM to generate a structural ensemble of 100,000 singly bound

1089   complexes with N-CoR$_{NID}$ bound to RXR/RAR through CoRNR1, but with CoRNR2

1090   structured as in the crystal. Secondly, this starting ensemble was refined by selecting

1091   those conformers with CoRNR2 in the vicinity of RXR binding site (i.e. ≤ 6.5),

1092   followed by a docking process driven by the surface contacts observed in the

44

1093 RXR/SMRT CoRNR2 complex (PDB:3R29), which were integrated as distance

1094 restraints in the HADDOCK docking approach (Dominguez et al., 2003). Structures

1095 were then ranked using the energy-based HADDOCK scoring function and a term

1096 quantifying the RMSD of the binding interface with the RXR/SMRT CoRNR2

1097 complex. In the apo-form and in the presence of BMS493, RAR/RXR was modelled

1098 with disordered H12s. For RAR/RXRI396E mutant and in the presence of Am580,

1099 RAR H12 was placed in the agonist position, and RXR H12 was removed in the

1100 complex between N-COR$_{NID}$ and RXRΔH12/RAR. All ensembles of N-

1101 CoR$_{NID}$:RXR/RAR complexes contained 2,000 conformers.

1102 **Two-hybrid experiments**

1103 COS cells derived from African green monkey kidney (ATCC) were cultured in

1104 DMEM with Glutamax and 5% (v/v) FCS and transfected using JetPei transfectant

1105 (Ozyme). After 24 h, the medium was changed to a medium containing the indicated

1106 ligands or vehicle. Cells were lysed and assayed for reporter expression 48 h after

1107 transfection. The luciferase assay system was used according to the maunfacturer's

1108 instruction (Promega). In each case, results were normalized to coexpressed β–

1109 galactosidase. Each transfection was carried out in duplicate and repeated three to six

1110 times each.

1111 **Data availability.** The NMR chemical shifts of N-CoR$_{NID}$ have been deposited at the

1112 BMRM databank under the entry 27848. The SAXS data have been deposited at the

1113 SASBDB under the project "Nuclear receptor CoRepressor (N-CoR) and retinoic acid

1114 receptor heterodimer(RXR/RAR) Complex". The accession codes are SASDF34,

1115 SASDF44, SASDF54, SASDF64, SASDF74 and SASDF84 for Free Nuclear receptor

1116 CoRepressor NID, RXR/RAR Heterodimer:N-CoRNID Complex, RXRΔH12/RAR

1117 Heterodimer:N-CoRNID Complex, RXR/RAR Heterodimer:N-CoRNID Complex

45

1118 with RAR inverse agonist (BMS493), RXR/RARI396E Heterodimer:N-CoRNID

1119 Complex and RXR/RAR Heterodimer:N-CoRNID Complex with RAR agonist

1120 (Am580), respectively.

1121 References

1122 Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindah, E.

1123 (2015). Gromacs: High performance molecular simulations through multi-level

1124 parallelism from laptops to supercomputers. SoftwareX *1–2*, 19–25.

1125 Barrett, P.J., Chen, J., Cho, M.-K., Kim, J.-H., Lu, Z., Mathew, S., Peng, D., Song,

1126 Y., Van Horn, W.D., Zhuang, T., et al. (2013). The quiet renaissance of protein

1127 nuclear magnetic resonance. Biochemistry *52*, 1303–1320.

1128 Beck, D.A.C., Alonso, D.O.V., Inoyama, D., and Daggett, V. (2008). The intrinsic

1129 conformational propensies of the 20 naturally occurring amino acids and reflection

1130 of these propensies in proteins. Proc. Natl. Acad. Sci. *105*, 12259–12264.

1131 Bernado, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R.W., and Blackledge,

1132 M. (2005). A structural model for unfolded proteins from residual dipolar couplings

1133 and small-angle x-ray scattering. Proc. Natl. Acad. Sci. U S A *102*, 17002–17007.

1134 Bernado, P., Mylonas, E., Petoukhov, M. V, Blackledge, M., and Svergun, D.I.

1135 (2007). Structural characterization of flexible proteins using small-angle X-ray

1136 scattering. J. Am. Chem. Soc. *129*, 5656–5664.

1137 Bernadó, P., and Svergun, D.I. (2012). Structural analysis of intrinsically disordered

1138 proteins by small-angle X-ray scattering. Mol. BioSyst. *8*, 151–167.

1139 Bienkiewicz, E.A., Adkins, J.N., and Lumb, K.J. (2002). Functional consequences of

1140 preorganized helical structure in the intrinsically disordered cell-cycle inhibitor

1141 p27Kip1. Biochemistry *41*, 752–759.

1142 Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995). Crystal

1143   structure of the ligand-binding domain of the human nuclear receptor RXR-α. Nature

1144   *375*, 377-82.

1145   Bourguet, W., Vivat, V., Wurtz, J.M., Chambon, P., Gronemeyer, H., and Moras, D.

1146   (2000). Crystal structure of a heterodimeric complex of RAR and RXR ligand-

1147   binding domains. Mol. Cell *5*, 289–298.

1148   Brautigam, C.A., Zhao, H., Vargas, C., Keller, S., and Schuck, P. (2016). Integration

1149   and global analysis of isothermal titration calorimetry data for studying

1150   macromolecular interactions. Nat. Protoc. *11*, 882–894.Buchan, D.W.A., Minneci, F.,

1151   Nugent, T.C.O., Bryson, K., and Jones, D.T. (2013). Scalable web services for the

1152   PSIPRED Protein Analysis Workbench. Nucleic Acids Res. *41*, W349-57.

1153   Chandra, V., Wu, D., Li, S., Potluri, N., Kim, Y., and Rastinejad, F. (2017). The

1154   quaternary architecture of RARβ-RXRα heterodimer facilitates domain-domain signal

1155   transmission. Nat. Commun. *8*, 868.

1156   Chen, J.D., and Evans, R.M. (1995). A transcriptional co-repressor that interacts with

1157   nuclear hormone receptors. Nature *377*, 454–457.

1158   Cordeiro, T.N., Herranz-Trillo, F., Urbanek, A., Estaña, A., Cortés, J., Sibille, N., and

1159   Bernadó, P. (2017a). Small-angle scattering studies of intrinsically disordered

1160   proteins and their complexes. Curr. Opin. Struct. Biol. *42*, 15–23.

1161   Cordeiro, T.N., Chen, P.C., De Biasio, A., Sibille, N., Blanco, F.J., Hub, J.S.,

1162   Crehuet, R., and Bernadó, P. (2017b). Disentangling polydispersity in the PCNA-

1163   p15PAFcomplex, a disordered, transient and multivalent macromolecular assembly.

1164   Nucleic Acids Res. *45*, 1501–1515.

1165   Csizmok, V., Follis, A.V., Kriwacki, R.W., and Forman-Kay, J.D. (2016). Dynamic

1166   Protein Interaction Networks and New Structural Paradigms in Signaling. Chem Rev

1167   *116*, 6424–6462.

47

1168   Dominguez, C., Boelens, R., and Bonvin, A.M.J.J. (2003). HADDOCK: A protein-

1169   protein docking approach based on biochemical or biophysical information. J. Am.

1170   Chem. Soc. *125*, 1731–1737.

1171   Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: Web server for

1172   the prediction of intrinsically unstructured regions of proteins based on estimated

1173   energy content. Bioinformatics *21*, 3433–3434.

1174   Dyson, H.J., and Wright, P.E. (2004). Unfolded proteins and protein folding studied

1175   by NMR. Chem. Rev. *104*, 3607–3622.

1176   Egea, P.F., Rochel, N., Birck, C., Vachette, P., Timmins, P.A., and Moras, D. (2001).

1177   Effects of ligand binding on the association properties and conformation in solution of

1178   retinoic acid receptors RXR and RAR. J. Mol. Biol. *307*, 557–576.

1179   Eyal, E., Najmanovich, R., Mcconkey, B.J., Edelman, M., and Sobolev, V. (2004).

1180   Importance of Solvent Accessibility and Contact Surfaces in Modeling Side-Chain

1181   Conformations in Proteins. J. Comput. Chem. *25*, 712–724.

1182   Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A.,

1183   Mertens, H.D.T., Kikhney, A.G., Hajizadeh, N.R., Franklin, J.M., Jeffries, C.M., et al.

1184   (2017). ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering

1185   from macromolecular solutions. J. Appl. Crystallogr. *50*, 1212–1225.

1186   Germain, P., Iyer, J., Zechel, C., and Gronemeyer, H. (2002). Co-regulator

1187   recruitment and the mechanism of retinoic acid receptor synergy. Nature *415*, 187–

1188   192.

1189   Germain, P., Staels, B., Dacquet, C., Spedding, M., and Laudet, V. (2006). Overview

1190   of Nomenclature of Nuclear Receptors. Pharmacol. Rev. *58*, 685–704.

1191   Germain, P., Gaudon, C., Pogenberg, V., Sanglier, S., Van Dorsselaer, A., Royer,

1192   C.A., Lazar, M.A., Bourguet, W., and Gronemeyer, H. (2009). Differential Action on

1193 Coregulator Interaction Defines Inverse Retinoid Agonists and Neutral Antagonists.

1194 Chem. Biol. *16*, 479–489.

1195 Glass, C.K., and Rosenfeld, M.G. (2000). The coregulator exchange in transcriptional

1196 functions of nuclear receptors. Genes Dev. *14*, 121–141.

1197 Gronemeyer, H., Gustafsson, J.Å., and Laudet, V. (2004). Principles for modulation

1198 of the nuclear receptor superfamily. Nat. Rev. Drug Discov. *3*, 950–964.

1199 Gront, D., Kmiecik, S., and Kolinski, A. (2007). Backbone building from

1200 quadrilaterals: A fast and accurate algorithm for protein backbone reconstruction from

1201 alpha carbon coordinates. J. Comput. Chem. *28*, 1593–1597.

1202 Harrus D, Déméné H, Vasquez E, Boulahtouf A, Germain P, Figueira AC, Privalsky

1203 ML, Bourguet W, le M.A. (2018). Pathological Interactions Between Mutant Thyroid

1204 Hormone Receptors and Corepressors and Their Modulation by a Thyroid Hormone

1205 Analogue with Therapeutic Potential. Thyroid *28*, 1708-1722.

1206 Hegyi, H., Schad, E., and Tompa, P. (2007). Structural disorder promotes assembly of

1207 protein complexes. BMC Struct Biol *7*, 65.

1208 Heinzel, T., Lavinsky, R.M., Mullen, T.M., Söderström, M., Laherty, C.D., Torchia,

1209 J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R., et al. (1997). A complex containing

1210 N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature

1211 *387*, 43–48.

1212 Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008). GROMACS 4:

1213 algorithms for highly efficient, load balanced, and scalable molecular simulations. J.

1214 Chem. Theory Comput. *4*, 435–447.

1215 Hörlein, A.J., Näär, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A.,

1216 Kamei, Y., Söderström, M., Glass, C.K., et al. (1995). Ligand-independent repression

1217 by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature

1218 *377*, 397–404.

1219 Hu, X., and Lazar, M.A. (1999). The CoRNR motif controls the recruitment of

1220 corepressors by nuclear hormone receptors. Nature *402*, 93–96.

1221 Imhof, a, Yang, X.J., Ogryzko, V. V, Nakatani, Y., Wolffe, a P., and Ge, H. (1997).

1222 Acetylation of general transcription factors by histone acetyltransferases. Curr. Biol.

1223 *7*, 689–692.

1224 Ishida, T., and Kinoshita, K. (2007). PrDOS: Prediction of disordered protein regions

1225 from amino acid sequence. Nucleic Acids Res. *35*.

1226 Iwahara, J., Schwieters, C.D., and Clore, G.M. (2004). Ensemble Approach for NMR

1227 Structure Refinement against 1H Paramagnetic Relaxation Enhancement Data Arising

1228 from a Flexible Paramagnetic Group Attached to a Macromolecule. J. Am. Chem.

1229 Soc. *126*, 5879–5896.

1230 Jensen, M.R., Ruigrok, R.W.H., and Blackledge, M. (2013). Describing intrinsically

1231 disordered proteins at atomic resolution by NMR. Curr. Opin. Struct. Biol. *23*, 426-

1232 35.

1233 Johnson, B.A., and Blevins, R.A. (1994). NMR View: A computer program for the

1234 visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

1235 Johnson, B.A., Wilson, E.M., Li, Y., Moller, D.E., Smith, R.G., and Zhou, G. (2000).

1236 Ligand-induced stabilization of PPARgamma monitored by NMR spectroscopy:

1237 implications for nuclear receptor activation. J. Mol. Biol. *298*, 187–194.

1238 Jones, D.T., and Cozzetto, D. (2015). DISOPRED3: Precise disordered region

1239 predictions with annotated protein-binding activity. Bioinformatics *31*, 857–863.

1240 Kallenberger, B.C., Love, J.D., Chatterjee, V.K.K., and Schwabe, J.W.R. (2003). A

1241 dynamic mechanism of nuclear receptor activation and its perturbation in a human

1242 disease. Nat. Struct. Biol. *10*, 136–140.

1243 Kamisetty, H., Ovchinnikov, S., and Baker, D. (2013). Assessing the utility of

1244 coevolution-based residue-residue contact predictions in a sequence- and structure-

1245 rich era. Proc. Natl. Acad. Sci. *110*, 15674–15679.

1246 Klaholz, B.P., Mitschler, A., and Moras, D. (2000). Structural basis for isotype

1247 selectivity of the human retinoic acid nuclear receptor. J. Mol. Biol. *302*, 155–170.

1248 Konarev, P. V, Volkov, V. V, Sokolova, A. V, Koch, M.H.J., Svergun, D.I., and

1249 Koch, H.J. (2003). PRIMUS : a Windows PC-based system for small-angle scattering

1250 data analysis PRIMUS : a Windows PC-based system for small- angle scattering data

1251 analysis. Primus *36*, 1277–1282.

1252 Krivov, G.G., Shapovalov, M. V., and Dunbrack, R.L. (2009). Improved prediction of

1253 protein side-chain conformations with SCWRL4. Proteins Struct. Funct. Bioinforma.

1254 *77*, 778–795.

1255 Kuhlman, B., Dantas, G., Ireton, G.C., Varani, G., Stoddard, B.L., and Baker, D.

1256 (2003). Design of a Novel Globular Protein Fold with Atomic-Level Accuracy.

1257 Science (80-.). *302*, 1364–1368.

1258 Lescop, E., Schanda, P., and Brutscher, B. (2007). A set of BEST triple-resonance

1259 experiments for time-optimized protein resonance assignment. J. Magn. Reson. *187*,

1260 163–169.

1261 Lonard, D.M., and O'Malley, B.W. (2007). Nuclear Receptor Coregulators: Judges,

1262 Juries, and Executioners of Cellular Regulation. Mol. Cell *27*, 691–700.

1263 Lu, J., Cistola, D.P., and Li, E. (2006). Analysis of ligand binding and protein

1264 dynamics of human retinoid X receptor alpha ligand-binding domain by nuclear

1265 magnetic resonance. Biochemistry *45*, 1629–1639.

1266 le Maire, A., and Bourguet, W. (2014). Retinoic acid receptors: Structural basis for

1267 coregulator interaction and exchange. Subcell. Biochem. *70*, 37–54.

51

1268   le Maire, A., Alvarez, S., Shankaranarayanan, P., R de Lera, A., Bourguet, W., and

1269   Gronemeyer, H. (2012). Retinoid Receptors and Therapeutic Applications of

1270   RAR/RXR Modulators. Curr. Top. Med. Chem. *12*, 505–527.

1271   le Maire, A., Teyssier, C., Erb, C., Grimaldi, M., Alvarez, S., De Lera, A.R.,

1272   Balaguer, P., Gronemeyer, H., Royer, C.A., Germain, P., et al. (2010). A unique

1273   secondary-structure switch controls constitutive gene repression by retinoic acid

1274   receptor. Nat. Struct. Mol. Biol. *17*, 801–807.

1275   Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright,

1276   P.E., and Wuethrich, K. (1998). Recommendations for the presentation of NMR

1277   structures of proteins and nucleic acids - (IUPAC Recommendations 1998). Pure

1278   Appl. Chem. *70*, 117–142.

1279   McKenna, N.J., Lanz, R.B., and O'Malley, B.W. (1999). Nuclear receptor

1280   coregulators: Cellular and molecular biology. Endocr. Rev. *20*, 321–344.

1281   Mohan, A., Oldfield, C.J., Radivojac, P., Vacic, V., Cortese, M.S., Dunker, A.K., and

1282   Uversky, V.N. (2006). Analysis of Molecular Recognition Features (MoRFs). J. Mol.

1283   Biol. *362*, 1043–1059.

1284   Nagy, L., Kao, H.-Y., Chakravarti, D., Lin, R.J., Hassig, C.A., Ayer, D.E., Schreiber,

1285   S.L., and Evans, R.M. (1997). Nuclear Receptor Reression Mediated by a Complex

1286   Containing SMRT, mSin3A, and Histone Deacetylase. Cell *89*, 373–380.

1287   Nagy, L., Kao, H.Y., Love, J.D., Li, C., Banayo, E., Gooch, J.T., Krishna, V.,

1288   Chatterjee, K., Evans, R.M., and Schwabe, J.W.R. (1999). Mechanism of corepressor

1289   binding and release from nuclear hormone receptors. Genes Dev. *13*, 3209–3216.

1290   Nahoum, V., de Lera, A.R., and Bourguet, W. (2007). Modulator structure dynamics

1291   retinoid X receptor rel function agonist antagonist. Proc. Natl. Acad. Sci. USA *104*,

1292   17323-8.

1293 Oteri, F., Nadalin, F., Oteri, F., Nadalin, F., and Carbone, A. (2017). BIS2Analyzer :

1294 A server for co-evolution analysis of conserved protein families BIS2Analyzer : a

1295 server for co-evolution analysis of. Nucleic Acids Res. *45*.

1296 Osz, J., Pethoukhov, M. V., Sirigu, S., Svergun, D.I., Moras, D., and Rochel, N.

1297 (2012). Solution structures of PPARγ2/RXRα complexes. PPAR Res. 701412.

1298 Ota, H., and Fukuchi, S. (2017). Sequence conservation of protein binding segments

1299 in intrinsically disordered regions. Biochem. Biophys. Res. Commun. *494*, 602–607.

1300 Ozenne, V., Bauer, F., Salmon, L., Huang, J.R., Jensen, M.R., Segard, S., Bernadó,

1301 P., Charavay, C., and Blackledge, M. (2012). Flexible-meccano: A tool for the

1302 generation of explicit ensemble descriptions of intrinsically disordered proteins and

1303 their associated experimental observables. Bioinformatics *28*, 1463–1470.

1304 Parigi, G., Rezaei-Ghaleh, N., Giachetti, A., Becker, S., Fernandez, C., Blackledge,

1305 M., Griesinger, C., Zweckstetter, M., and Luchinat, C. (2014). Long-range correlated

1306 dynamics in intrinsically disordered proteins. J. Am. Chem. Soc. *136*, 16201–16209.

1307 Pavlin, M.R., Brunzelle, J.S., and Fernandez, E.J. (2014). Agonist ligands mediate the

1308 transcriptional response of nuclear receptor heterodimers through distinct

1309 stoichiometric assemblies with coactivators. J. Biol. Chem. *289*, 24771–24778.

1310 Perissi, V., and Rosenfeld, M.G. (2005). Controlling nuclear receptors: The circular

1311 logic of cofactor cycles. Nat. Rev. Mol. Cell Biol. *6*, 542–554.

1312 Perissi, V., Staszewski, L.M., Mcinerney, E.M., Kurokawa, R., Krones, A., Rose,

1313 D.W., Lambert, M.H., Milburn, M. V, Glass, C.K., and Rosenfeld, M.G. (1999).

1314 Molecular determinants of nuclear receptor − corepressor interaction Molecular

1315 determinants of nuclear receptor – corepressor interaction. Genes Dev. *13*, 3198–

1316 3208.

1317 Pernot, P., Round, A., Barrett, R., De Maria Antolinos, A., Gobbo, A., Gordon, E.,

1318   Huet, J., Kieffer, J., Lentini, M., Mattenet, M., et al. (2013). Upgraded ESRF BM29

1319   beamline for SAXS on macromolecules in solution. J. Synchrotron Radiat. *20*, 660–

1320   664.

1321   Pogenberg, V., Guichoul, J.F., Vivat-Hannah, V., Kammerer, S., Pérez, E., Germain,

1322   P., De Lera, A.R., Gronemeyer, H., Royer, C.A., and Bourguet, W. (2005).

1323   Characterization of the interaction between retinoic acid receptor/retinoid X receptor

1324   (RAR/RXR) heterodimers and transcriptional coactivators through structural and

1325   fluorescence anisotropy studies. J. Biol. Chem. *280*, 1625–1633.

1326   Polyhach, Y., Bordignon, E., and Jeschke, G. (2011). Rotamer libraries of spin

1327   labelled cysteines for protein studies. Phys. Chem. Chem. Phys. *13*, 2356–2366.

1328   Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2012). HHblits: Lightning-fast

1329   iterative protein sequence searching by HMM-HMM alignment. Nat. Methods *9*,

1330   173–175.

1331   Renaud, J.P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H., and

1332   Moras, D. (1995). Crystal structure of the RAR-gamma ligand-binding domain bound

1333   to all-trans retinoic acid. Nature *378*, 681–689.

1334   Rochel, N., Ciesielski, F., Godet, J., Moman, E., Roessle, M., Peluso-Iltis, C., Moulin,

1335   M., Haertlein, M., Callow, P., Mély, Y., et al. (2011). Common architecture of

1336   nuclear receptor heterodimers on DNA direct repeat elements with different spacings.

1337   Nat. Struct. Mol. Biol. *18*, 564–570.

1338   Roeder, R.G. (1998). Role of general and gene-specific cofactors in the regulation of

1339   eukaryotic transcription. Cold Spring Harbor Symposia on Quantitative Biology. *63*,

1340   201–218.

1341   Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation

1342   velocity ultracentrifugation and Lamm equation modeling. Biophys. J.

1343 Van Roey, K., Uyar, B., Weatheritt, R.J., Dinkel, H., Seiler, M., Budd, A., Gibson,

1344 T.J., and Davey, N.E. (2014). Short linear motifs: Ubiquitous and functionally diverse

1345 protein interaction modules directing cell regulation. Chem. Rev. *114*, 6733-78.

1346 Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M.R., Zweckstetter, M., and

1347 Blackledge, M. (2010). NMR characterization of long-range order in intrinsically

1348 disordered proteins. J. Am. Chem. Soc. *132*, 8407–8418.

1349 Sato, Y., Ramalanjaona, N., Huet, T., Potier, N., Osz, J., Antony, P., Peluso-Iltis, C.,

1350 Poussin-Courmontagne, P., Ennifar, E., Mély, Y., et al. (2010). The "phantom effect"

1351 of the rexinoid LG100754: Structural and functional insights. PLoS One *5*, e15119.

1352 Schulman, I.G., Juguilon, H., and Evans, R.M. (1996). Activation and repression by

1353 nuclear hormone receptors: hormone modulates an equilibrium between active and

1354 repressive states. Mol. Cell. Biol. *16*, 3807–3813.

1355 Smith, W.W., Ho, P.Y., and O'Hern, C.S. (2014). Calibrated Langevin-dynamics

1356 simulations of intrinsically disordered proteins. Phys. Rev. E - Stat. Nonlinear, Soft

1357 Matter Phys. *90*, 042709.

1358 Svergun, D., Barberato, C., and Koch, M.H. (1995). CRYSOL - A program to

1359 evaluate X-ray solution scattering of biological macromolecules from atomic

1360 coordinates. J. Appl. Crystallogr. *28*, 768–773.

1361 Svergun, D.I., Semenyuk, A. V., and Feigin, L.A. (1988). Small- angle- scattering-

1362 data treatment by the regularization method. Acta Crystallogr. Sect. A *44*, 244–251.

1363 Uversky, V.N., and Gillespie, J.R. (2000). Why are "natively unfolded" proteins

1364 unstructured under physiologic conditions? Proteins *41*, 415-27.

1365 de Vera, I.M.S., Zheng, J., Novick, S., Shang, J., Hughes, T.S., Brust, R., Munoz-

1366 Tello, P., Gardner, W.J., Marciano, D.P., Kong, X., et al. (2017). Synergistic

1367 Regulation of Coregulator/Nuclear Receptor Interaction by Ligand and DNA.

1368   Structure *25*, 1506–1518.e4.

1369   Westin, S., Kurokawa, R., Nolte, R.T., Wisely, G.B., McInerney, E.M., Rose, D.W.,

1370   Milburn, M. V., Rosenfeld, M.G., and Glass, C.K. (1998). Interactions controlling the

1371   assembly of nuclear-receptor heterodimers and co-activators. Nature *395*, 199–202.

1372   Wheeler, T.J., Clements, J., and Finn, R.D. (2014). Skylign: A tool for creating

1373   informative, interactive logos representing sequence alignments and profile hidden

1374   Markov models. BMC Bioinformatics *15*.

1375   Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010).

1376   PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. Biochim.

1377   Biophys. Acta - Proteins Proteomics *1804*, 996–1010.

1378   Zhang, H., Chen, L., Chen, J., Jiang, H., and Shen, X. (2011). Structural basis for

1379   retinoic X receptor repression on the tetramer. J. Biol. Chem. *286*, 24593–24598.

1380   Zhang, J., Hu, X., and Lazar, M. (1999). A novel role for helix 12 of retinoid X

1381   receptor in regulating repression. Mol. Cell. Biol. *19*, 6448–6457.

1382   Zheng, J., Chang, M.R., Stites, R.E., Wang, Y., Bruning, J.B., Pascal, B.D., Novick,

1383   S.J., Garcia-Ordonez, R.D., Stayrook, K.R., Chalmers, M.J., et al. (2017). HDX

1384   reveals the conformational dynamics of DNA sequence specific VDR co-activator

1385   interactions. Nat. Commun. *8*, 923.

1386

1387

1388

1389

1390

Figure with panels A–E showing Asymmetric-model (Singly-bound complexes) and Deck-model (Doubly-bound complex). Panel B plots χ^2 versus % Deck-model. Panel C plots $I(s)$ (a.u.) versus s (Å^{-1}) with $\chi^2 = 0.72$ and $\chi^2 = 13.97$, and Residuals. Panels D and E plot $\Delta^1H/^{15}N$ peaks and PRE ratio versus Residue number.
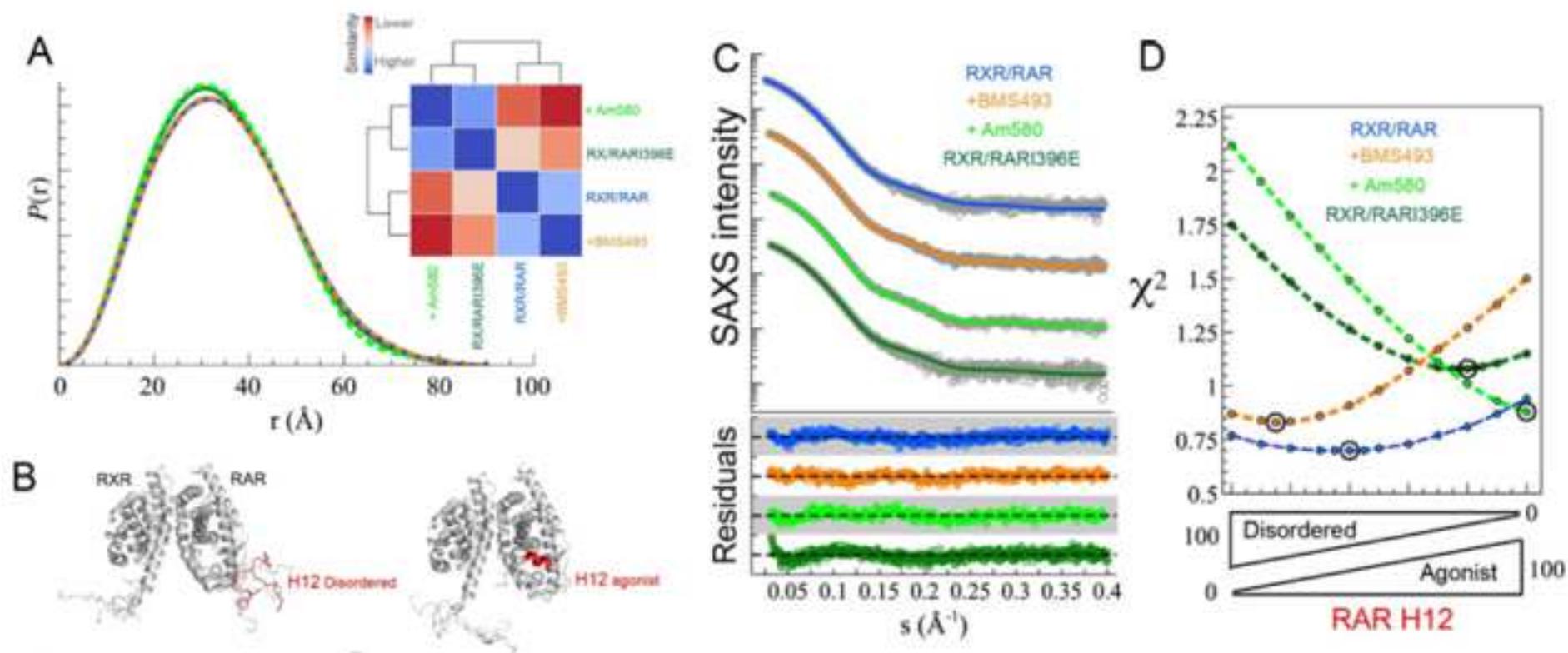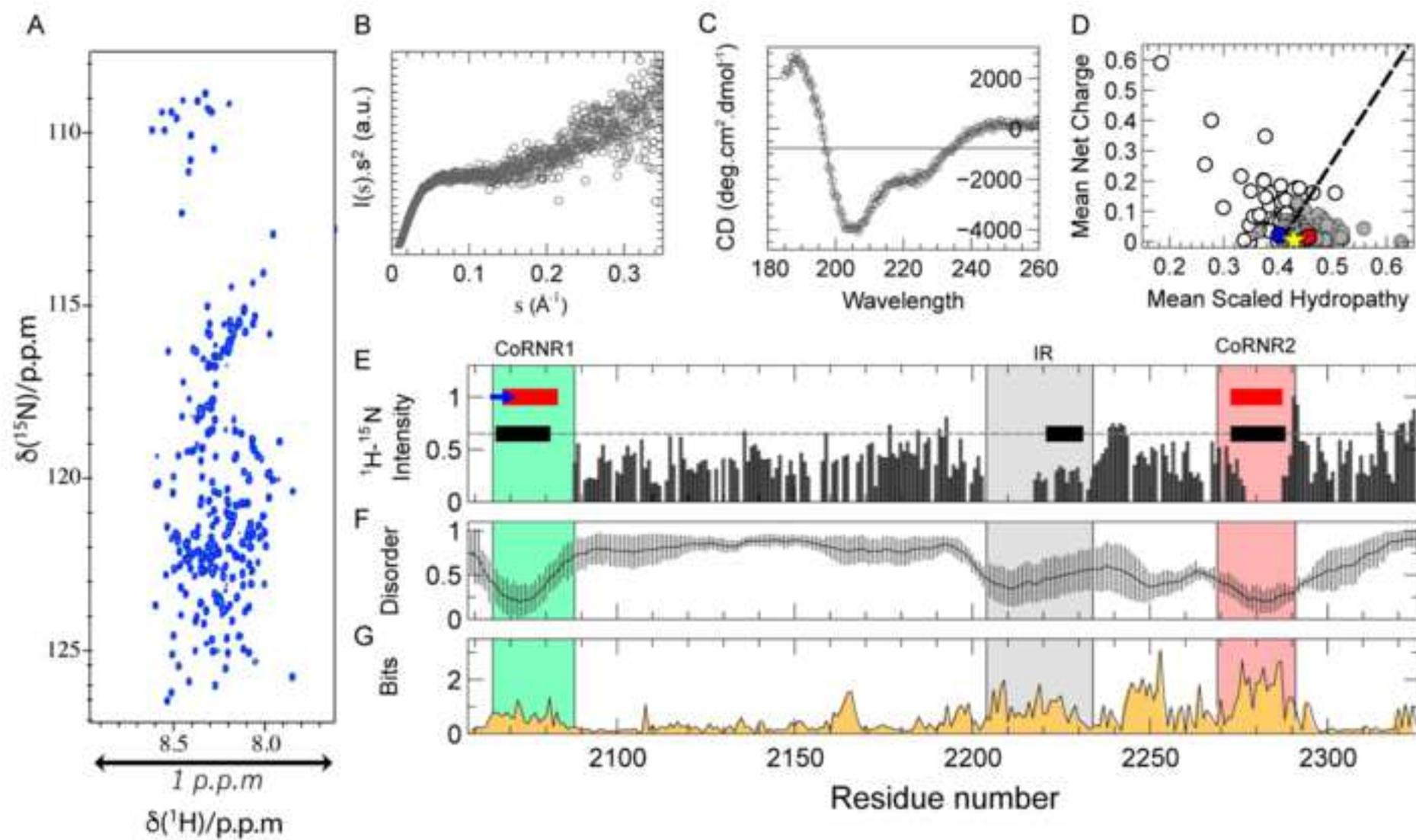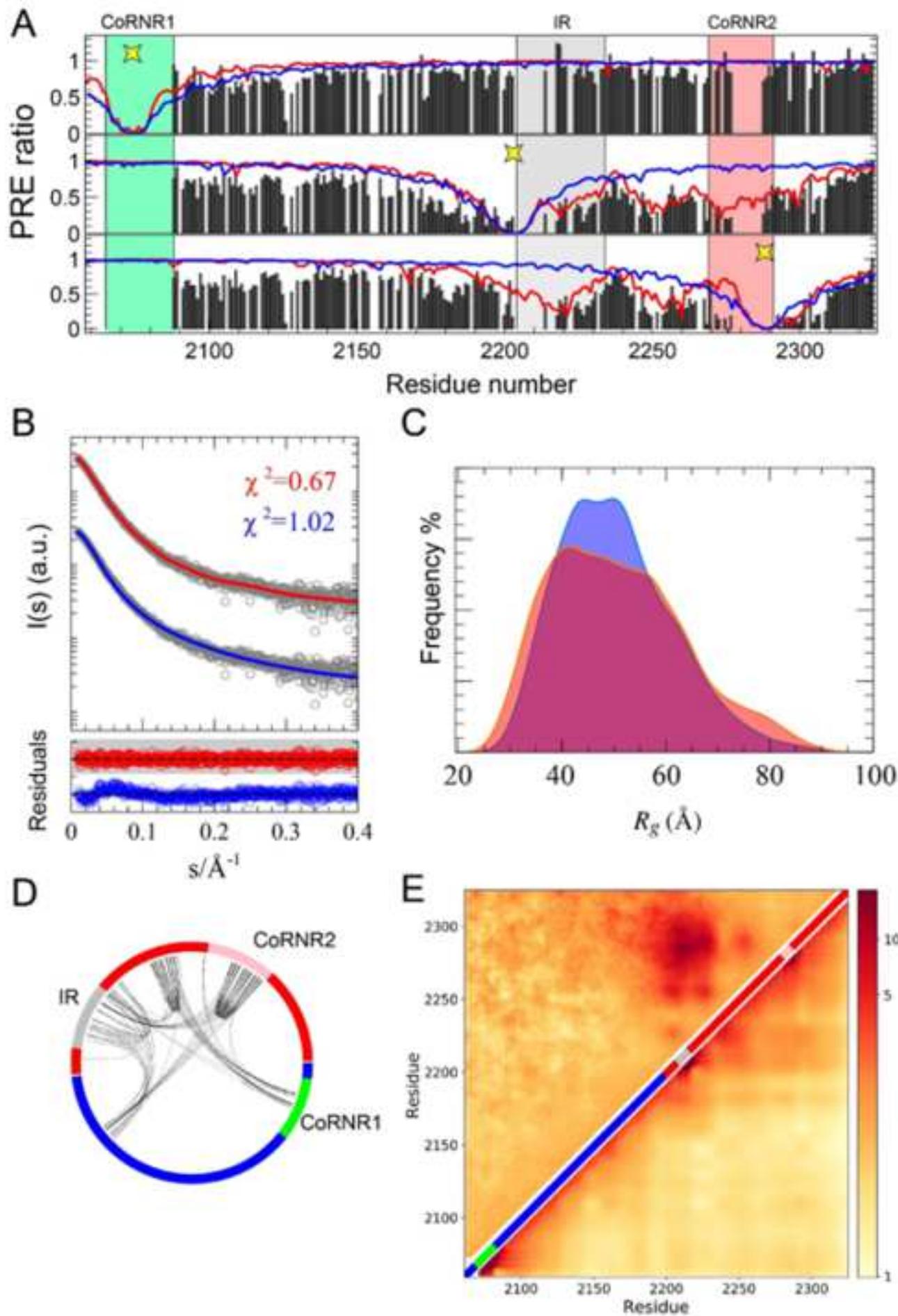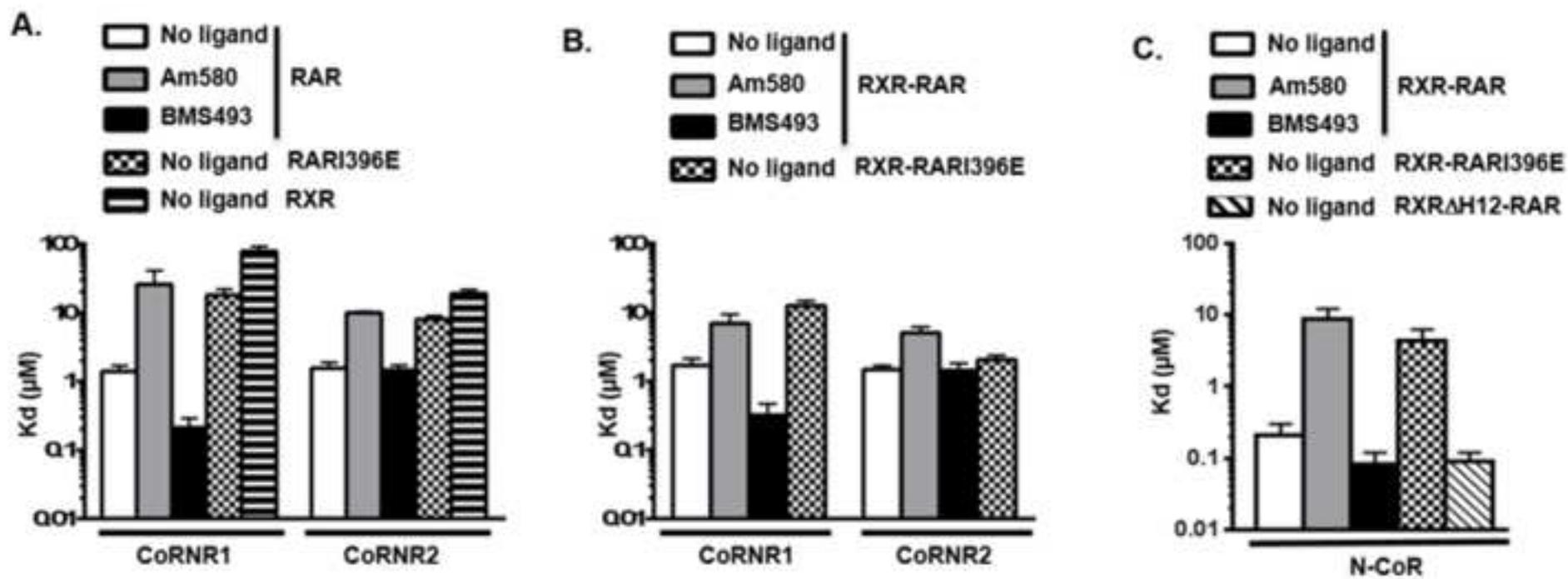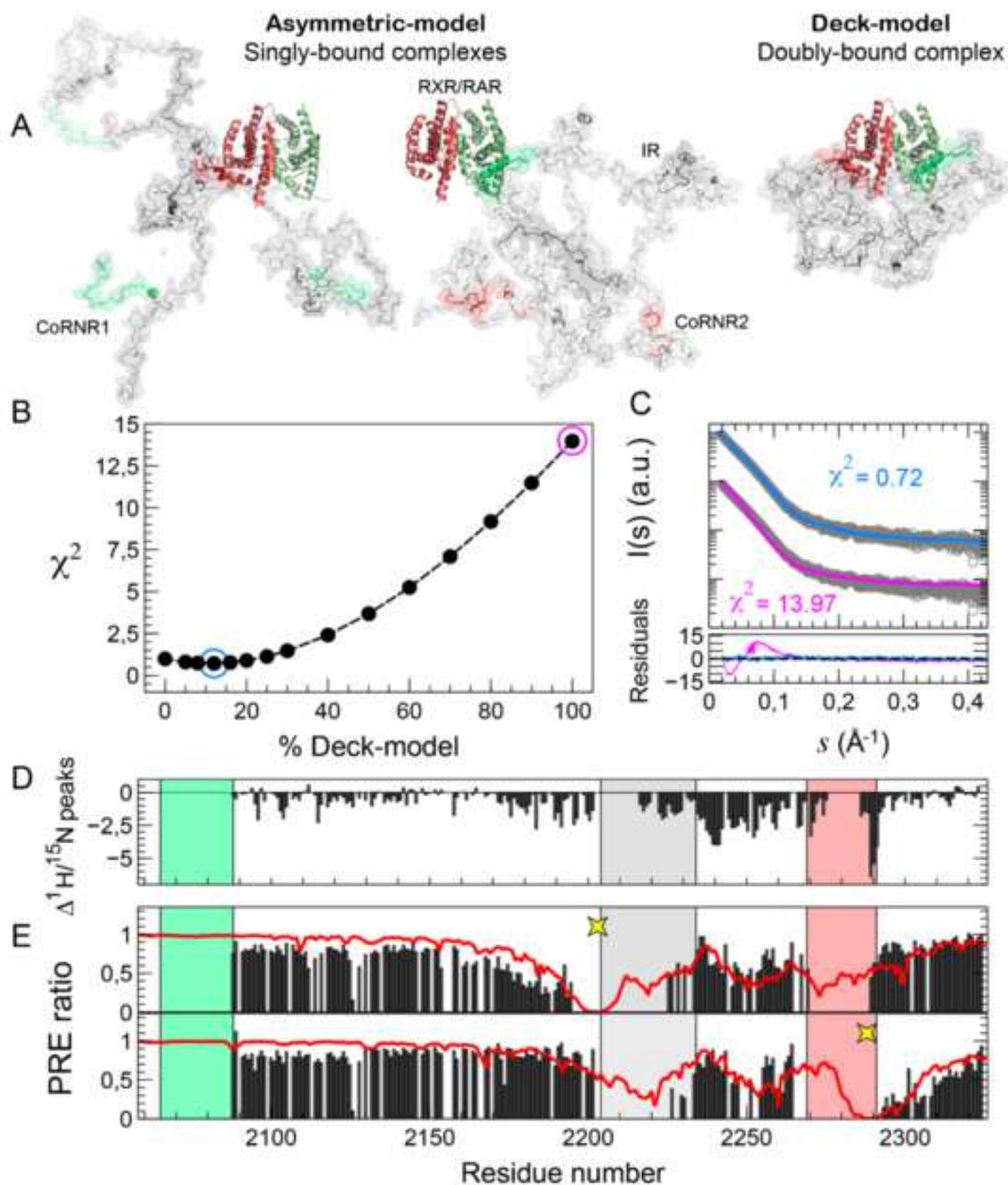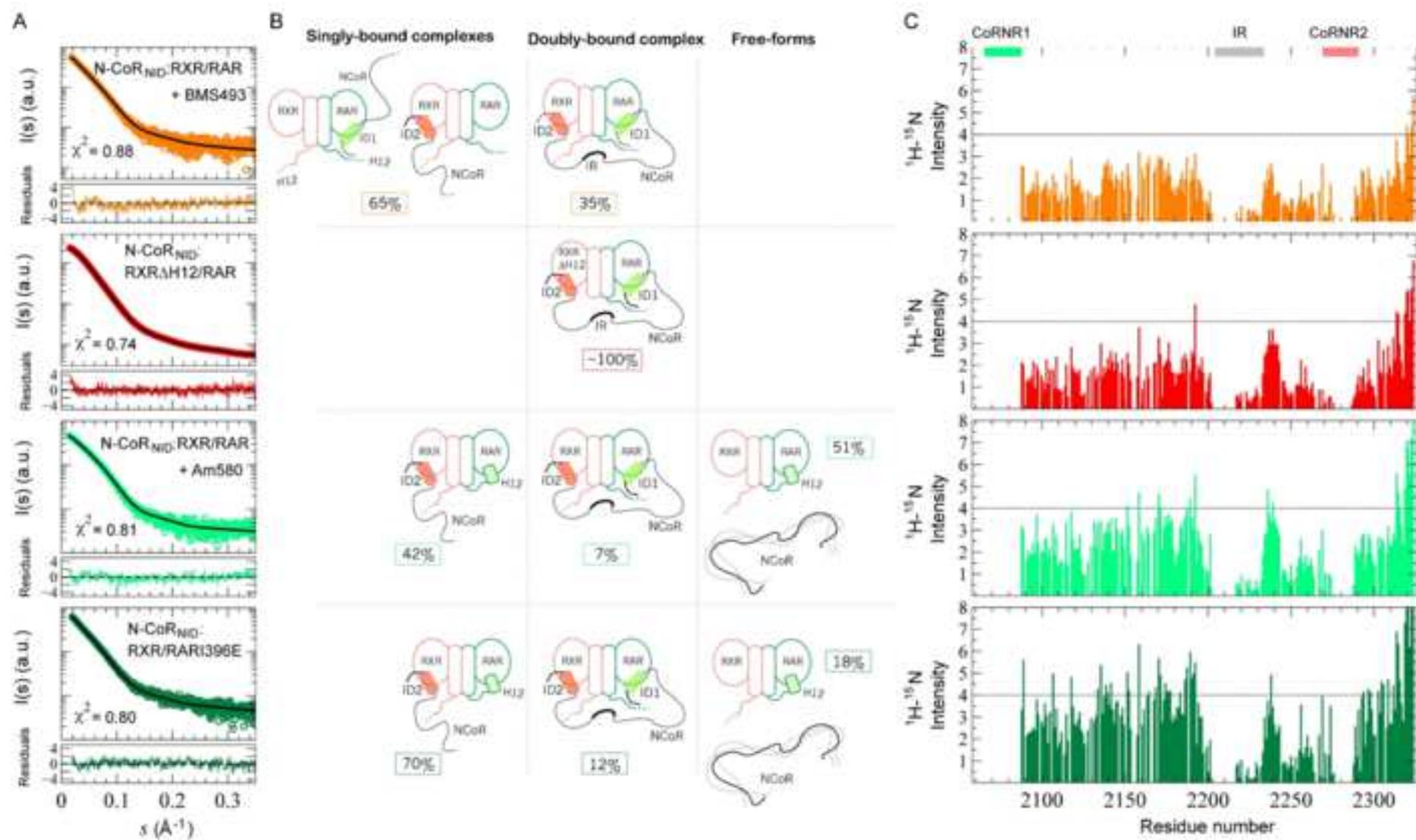
Figure
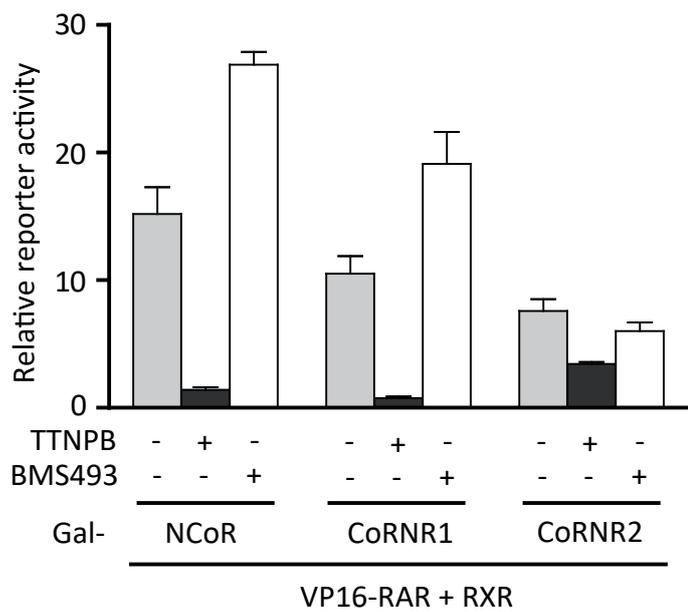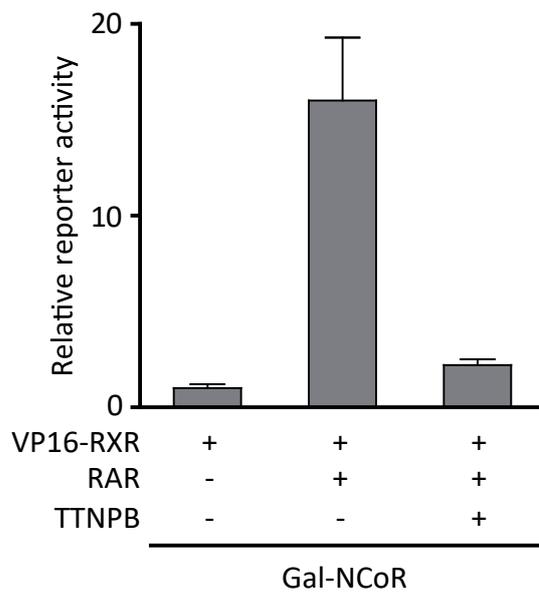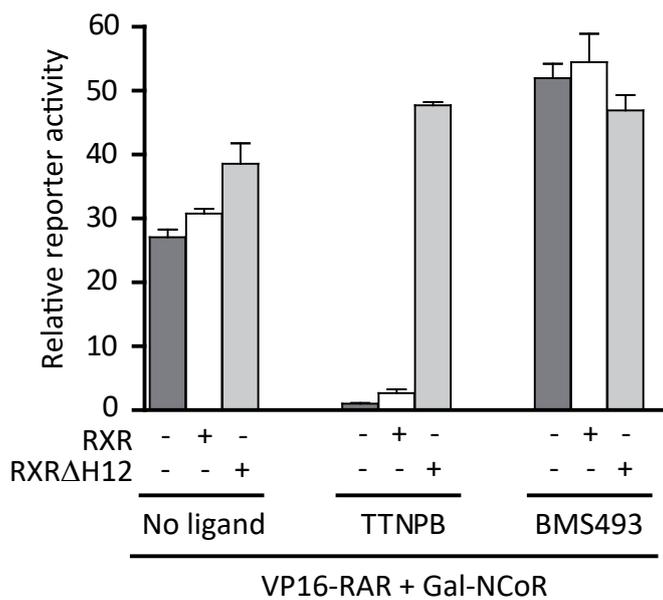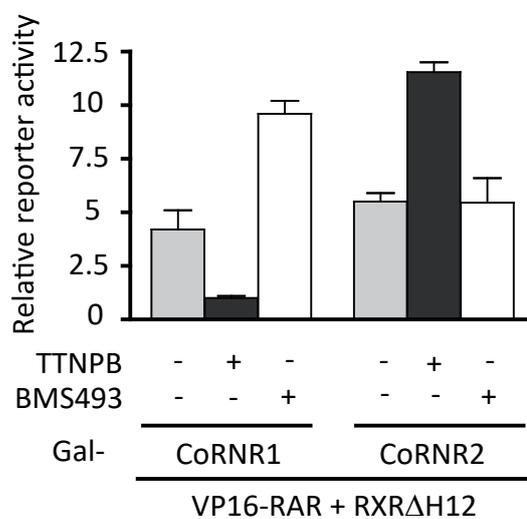
KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
BL21(*DE3*)	Novagen	Cat # 69450-4
Chemicals, Peptides, and Recombinant Proteins		
AM580	Tocris Bioscience	**Cat. No. 0760/10**
BMS493	Tocris Bioscience	**Cat. No. 3509/10**
TTNPB	Tocris Bioscience	**Cat. No. 0761/10**
Fluo-N-CoRNR1 (fluorescein-RLITLADHICQIITQDFAR)	EZbiolab	N/A
Fluo-N-CoRNR2 (fluorescein-DPASNLGLEDIIRKALMGSFD)	EZbiolab	N/A
Atto647N maleimide	SIGMA-ALDRICH	#05316
MST buffer	NanoTemper Technologies GmbH	N/A
3-(2-Iodoacetamido)-2,2,5,5-tetramethyl-1-pyrrolidinyloxy (IA-PROXYL)	Sigma	#253421
JetPei transfectant	Ozyme	POL101-10N
Critical Commercial Assays		
Luciferase assay system	Promega	E1500
Deposited Data		
Free Nuclear receptor CoRepressor NID	This paper	SASBDB: SASDF34
RXR/RAR Heterodimer:N-CoRNID Complex	This paper	SASBDB: SASDF44
RXRΔH12/RAR Heterodimer:N-CoRNID Complex	This paper	SASBDB: SASDF54
RXR/RAR Heterodimer:N-CoRNID Complex with RAR inverse agonist (BMS493)	This paper	SASBDB: SASDF64
RXR/RARI396E Heterodimer:N-CoRNID Complex	This paper	SASBDB: SASDF74
RXR/RAR Heterodimer:N-CoRNID Complex with RAR agonist (Am580)	This paper	SASBDB: SASDF84
NMR chemical shifts of N-CoR$_{NID}$	This paper	BMRM: 27848
RXRα (fatty acid)-RARα (antagonist) structure	(Bourguet et al., 2000)	PDB:**1DKF**
RXRα (agonist, CoA)-RARβ (agonist, CoA) structure	(Pogenberg et al., 2005)	PDB:1XDK
RARα (agonist, CoA) structure	(le Maire et al., 2010)	PDB:3KMR

RARα (inverse agonist, CoR) structure	(le Maire et al., 2010)	PDB:3KMZ
RXRα (S-CoR) structure	(Zhang et al., 2011)	PDB:3R29
Experimental Models: Cell Lines		
COS cells	ATCC	CRL-1650
Recombinant DNA		
mRXRα LBD-pET3a	Pogenberg et al. (2005)	N/A
hRARα LBD-pET15b	Pogenberg et al. (2005)	N/A
hRARαI396E LBD-pET15b	le Maire et al. (2010)	N/A
mRXRα ΔH12 LBD-pET3a	le Maire et al. (2010)	N/A
N-CORNID-PETTEV	Harrus et al. (2018)	N/A
Software and Algorithms		
GraphPad Prism	GraphPad Software	N/A
NT Analysis software	NanoTemper Technologies GmbH	N/A
PSIPRED3.2 server	(Buchan et al., 2013)	N/A
IUPRED	(Dosztányi et al. 2005)	N/A
PrDOS	(Ishida et al. 2007)	N/A
PONDR-FIT	(Xue et al. 2010)	N/A
DISOPRED3	(Jones et al, 2015)	N/A
GREMLIN software	(Kamisetty et al. 2013)	N/A
HHblits	(Remmert et al. 2012)	N/A
Skylign	(Wheeler et al. 2014)	N/A
BIS2Analyzer	(Oteri et al. 2017)	N/A
ATSAS Software	(Franke et al. 2017)	N/A
GNOM	(Svergun et al. 1988)	N/A
CRYSOL	(Svergun et al. 1995)	N/A
OLIGOMER	(Konarev et al. 2003)	N/A
NMR View	(Johnson et al. 1994)	N/A
Flexible-Meccano	(Bernado et al. 2005; Ozenne et al. 2012)	N/A
SCCOMP	(Eyal et al. 2004)	N/A
Rosetta 3.5 fixbb module	(Kuhlman et al. 2003)	N/A
Gromacs 5.0.2	(Hess et al. 2008)	N/A
EOM	(Bernado et al. 2007)	N/A
HYCUD	(Parigi et al. 2014)	N/A
GROMACS 5.4	(Abraham et al. 2015)	N/A
HADDOCK	(Dominguez et al. 2003)	N/A
Sedfit software	http://www.analyticalultracentrifugation.com	N/A
Gussi	http://biophysics.swmed.edu/MBR/software.html	N/A
Sednterp	http://sednterp.unh.edu/	

Other		
HiLoad 16/600 Superdex 200	GE Healthcare	Cat#28-9893-35
Safire microplate reader	TECAN	N/A
Monolith NT.115 Premium Coated Capillaries	NanoTemper Technologies GmbH	MO-K025
Monolith NT.115 Microscale Thermophoresis device	NanoTemper Technologies GmbH	N/A
BM29 beamline	ESRF	N/A
950 MHz Bruker Avance III spectrometer equipped with a cryogenic triple-resonance probe	IR-RMN Gif/Yvette	N/A
700 MHz Bruker Avance III spectrometer equipped with a cryogenic triple-resonance probe	In house	N/A
QuikChange Lightning Site-Directed mutagenesis kit	Agilent	# 210518
PD 10 Desalting Columns	GE Healthcare	# 17-0851-01
XLI analytical ultracentrifuge	Beckman, Palo Alto, CA	N/A

SUPPLEMENTARY INFORMATION FOR

Interplay of protein disorder in retinoic acid receptor heterodimer and its corepressor

regulates gene expression

Tiago N. Cordeiro[1,2], Nathalie Sibille[1], Pierre Germain[1], Philippe Barthe[1], Abdelhay Boulahtouf[3], Fréderic Allemand[1], Rémy Bailly[1], Valérie Vivat[4], Christine Ebel[5], Alessandro Barducci[1], William Bourguet[1], Albane le Maire[1,6,*], Pau Bernadó[1,*]

[1]Centre de Biochimie Structurale (CBS). CNRS, INSERM, Université de Montpellier. 29, rue de Navacelles. 34090-Montpellier, France

[2]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2790-157 Oeiras, Portugal

[3]IRCM, INSERM, ICM, Univ Montpellier, 34298 Montpellier, France

[4]NovAliX, 67400 Illkirch, France

[5]IBS, Univ. Grenoble Alpes, CEA, CNRS, 38000 Grenoble, France

[6]Brazilian Biosciences National Laboratory (LNBio), Brazilian Center for Research in Energy and Materials (CNPEM), 13083-970 Campinas, Sao Paulo, Brazil.

*Correspondence: albane.lemaire@cbs.cnrs.fr; pau.bernado@cbs.cnrs.fr

TABLE S1.

Table S1, related to Figure 1. SAXS acquisition and analysis of RXR/RAR heterodimers

Samples	RXR/RAR	RXR/RAR +BMS493	RXR/RAR +AM580	RXR/RARI396E
Data Collection Parameters				
Instrument (Pernot et al., 2013)	ESRF-BM29	ESRF-BM29	ESRF-BM29	ESRF-BM29
Wavelength (Å)	0.99	0.99	0.99	0.99
s range (Å^{-1})	0.0031-0.49	0.0031-0.49	0.0031-0.49	0.0032-0.493
Concentration range (mg·mL^{-1})	0.7, 1.4, 2.7,5.1	0.7, 1.4, 2.7, 5.1	0.7, 1.4, 2.7, 5.1	0.9,1.93,3.3,4.0
Exposure Time (s)	10	10	10	10
Temperature (K)	283	283	283	283
Structural parameters				
I_0 (cm^{-1}) [from $P(r)$]	0.3965±0.05	0.4407±0.1	0.3511±0.1	0.4224±0.1
R_g (Å) [from $P(r)$]	26.6±0.4	26.5±0.3	25.6±0.2	26.1±0.2
I_0 (cm^{-1}) [from Guinier]	0.4007	0.4483	0.3584	0.4266
R_g (Å) [from Guinier]	27.5± 0.2	27.1± 0.3	26.2± 0.4	26.8± 0.3
D_{max} (Å)	89.0± 3.0	90.0± 4.0	84.0±2.0	86.0±2.0
Porod volume estimate (Å^3)	92709.7	91962.5	86092.5	91008.5
Molecular mass determination				
Molecular mass (from I_0) (KDa)	53.1±5.0	59.3±6.0	48. 7±5.0	55.058±5.0
Calculated MW from sequence (KDa)	54.7	54.7	54.7	54.7
Software (Franke et al., 2017)				
1D data processing	PRIMUS	PRIMUS	PRIMUS	PRIMUS
$P(r)$	GNOM	GNOM	GNOM	GNOM
Simulated SAXS	CRYSOL	CRYSOL	CRYSOL	CRYSOL

TABLE S2.

Table S2, related to Figure 1 and Figure S1. List of X-ray structures used to generate ensemble models of RXR/RAR for SAXS data interpretation.

	X-ray structures				
	1DKF RXRα (fatty acid)-RARα (antagonist)(Bourguet et al., 2000)	**1XDK** RXRα (agonist, CoA)-RARβ (agonist, CoA)(Pogenberg et al., 2005)	**3KMR** RARα (agonist, CoA)(le Maire et al., 2010)	**3KMZ** RARα (inverse agonist, CoR)(le Maire et al., 2010)	**3R29,** RXRα (S-CoR)(Zhang et al., 2011)
Models Figure S1					
RXR/RAR-apo, disordered H12s	RXR/RAR				
	disordered RXR (residues 441-467) and disordered RAR H12 (residues 394-421)				
RXR/RAR-apo, antagonist H12s	RXR/RAR				
	RXR H12 antagonist conformation (residues 452-461) and RAR H12 antagonist conformation (residues 410-415)				
RXR/RAR-apo, agonist H12s			RXR RAR		
	RXR H12 agonist conformation (residues 454-459) and RAR H12 agonist conformation (residues 408-414)				
Models Figure 1					
RXR/RAR "agonist"	RXR		RAR		
	disordered RXR (residues 441-467) and RAR H12 agonist conformation (residues 408-414)				
RXR/RAR "disordered"	RXR			RAR	
	disordered RXR (residues 441-467) and disordered RAR H12 (residues 394-421)				
N-CoR$_{NID}$:RXR/RAR complexes					
Singly bound via RAR, disordered H12s	RXR			RAR:N-CoRNR1	
Singly bound via RXR, disordered H12s	RXR/RAR				S-CoRNR2
Singly bound via RXR, RAR-agonist	RXR		RAR		S-CoRNR2
Doubly bound	RXR			RAR:CoRNR1	S-CoRNR2

TABLE S3.

Table S3, related to Figure 6. SAXS acquisition and analysis of free N-CoR$_{NID}$ and N-CoR$_{NID}$:RXR/RAR complexes.

|	N-CoR$_{NID}$	N-CoR$_{NID}$: RXR/RAR	N-CoR$_{NID}$: RXR/RAR	N-CoR$_{NID}$: RXRΔH12/RAR	N-CoR$_{NID}$: RXR/RARI396E	N-CoR$_{NID}$: RXR/RAR
Ligand	—	—	BMS493	—	—	Am580
SASBDB code	SASDF34	SASDF44	SASDF64	SASDF54	SASDF74	SASDF84
Data-collection parameters						
Instrument	ESRF - BM29	ESRF - BM29	ESRF - BM29	ESRF - BM29	ESRF - BM29	ESRF - BM29
Wavelength (Å)	0.99	0.99	0.99	0.99	0.99	0.99
s range (Å^{-1})	0.0031-0.4952	0.0031-0.4952	0.0031-0.49 52	0.0027-0.445	0.0033-0.497	0.0031-0.497
Concentration range (mg·mL^{-1})	1.3, 1.6, 2.0	0.9, 1.7, 2.6, 4.0, 5.0	0.9, 1.3, 3.5	0.8, 1.2, 2.2, 4.9, 11.2	0.6, 0.9, 3.0, 3.5	0.9, 1.3, 3.5
Exposure Time (s)	10	10	10	10	10	10
Temperature (K)	283	283	283	283	283	283
Structural parameters						
I0 (cm^{-1}) [from Pr]	0.271± 0.04	1.155± 0.1	0.751± 0.1	2.715±0.1	0.849± 0.05	0.507± 0.05
R_g (Å) [from Pr]	50.3± 1	52.8± 2	52.6± 2	42.4± 2	51.3± 2	43.2± 1
I0 (cm^{-1}) [from Guinier]	0.2664	1.1859± 0.1	0.731± 0.1	2.77±0.1	0.804± 0.05	0.5234
R_g (Å) [from Guinier]	47.2± 1.2	48.4± 1.1	47.5± 1.0	42.0± 2	44.9± 1.0	44.2± 0.5
D_{max} (Å)	177± 6	194± 10	195± 10	157.0± 5	190± 5	172± 10
Porod volume estimate (Å^3)	101.897	163.06	178.3	182.771	170.65	131.09
Molecular mass determination						
Molecular mass [from I0] (kDa)	26.3± 3.0	120± 10.0	75.6± 8.0	N.D.	82.9± 9.0	50.3± 5.0
Molecular Mass [from sequence] (kDa)	29.17	83.53	83.53	83.53	85.44	83.53
Software						
1D data processing	PRIMUS	PRIMUS	PRIMUS	PRIMUS	PRIMUS	PRIMUS
$P(r)$	GNOM	GNOM	GNOM	GNOM	GNOM	GNOM
Simulated SAXS	CRYSOL	CRYSOL	CRYSOL	CRYSOL	CRYSOL	CRYSOL

FIGURE S1.
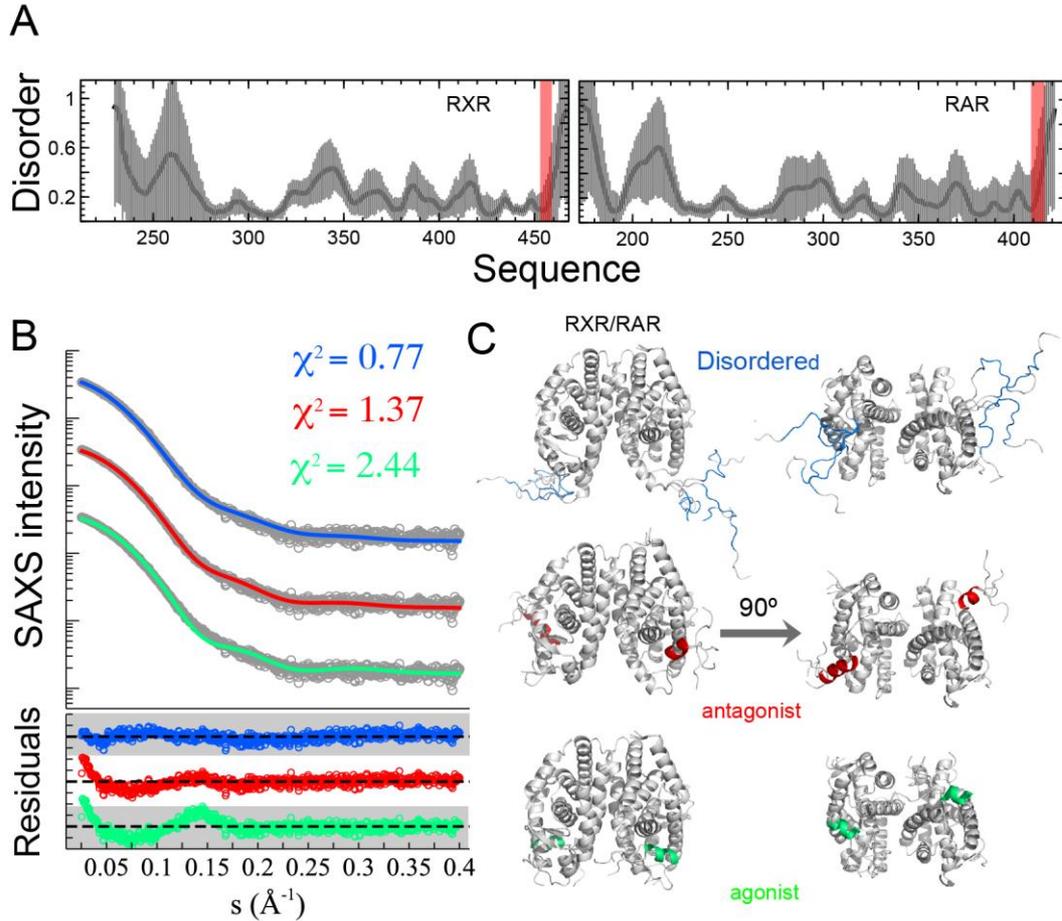
Figure S1, related to Figure 1: SAXS analysis of RXR/RAR heterodimer. (A) Average disorder prediction (solid grey line) and its standard deviation computed using different computational tools along the RXR and RAR LDB primary sequence. The limits of RXR H12 (residues 454-459) and RAR H12 (residues 408-414) are displayed in red. (B) The experimental SAXS profile for RXR/RAR in solution (black open circles) is compared with profiles derived from the structural models of the heterodimer displayed in (C). Point-by-point residuals for each model are shown at the bottom with the same colour code. The shaded regions correspond to ±5 residual points. (C) Cartoon representations of the ligand-binding domain (LBD) of RXR/RAR models generated based on the X-ray structures referred on Table S2 (both helices H12 disordered, in antagonist and agonist conformations in the blue, red and green models, respectively).

FIGURE S2.

Sample	Ligand	s (S)	s_{20w} (S)
RXR/RAR	-	3.02 +/- 0.10	3.78
RXRΔH12/RAR	-	4.65 +/- 0.10	5.81*
N-CoR$_{NID}$	-	1.9 +/- 0.15	2.37
N-CoR$_{NID}$:RXR/RAR	-	3.96 +/- 0.10	4.93
N-CoR$_{NID}$:RXR/RAR	BMS493	3.63 +/- 0.10	4.52
N-CoR$_{NID}$:RXRΔH12/RAR	-	3.51+/- 0.10	4.37
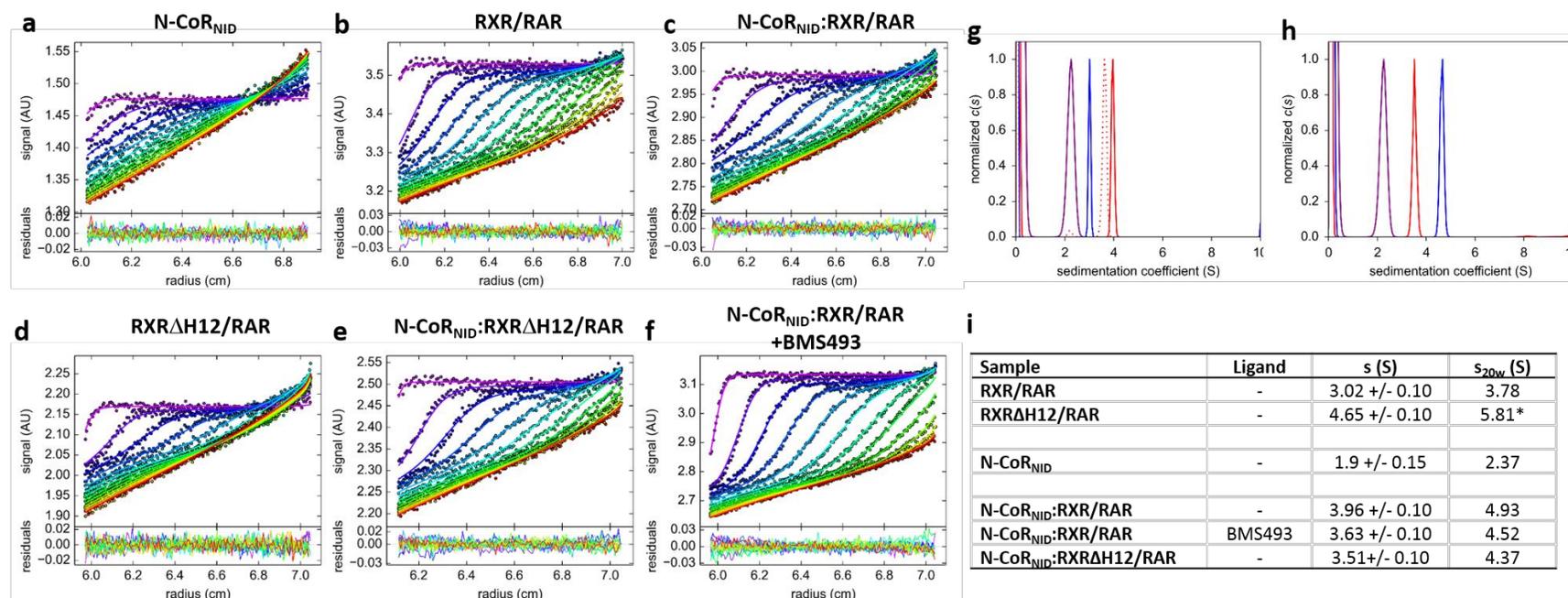
Figure S2, related to Figures 1 and 2: Biophysical parameters derived from Analytical Ultracentrifugation Sedimentation velocity experiments. (a-f) Sedimentation velocity of N-CoR$_{NID}$, RXR/RAR, RXRΔH12/RAR and complexes. Superposition of selected experimental and fitted sedimentation velocity profiles obtained every 45 min. at 42,000 rpm, 12 °C, at 295 nm (top subpanels) and their differences (bottom subpanels). (g-h) Sedimentation coefficient distributions c(s) for N-CoR$_{NID}$ (purple), RXR/RAR (blue, g) or RXRΔH12/RAR (blue, h), N-CoR$_{NID}$:RXR/RAR (continuous red, g) or N-CoR$_{NID}$:RXRΔH12/RAR (continuous red, h) and N-CoR$_{NID}$:RXR/RAR/BMS493 (dotted red) complexes. Protein concentration in each sample was about 1 g/L. (i) Experimental (*s*) and corrected to water at 20°C (s_{20w}) sedimentation coefficients for free and bound forms of RXR/RAX and N-CoR$_{NID}$. All species except RXRΔH12-RAR were considered with the lowest possible association state, since other stoichiometry would correspond to very elongated shapes. Note that in its unliganded form, the RXRΔH12/RAR heterodimer forms a dimer in solution, as observed in gel filtration.

FIGURE S3.

A)

B)

Figure S3, related to Figure 2: NMR assignment of free N-CoR$_{NID}$ in solution. (A) Primary sequence of N-CoR$_{NID}$, where assigned residues are underlined. Regions coloured in green, grey and pink correspond to CoRNR1, IR and CoRNR2 regions, respectively. The conserved LXXIIXXXL motifs of the CoRNR boxes are displayed in bold, and their secondary structure found in crystallographic structures are represented as a blue arrow for β-strand and red cylinder for α-helix. (B) Assigned ^{1}H-^{15}N HSQC spectrum measured for ^{15}N/^{13}C labelled N-CoR$_{NID}$ (950 MHz, 293 K). The NMR assignment was performed for the fragment Gly2057-Glu2325 of mouse N-CoR.

FIGURE S4.

A.

N-CoR_NID:RXR/RAR/BMS493 (600pmoles)

BMS493

RXR

RAR

N-CoR_NID

λ = 280 nM

Protein	Extinction Coefficient (M-1 cm-1)	UV peak area	Ratio
RXR	16960	23374	[RXR]/[RAR] = 1.1
RAR	11460	14142	[RAR]/[N-CoR_NID] = 1.1
N-CoR_NID	4470	4982	[RXR]/[N-CorR_NID] = 1.2

B.

Figure S4, related to Figure 5: Characterization of N-CoR_NID:RXR-RAR complex. (A) UV chromatogram (λ 280nm) of the complex N-CoR_NID:RXR-RAR in the presence of BMS493 submitted to RPLC. Theoretical extinction coefficients and UV peak area were used to determine the relative abundance of RXR, RAR and N-CoR_NID. As shown in the table, quasi stoichiometric ratios were obtained for all three proteins. (B) Coomassie-stained SDS-PAGE of N-CoR_NID, RXR-RAR, RXRΔH12-RAR and the corresponding complexes after the final purification step.

FIGURE S5.
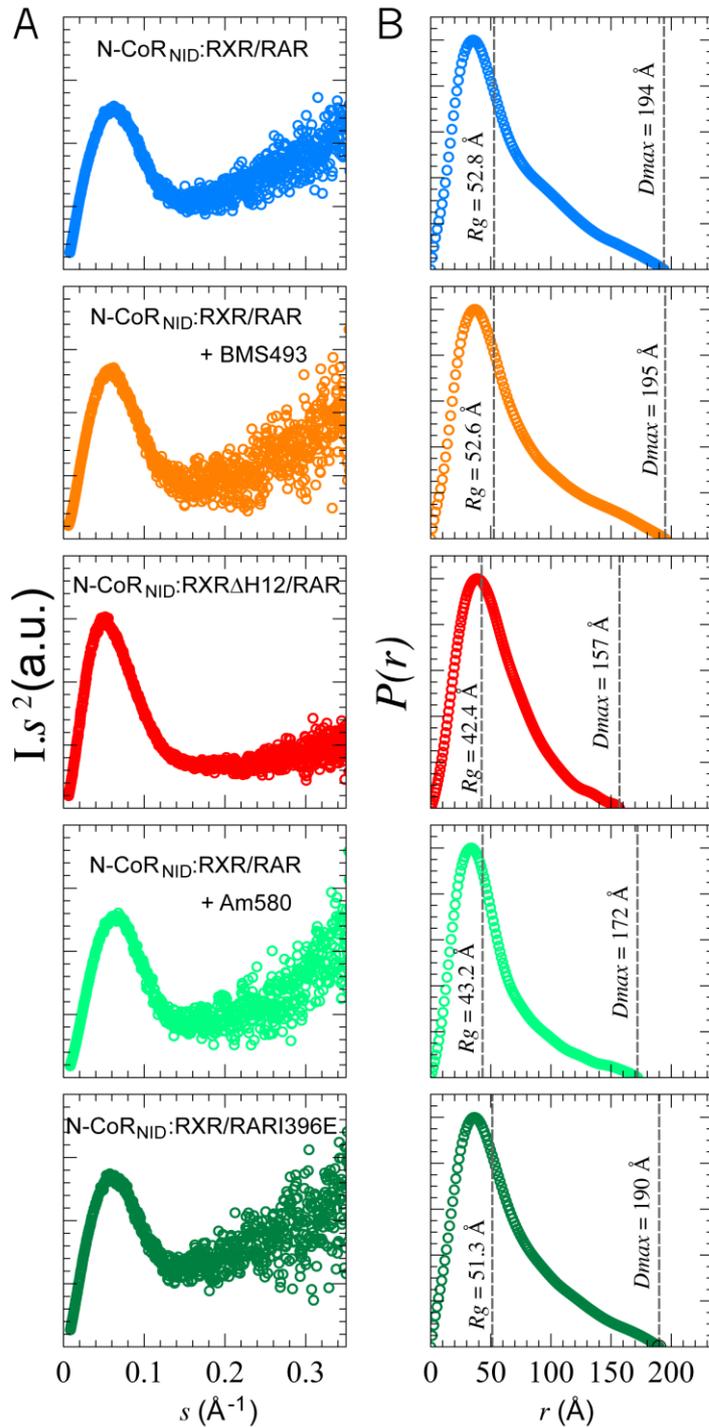
Figure S5, related to Figures 5 and 6: SAXS analysis of N-CoRNID: RXR/RAR complexes. (A) Kratky representations of SAXS intensity versus momentum transfers (open circles) of N-CoR$_{NID}$:RXR/RAR in the absence of ligand (blue), in the presence of inverse agonist BMS493 (orange) or agonist Am580 (light green), as well as N-CoR$_{NID}$:RXRΔH12/RAR (red) and N-CoR$_{NID}$:RXR/RARI39E (green) variants. (B) Asymmetric P(r) functions determined from each SAXS data are in the same colour code. The derived Rg and D_{max} values are in displayed in dashed lines.

FIGURE S6.
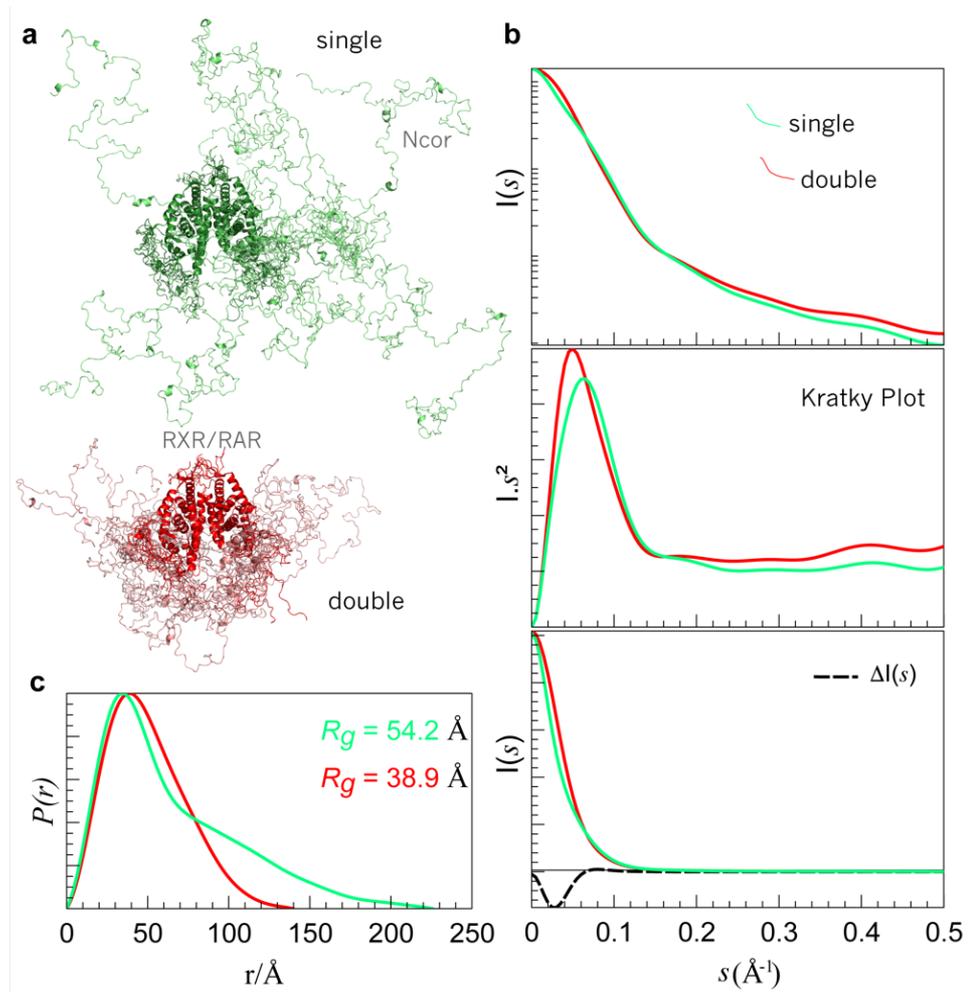
Figure S6, related to Figure 5: SAXS properties of the simulated N-CoR$_{NID}$:RXR-RAR complexes.
(A) Cartoon representations and ensemble models for both structural models: asymmetric singly-bound (green) and deck doubly-bound model (red). Theoretical SAXS profiles and Kratky plots (B) and *P(r)* (C) from both models with the same colour code. The difference between the two SAXS curves in absolute scale is depicted as a dashed black line in the bottom panel of (B).

FIGURE S7.
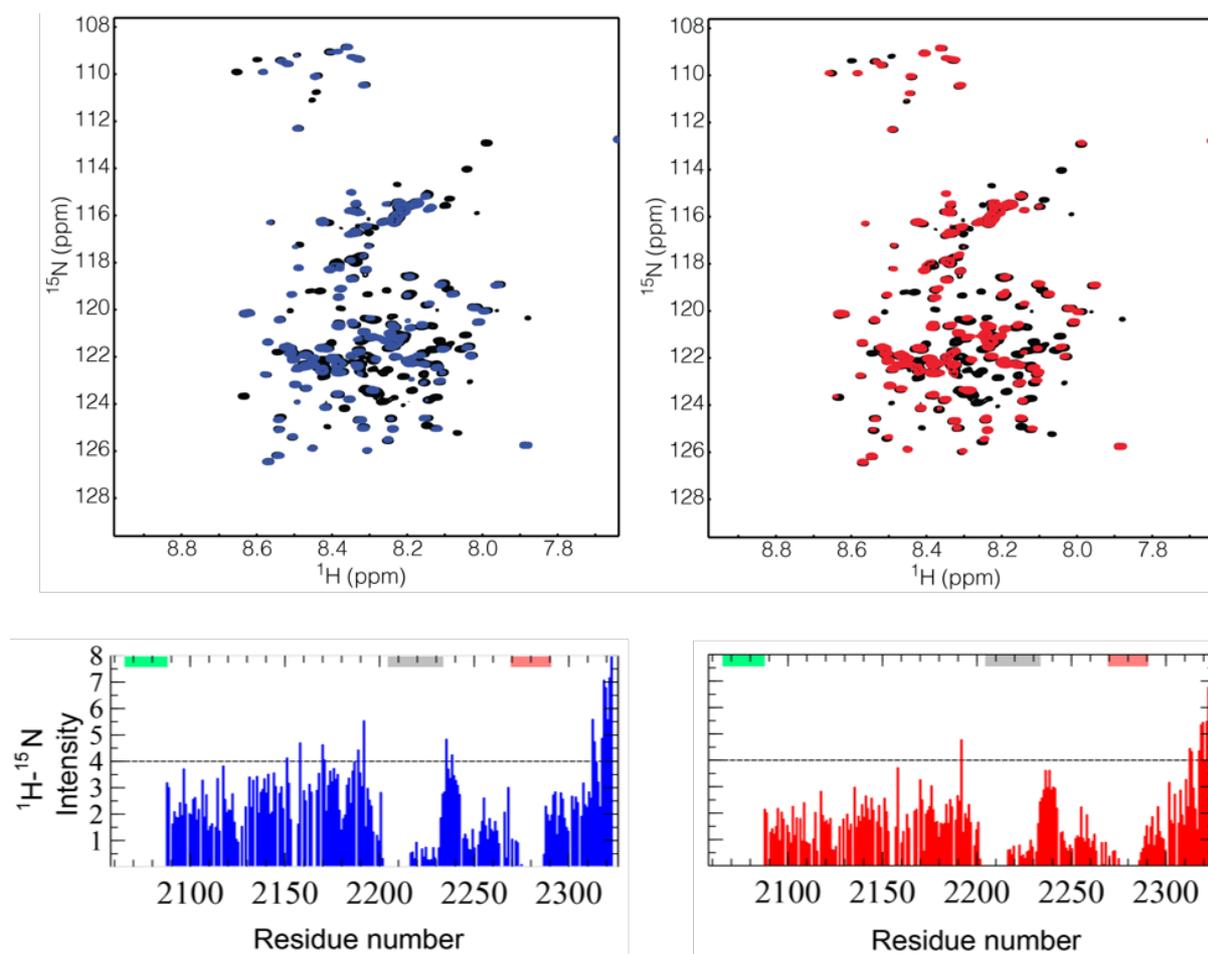
Figure S7, related to Figure 6: Interaction of N-CoR$_{NID}$ with RXR/RAR heterodimers. (A) ^{1}H-^{15}N HSQC's of ^{15}N labelled N-CoR$_{NID}$ free in solution (black) and in the presence of equimolar concentration of RXR-RAR (blue, left), and RXRΔH12-RAR (red, right). The spectra show that no chemical shift variation or the appearance of additional peaks in the presence of the partners, and only a decrease of intensity due to the increased molecular tumbling of N-CoR$_{NID}$ as observed in the bottom panels for both complexes with the same color code.