



**HAL**  
open science

## **IDPs and their complexes in GPCR and nuclear receptor signaling**

Myriam Guillien, Albane Le Maire, Assia Mouhand, Pau Bernadó, William Bourguet, Jean-louis Baneres, Nathalie Sibille

► **To cite this version:**

Myriam Guillien, Albane Le Maire, Assia Mouhand, Pau Bernadó, William Bourguet, et al.. IDPs and their complexes in GPCR and nuclear receptor signaling. *Progress in Molecular Biology and Translational Science*, 2020, Dancing Protein Clouds: Intrinsically Disordered Proteins in Health and Disease, Part B, 174, pp.105 - 155. 10.1016/bs.pmbts.2020.05.001 . hal-03749862

**HAL Id: hal-03749862**

**<https://hal.science/hal-03749862>**

Submitted on 12 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **IDPs and their complexes in GPCR and nuclear receptor signaling**

**Myriam Guillien<sup>a</sup>, Albane le Maire<sup>a</sup>, Assia Mouhand<sup>a</sup>, Pau Bernadó<sup>a</sup>, William Bourguet<sup>a</sup>, Jean-Louis Banères<sup>b</sup>, Nathalie Sibille<sup>a,\*</sup>**

<sup>a</sup>Centre de Biochimie Structurale (CBS), Inserm, CNRS, Univ Montpellier, Montpellier, France

<sup>b</sup>Institut des Biomolécules Max Mousseron (IBMM), UMR 5247 CNRS, Université Montpellier, ENSCM, 15 av. Charles Flahault, 34093, Montpellier, France.

\*Corresponding author: e-mail address: nathalie.sibille@cbs.cnrs.fr

## **Contents**

1. Introduction
  2. G protein-coupled receptors
    - 2.1 GPCRs : general trends and function
    - 2.2 Structural organization of GPCRs and activation process
    - 2.3 Extramembrane domains
      - 2.3.1 Intrinsic disorder
      - 2.3.2 Post-translational modifications
      - 2.3.3 Ligand binding and coupling to signaling partners
    - 2.4 Therapeutic implications
    - 2.5 Conclusions
  3. Nuclear receptors and their transcriptional coregulators
    - 3.1 Biological context
    - 3.2 The NTD/AF-1 of NRs and its interactions with transcriptional coregulators
    - 3.3 Transcriptional coregulators and their interaction with the CTD/AF-2 of NRs
    - 3.4 Importance of NR-NTDs in diseases and as pharmaceutical targets
    - 3.5 Liquid-Liquid Phase Separation (LLPS) in NRs
  4. Conclusions & Perspectives
- Acknowledgments  
References

## **Abstract**

G protein-coupled receptors (GPCRs) and Nuclear Receptors (NRs) are two signaling machineries that are involved in major physiological processes and, as a consequence, in a substantial number of diseases. Therefore, they actually represent two major targets for drugs with potential applications in almost all public health issues. Full exploitation of these targets for therapeutic purposes nevertheless requires opening original avenues in drug design, and this in turn implies a better understanding of the molecular mechanisms underlying their functioning. However, full comprehension of how these complex systems function and how they are deregulated in a physiopathological context is obscured by the fact that these proteins include a substantial number of disordered regions that are central to their mechanism of action but whose structural and functional properties are still largely unexplored. In this chapter, we describe how these intrinsically disordered regions (IDR) or proteins (IDP) intervene, control and finely modulate the thermodynamics of complexes involved in GPCR and NR regulation, which in turn triggers a multitude of cascade of events that are exquisitely orchestrated to ultimately control the biological output.

## **Keywords**

Intrinsically disordered proteins (IDPs), G protein-coupled receptors (GPCR), nuclear receptors (NRs), macromolecular complexes, signaling, G protein-coupled receptor kinase (GRK), arrestins, transcriptional coregulators, post-translational modifications (PTMs), ligands

# 1. Introduction

Why are Intrinsically Disordered Proteins (IDPs) so fascinating? They were discovered in the mid-90s and it is now well established that IDPs are very common in eukaryotes (40%)<sup>1</sup> where they confer complexity to these organisms.<sup>2,3</sup> Moreover, their inherent flexibility renders IDPs accessible to post-translational modifications (PTMs)<sup>4</sup> (phosphorylation,<sup>5</sup> ubiquitination,<sup>6</sup> palmitoylation, acetylation, sumoylation,<sup>7</sup> methylation). More than 300 PTMs can occur physiologically,<sup>8</sup> leading to the availability of more than 140 amino acids produced by these various modifications.<sup>9</sup> IDPs lack stable tertiary and secondary structure.<sup>10</sup> This is due to their specific amino acid sequences depleted in hydrophobic residues and enriched in polar and charged residues that are uncompensated; and that often have low complexity.<sup>11</sup> Their transient secondary structures are often conserved and linked to function.<sup>12</sup> Indeed, IDPs can adapt to many partners by changing their structure (promiscuity) and their ability to interact with high specificity and low affinity allow them to finely tune the thermodynamic properties of their complexes.<sup>13</sup>

With all these general biophysical features, it is not surprising that IDPs are linked to very specific functions such as signaling, transcription, etc.<sup>14-16</sup> As a consequence, their deregulation is at the origin of a multitude of diseases,<sup>17</sup> such as cancer,<sup>18,19</sup> neurodegeneration,<sup>20,21</sup> cardiovascular diseases<sup>22</sup> and diabetes<sup>23</sup>. In order to understand the role of disorder in function and disease, we have to characterize their structure and partner-recognition mechanisms. IDPs and their complexes are very difficult to characterize at the structural level due to their inherent flexibility.<sup>24</sup> The most suited way to address their study is by combining data from complementary techniques.<sup>25</sup> Indeed, Nuclear Magnetic Resonance (NMR) is the best-adapted technique to highlight the hot spots of IDPs in their free state.<sup>26,27</sup> Most of the times, these identified Short Linear Motifs (SLiMs)<sup>28</sup> or Molecular Recognition Features (MoRFs)<sup>29</sup> are evolutionary conserved<sup>30</sup> and are the functional part of IDPs.<sup>31</sup> These interacting regions identified in the complex, which usually fold upon binding,<sup>32,33</sup> can then be delimited by NMR, synthesized and submitted to co-crystallization trials. In parallel, Small Angle X-ray Scattering (SAXS) and Small Angle Neutron Scattering (SANS) can provide the overall envelope of the entire system.<sup>34,35</sup> Finally, all these data are generally integrated into computational tools to deliver accurate structural models.<sup>36,37</sup> These data can be complemented or validated by other biophysical methods such as FRET, interactions studies *in vitro* (fluorescence anisotropy, ITC, MST, etc), mutational analyses and cell-based assays allowing connecting detailed structural models with biological function. Such an integrative

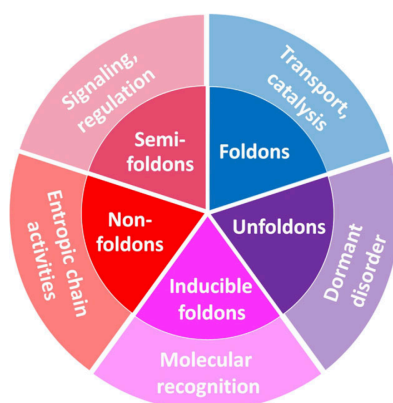


strategy where experimental information obtained from complementary techniques is integrated and simultaneously interpreted to derive a unified model of the biological system has already allowed the understanding of the structural bases of crucial biological processes involving IDPs.<sup>14,38,39</sup>

In this chapter, we will present two biological machineries involved in cell signaling for which structural disorder plays key roles: G Protein-Coupled Receptors (GPCRs) and Nuclear Receptors (NRs). GPCRs are the target of almost 40% of drugs on the market,<sup>40-42</sup> and NR ligands constitute 10-20% of the worldwide pharmaceutical market<sup>43</sup> and, as a consequence, it is essential to better understand their functioning to design more effective drugs without side effects. GPCRs and NRs are representative examples of the complementarity of ordered and disordered regions associated with distinct and complementary functions.<sup>15</sup> Both GPCRs and NRs (Figs. 2 and 5) share many structural features. They have a well-folded core domain whose high-resolution tridimensional structures have been solved by X-ray crystallography and cryo-EM, and unfolded domains that are not well characterized because they are missing from the electron densities (see below §2.1 and 3.1). For class A or (rhodopsin-class) GPCRs, the disordered domains are the N- and C-terminal domains plus extra- and intra-cellular loops. For NRs, the main disordered domains are the N-terminal Activation Domain (NTD), a linker between the DNA and ligand binding domains (DBD and LBD, respectively) and the H12 helices at the C-terminus of LBD. Both N-terminal domain of GPCRs and H12 of NRs undergo folding upon ligand binding processes that trigger different functional outputs. For GPCRs, this disorder to order transition triggers the transduction of signal via a transmembrane core reorganization, and for NRs it triggers the switch of the repressor complex by the activator one to initiate transcription of the target genes. NR coregulators are among the first proteins reported to be functional despite being mainly disordered.<sup>44,45</sup> Transcriptional coregulators are very large IDPs that include corepressors and coactivators, containing around 2,400 and 1,400 amino acids, respectively. Concerning the cytoplasmic C-terminal domain of GPCRs, they were discovered with the tridimensional structure of the beta-2 adrenergic receptor ( $\beta$ 2AR)<sup>46</sup>. This structure came out seven years after the first membrane protein solved by X-ray crystallography, the Rhodopsin.<sup>47</sup> Similarly as NTD of NRs, C-terminal domains of class A GPCRs are very variable in length and amino acid composition (from 30 to 325 residues), and are recognized as fully disordered.<sup>48</sup> Intrinsically disordered coregulators, NTDs of NRs and C-terminal domains of GPCRs are submitted to a multitude of PTMs. Together with these PTM-sites, these highly flexible IDPs and IDRs contain SLiMs and/or MoRFs regions (for a review, see <sup>49</sup> for GPCRs and <sup>50</sup> for NRs). This

combination increases the degree of diversity of binding sites, thus making the modulation of their interactions with their physiological partners more multifaceted, which ultimately allows these complexes to play their role of super regulators. These regions were assigned in the literature (described in <sup>51</sup> and illustrated in Fig. 1) as functional semi-foldons (i.e. partially folded protein segments with transient residual structure) and inducible foldons (i.e. segments that are able to fold upon interacting with specific binding partners). Moreover, the functions of non-foldons (i.e. protein segment with intrinsically disordered nature), which are found in this kind of regulating systems, are very important because of their obviously entropic chain activities. Indeed, this activity is remarkably well exemplified by coregulators that encompass multiple binding sites that increase the cooperativity in their interaction with NRs. This conformational variability complements the dynamical assembly of foldons (i.e. ordered segments or independently foldable regions) represented by the central core of GPCR and NRs (Figs. 2 and 5), which hold complementary function (transduction of signal for GPCRs and DNA and ligand recognition for NRs). The last piece of the mosaic of the protein structure-function space (Fig. 1) is illustrated by the functional unfoldons (a part of the ordered protein that has to undergo order-to-disorder transition to make protein active). This activity is exemplified in GPCRs interaction with arrestin, where the flexible cytosolic domain of GPCRs releases the C-terminal tail of arrestin that in turn will expose some specific interaction regions that will trigger different downstream effectors in the cell for specific functional responses.

The two complex machineries presented in this review, GPCRs and NRs, are illustrative of the structure-function continuum that undergoes IDPs and IDRs in these signaling and regulation systems to elegantly control cell signaling by orchestrating a combination of modular assembly of different type of foldons<sup>51</sup>.



**Fig. 1** Schematic representation of the mosaic nature of the protein structure–function space, where differently folded segments of a protein (foldons, semi-foldons, non-foldons, inducible foldons, and unfoldons) might possess different biological functions. Reproduced with permission from <sup>51</sup>.

## 2. G protein-coupled receptors

### 2.1 GPCRs: general trends and function

G protein-coupled receptors (GPCRs) are one of the largest class of membrane receptors. They play a central role in most intercellular communication processes, including vision, metabolism and neurological communication.<sup>52,53</sup> As such, they represent a major target family for drugs, with potential applications in many clinical spheres including cancer, metabolic, inflammatory and neurological disorders or hypertension, among many others.<sup>41</sup>

GPCRs transmit information through a mechanism that primarily involves activation by an external agonist compound and subsequent coupling to intracellular signaling partners. Historically, the first of these partners to be identified were G proteins, from which GPCRs got their name. G proteins are  $\alpha\beta\gamma$  heterotrimers composed of diverse combinations of the twenty-one  $\alpha$ , five  $\beta$  and twelve  $\gamma$  subunit isoforms.<sup>54</sup> Coupling to GPCRs triggers GDP-to-GTP exchange in the  $\alpha$  subunit, resulting in the activation of the G protein. Activated G proteins in turn activate their downstream effectors to trigger signaling.<sup>54</sup> Besides G proteins, GPCRs can trigger signaling through an extensive set of cytoplasmic partners. Among them are arrestins. There are four isoforms of arrestins, including the two visual ones (arrestin 1 and 4), which are involved in both signaling and receptor desensitization.<sup>55</sup> In the classical paradigm, activated GPCRs are first recognized by G protein-coupled Receptor Kinases (GRKs), leading to receptor phosphorylation at specific sites.<sup>56</sup> The phosphorylated receptor becomes competent for interacting with arrestins. Arrestin binding then triggers recruitment of the internalization machinery to the activated GPCR, leading either to signal termination or to a second wave of signaling through internalized receptors.<sup>57</sup> In addition, arrestin recruitment to the phosphorylated receptor can initiate an alternate intracellular signal.<sup>55,58</sup> Whether this additional signal is correlated or not to G protein activation is still under debate, however.<sup>59</sup> Finally, in addition to G proteins, GRKs and arrestins, GPCRs can interact with a full network of cytoplasmic proteins, including scaffolding ones, that further regulate the transmission of the signal. This interaction network has been extensively described in particular for the serotonergic system in the brain.<sup>60</sup>

### 2.2 Structural organization of GPCRs and activation process

The GPCR family is composed of almost 800 different members separated into five different classes (Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin).<sup>61</sup> Although these classes differ in their structure and function, they all share common structural elements. Indeed, GPCRs are primarily composed of two separated domains, the intra- and

extramembrane segments. For all GPCRs, the membrane domain is composed of seven transmembrane helices labeled TM1 to TM7.<sup>61</sup> The extramembrane region differs, depending on the class of GPCR considered. In some classes, *i.e.* glutamate and secretin, these regions include large extracellular domains that participate to ligand binding and receptor activation. In others, in particular the rhodopsin-like, the extramembrane regions essentially consist in relatively short N- and C-extensions, including a cytoplasmic H8 helix, and extra- and intracellular loops connecting the transmembrane helices. These segments in the rhodopsin class of GPCRs will be the main focus of the present chapter, as the large N-terminal parts of the glutamate and secretin receptors essentially display the characteristics of folded protein domains. In contrast to the transmembrane sequences that include highly conserved motifs, the extracellular and cytoplasmic regions of GPCRs are far more divergent in sequence.<sup>62</sup>

The three-dimensional organization of the membrane domain of GPCRs will be described here briefly, as it is not the focus of the present chapter. Starting from the first crystal structures of rhodopsin<sup>47</sup> and the  $\beta_2$ -adrenergic receptor,<sup>46,63</sup> the structural arrangement of the membrane domain of GPCRs has been progressively illuminated during the last decade through a series of crystal structures of receptors loaded with either agonists or antagonists (for a review, see <sup>64</sup>). These structures provided a detailed description of the general architecture of the helical domains, of the orthosteric and/or allosteric ligand binding sites, as well as the conformational changes associated with agonist binding. Structural analyses subsequently provided the arrangement of the complexes of receptors with their signaling partners. The first structure was obtained through X-ray crystallography with the  $\beta_2$ -adrenergic receptor associated to Gs.<sup>65</sup> The subsequent explosion in cryo-electron microscopy (cryoEM) approaches then provided us with a plethora of structures of receptors from different classes, *i.e.* rhodopsin (for a review, see <sup>66</sup>), secretin (for a review, see <sup>67</sup>), glutamate<sup>68</sup> and Smoothed<sup>69</sup>, in complex with different G protein partners (G11, Gs, Gi/o). More recently, cryoEM structures were obtained for receptors in complex with arrestins<sup>70-72</sup> or even with both G proteins and arrestins<sup>73</sup>. All these structures provide a detailed description of the three-dimensional arrangement of the complexes as well as the changes in receptor and signaling protein conformation associated to the activation process.

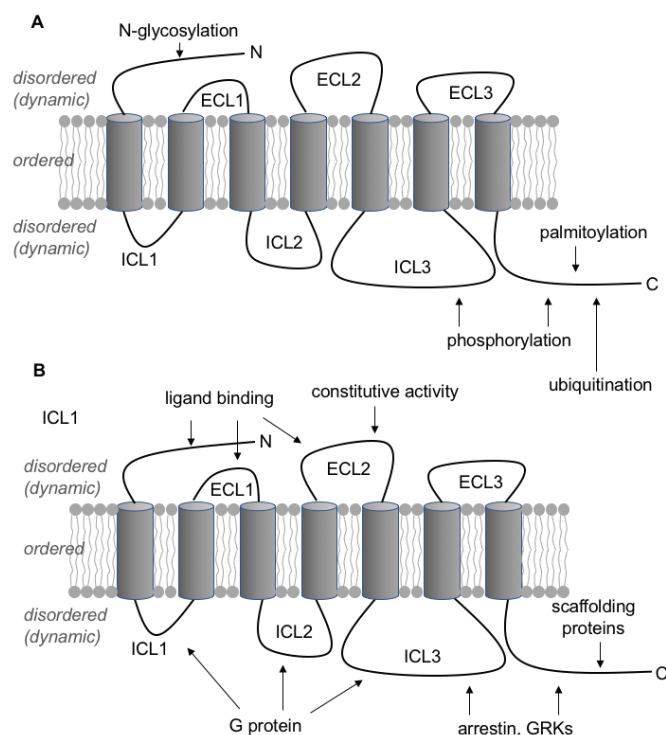
Besides the transmembrane domains, GPCRs include a significant proportion of residues that lie out of the lipid bilayer. These include the N- and C-terminal domains as well as in the intra- and extracellular loops. In contrast to our initial view, these regions are not to be considered as ancillary domains. Depending on the class of GPCR considered, they range from ca. 40% (rhodopsin-like receptors) to ca. 70% (glutamate) of the total receptor

residues.<sup>74</sup> However, so far, their structural arrangement remains obscure. Indeed, both their conformational dynamics and their modification for crystallization purposes (truncation of the N- and C-terminal regions, fusion of protein partners in the intracellular loops or N-terminus) precluded their observation in most of the GPCR crystal structures. This dynamic character also makes the use of modeling methods uncertain. In a few crystal- and cryoEM structures of the receptor:effector complexes, some of these regions are visible, however, in particular when their dynamics is restrained because of the direct interaction with the signaling partner (see below).

GPCRs transmit an information through concerted changes in their conformation. The three-dimensional structure of receptors in complex with antagonists, agonists and signaling proteins provided a detailed description of these changes. These include large amplitude motions in the TM helices as well as specific conserved microswitches. However, the diversity in signaling properties of GPCRs strongly suggests that this process can be hardly accounted for with the limited number of conformational states described in the crystal and cryoEM structures. Combination of computational and experimental approaches, including fluorescence and NMR, rather indicates that control and modulation of signaling involves a strong conformational plasticity GPCRs. The current model proposes that receptors explore complex conformational landscapes populated with multiple states, and that they adapt their shape in response to their environment through a series of conformational rearrangements that are central to the selective activation of downstream signaling partners.<sup>75,76</sup>

### **2.3 Extramembrane regions**

As state above, a significant part of GPCRs lies outside the membrane. These include the N- and C-terminus as well as the intra- and extracellular loops connecting the TM segments. In the case of rhodopsin-like receptors, they represent about 40% of the total receptor residues. Initially, these loop segments were essentially considered as necessary to maintain the architecture of the receptor central core. This is the case, as they likely constrain the respective arrangement of the transmembrane helices,<sup>77</sup> but not only. Indeed, in addition to their structural role, these regions are deeply involved in the functional properties of GPCRs and are subject to specific post-translational modifications that modulate GPCR function. As such these regions display some of the characteristics of intrinsically disordered regions (IDRs).



**Fig. 2 Schematic representation of class A GPCR structural and functional organization.** GPCRs are primarily organized in two kinds of domains, the ordered transmembrane helices and the dynamic extra- and intracellular loops and N- and C-terminal extensions. The general (but not restrictive) nature of PTMs (A) and binding properties (B) in the extra-membrane regions are indicated.

### 2.3.1 Intrinsic disorder

Using the FoldIndex program, Jaakola and co-workers<sup>62</sup> concluded that these regions show the characteristic patterns of IDRs. Accordingly, in the case of the dopamine receptor for instance, the occurrence of repeats in the region encoding the third intracellular loop cause this region to have a high disorder index.<sup>78</sup> Of importance, the amino-acid content in these unstructured regions is significantly different from that found in disordered proteins, with a particularly high content in basic (R, K) and histidine residues.<sup>62</sup> This specific composition is probably directly related to the functional properties of these regions (*e.g.* interaction with signaling partners and scaffolding proteins). Interestingly, independently of their intrinsic properties, the structural and dynamics features of these regions also appear to be intimately linked to the dynamics of the transmembrane domains. For instance, the structural arrangement of the C-terminal domain of the  $\beta_2$ -adrenergic<sup>48</sup> or vasopressin<sup>79</sup> receptors is dependent on the activation state triggered by pharmacologically distinct ligands. Specifically, biased ligands triggered a different arrangement of the C-terminal end of the receptor relative to the transmembrane domain. It is then tempting to speculate that the arrangement of the C-terminal domain occurs as in concerted movement with the dynamics of the TM helices, ultimately modulating the interaction with cognate cytoplasmic proteins.

### 2.3.2 *Post-translational modifications*

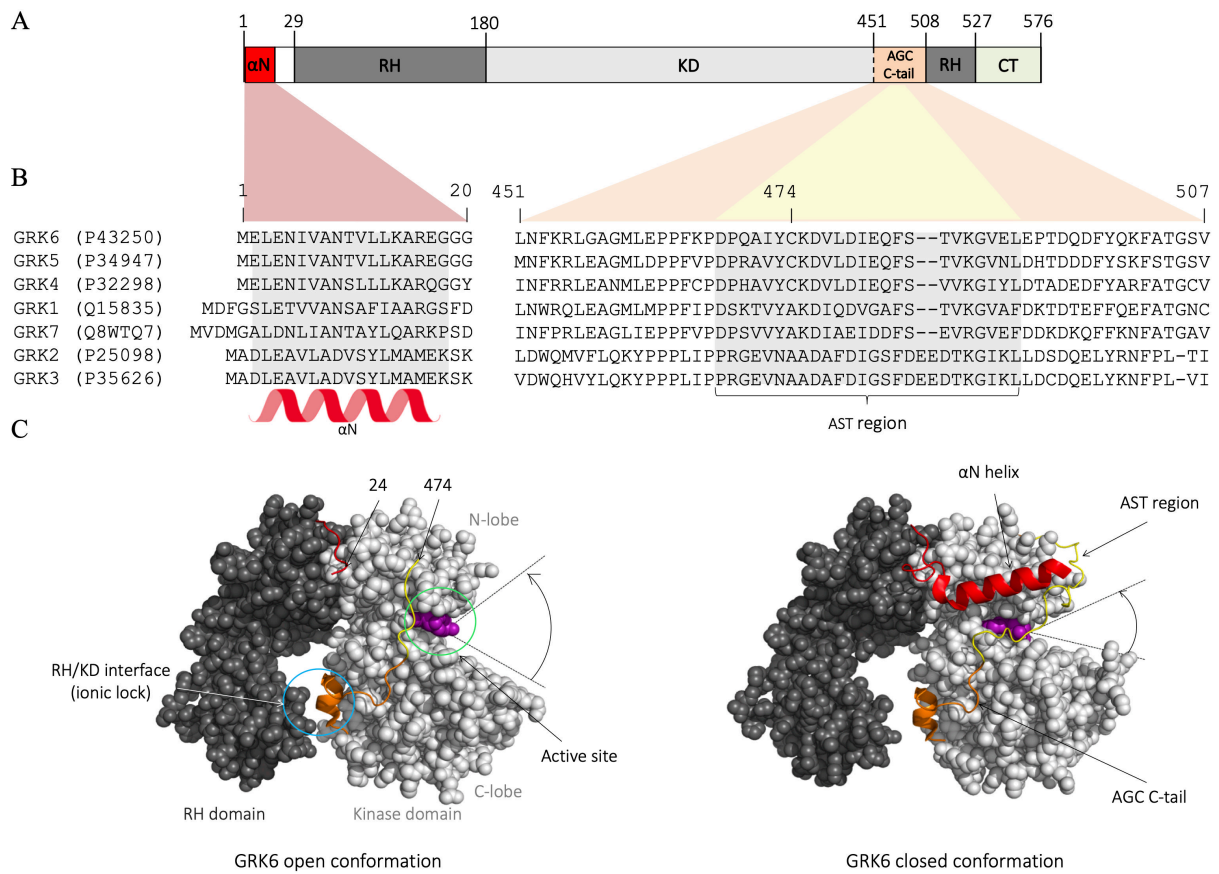
GPCRs undergo a full series of post-translational modifications (PTM) that primarily involve the unstructured extra-membrane domains, in particular the intracellular loops and C-terminal domain. Initially considered as an ancillary process, PTM now appear as central to GPCR functioning and regulation. Accordingly, dysregulation of the PTM machinery has been linked to pathological conditions, *e.g.* nephrogenic diabetes insipidus<sup>80</sup> or obesity<sup>81</sup>. The major GPCR modifications include glycosylation, ubiquitination, lipidation or phosphorylation, in addition to other far less described, more limited modifications (*e.g.* tyrosine sulfation). However, the impact of these modifications on the structure of GPCRs is largely unexplored, as most modification sites were removed in many structural studies, in particular in crystallization assays, to increase the homogeneity of receptor preparations.

GPCR glycosylation has been recently reviewed in details.<sup>82</sup> Briefly, glycosylation, and in particular N-glycosylation, essentially affects the N-terminal region of GPCRs and is primarily involved in their folding and transport to the cell surface.<sup>83-87</sup> As such, it may indirectly affect signaling by modulating the number of receptors present at the cell surface. In addition, glycosylation can have a direct impact on receptor functioning. For instance, activation of the  $\beta_2$ -adrenergic receptor by pathogens has been shown to involve traction forces on the receptor N-terminus *via* N-glycan chains.<sup>88</sup> GPCRs can also be ubiquitinated, and ubiquitination has been proposed to regulate GPCR signaling through degradation of cell-surface receptors.<sup>89</sup> In addition, this modification could also directly impact on signaling, as demonstrated for instance in the case of CXCR4-dependent MAPK activation.<sup>90</sup> Lipidation of GPCRs essentially involves addition of a palmitoyl chain on a conserved cysteine of the receptor C-terminus. This results in the anchoring of this region to the membrane, creating an additional intracellular loop. As such, it modulates the geometrical features of this region and, as a consequence, the receptor signaling and internalization properties.<sup>91</sup>

Phosphorylation actually appears as one of the most important PTM in the context of GPCR-dependent signal regulation. This has pushed forward, during the last decade, the development of novel analytical techniques and strategies, based in particular on mass spectrometry and specific antibodies, to identify residues subjected to phosphorylation under a variety of conditions.<sup>92</sup> Phosphorylation occurs in response to agonist-mediated receptor activation and is carried out by different classes of kinases. These essentially include second-messenger dependent protein kinases (PKA, PKC), and GRKs. Both phosphorylate specific Ser and Thr residues on the extra-membrane, intracellular regions of GPCRs. We will put the



focus here on GRKs, as they appear as a class of kinases that phosphorylate GPCRs specifically.



**Fig. 3 Importance of the conserved  $\alpha$ N domain and AGC C-tail in GRK activation.** (A) GRKs share a common structural organization. Residue numbers correspond to those of human GRK6.<sup>108</sup> The N-terminal domain ( $\alpha$ N, red) is followed by a regulator of G-protein signaling homology domain (RH, dark grey) surrounding a kinase domain (KD, light gray) and its AGC C-tail (orange). The C-terminal domain (CT, green) contains elements for membrane targeting specific to each subfamily.<sup>56</sup> (B) BLASTp **alignment and conservation** of the  $\alpha$ N and the AGC C-tail domains of human GRK<sup>100,105</sup> (sequence numbering based on GRK6 and Uniprot codes of human GRKs are in brackets). The  $\alpha$ N helix and the AST region are shown below the alignment. (C) Structures of GRK6 in open (PDB 2ACX, on the left)<sup>108</sup> and closed (PDB 3NYN, on the right)<sup>102</sup> conformations. The  $\alpha$ N domain and the AGC C-tail are represented in cartoon, the numbering on the left figure correspond to the end and the beginning of segments 1-23 and 475-491 which are absent of the electron density. C-terminal domains (CT) are not represented. The active site (green circle) contains adenylyl-imidodiphosphate (AMP-PNP, ATP analogue) and sangivamycin (nucleoside analogue, kinase inhibitor) (in purple) in open and closed conformation, respectively and its closure is represented by the angle between the N- and C-lobes of the kinase domain. The RH/KD interface containing the ionic lock (blue circle).



GRK is a family composed of seven members (1 to 7) that are subdivided into three subfamilies (GRK1/7, GRK2/3 and GRK4/5/6).<sup>93</sup> They share a common structural organization (Fig. 3A) composed of a specific and conserved N-Terminal domain ( $\alpha$ N), a discontinuous Regulator of G-protein signaling Homology domain (RH), a kinase domain (KD) belonging to the family of AGC kinases, and a C-terminal domain (CT) containing specific structural elements to each subfamily that is responsible for their membrane targeting.<sup>56</sup> GRKs require induced conformational rearrangements to become active (Fig. 3C). In the inactive conformation of GRKs the two N- and C-lobes of the kinase domain are splayed apart ('open conformation') compared to active AGC kinase structures ('closed conformation'),<sup>94,95</sup> and is stabilized by the interaction between the RH and the kinase domains.<sup>96,97</sup> The activation of GRKs is promoted by docking to activated GPCRs.<sup>98,99</sup> Indeed, in the model of GRK5/ $\beta$ 2AR complex,<sup>97</sup> the binding of  $\beta$ 2AR induce GRK5 conformational changes caused by the disruption of a transient 'ionic lock' between the RH and the kinase domains, and by the closure of the catalytic domain ('closed conformation'). The stabilization of this active state of GRKs involves two conserved elements (Fig. 3B): the C-tail extension of the kinase domain (AGC C-tail) and the  $\alpha$ N domain. During the activation and thus the closure of the kinase domain, the central part of the AGC C-tail (called the active site tether (AST)) and the  $\alpha$ N domain become more ordered.<sup>100-102</sup> Indeed, the  $\alpha$ N domain that folds in an  $\alpha$ N-helix and the AGC C-tail bring together the N- and C-lobes of the KD, stabilizing it in its 'closed' and active conformation (Fig. 3C).<sup>96,101,102</sup> We see here again, the functional importance of the disorder to order transition for the kinase activation and receptor phosphorylation.<sup>103-107</sup>

Phosphorylation of GPCRs essentially occurs in their third intracellular loop and C-terminal region, two regions that are central to the coupling to signaling partners. It is required for recruitment of arrestins, which ultimately leads to either receptor internalization or activation of alternate signaling pathways. Importantly, besides modulating the interaction with signaling partners, phosphorylation of the GPCR C-terminal tail could also affect the way it interacts with the membrane, and therefore its structural and dynamics features.<sup>109</sup> Even if it was initially considered that the most important feature for arrestin recruitment was the number of the phosphate residues and their spacing in the C-terminal domain,<sup>110</sup> GPCR phosphorylation appears to be a more complex process. First, the different GRK isoforms may regulate distinct, specific physiological functions of GPCRs.<sup>111</sup> In addition, these proteins appear to be able to specifically interact with different conformational states of the receptors. For instance, GRK2 can phosphorylate the D2 dopamine receptor without G protein

activation.<sup>112</sup> In this context, an emerging model is that of the “phospho-barcode” that states that distinct receptor conformations trigger interaction with different GRKs that in turn phosphorylate the receptor at distinct sites, establishing a “barcode” that ultimately dictates the functional outcome. This model was established by R. Lefkowitz and co-workers for the  $\beta_2$ -adrenergic receptor.<sup>111</sup> It has since then been extended to a full series of GPCRs including GPR120<sup>113</sup> and the ghrelin receptor<sup>114</sup>.

### **2.3.3 Ligand binding and coupling to signaling partners**

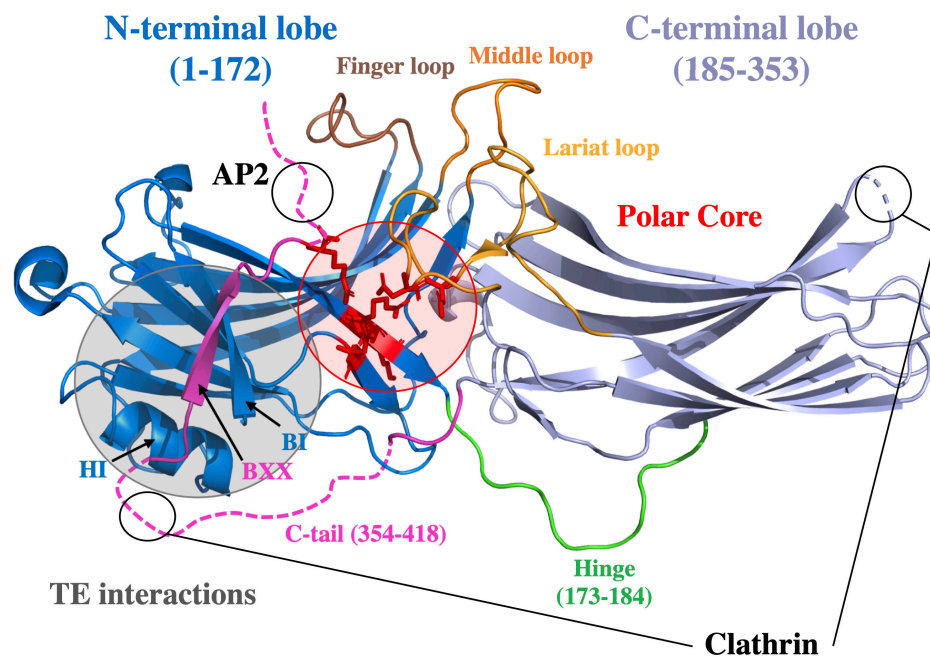
As stated above, besides being the site for PTMs, in particular phosphorylation, the extramembrane regions of GPCRs display one of the main characteristics of intrinsically disordered regions, i.e. interaction with cognate protein partners. This is the case, at least to some extent, for the extracellular regions. First, the extracellular loops can regulate the constitutive, ligand-independent, activity of some GPCRs such as the ghrelin receptor<sup>115</sup> and the orphan receptor GPR52<sup>116</sup>. In the latter case, the second extracellular loop directly plays the role of an intrinsic agonist for this otherwise orphan receptor. In addition, these regions can participate to ligand recognition and binding. This can occur through a direct interaction with the ligand, as demonstrated in particular for the chemokine<sup>117–119</sup> or adenosine<sup>120</sup> receptors. They can also play the role of a funnel delineating the binding pathway of ligands, as shown for the leukotriene<sup>121</sup> and ghrelin<sup>122</sup> receptors. Finally, the extracellular loops, in particular ECL2, can also affect indirectly the binding parameters for ligands by playing the role of a lid that precludes dissociation of compounds from their binding site within the transmembrane domains.<sup>123</sup>

Such a role as a “binding platform” for associated partners is even more relevant for the intracellular parts of GPCRs. First, the C-terminus appears to serve as a platform for assembling a full series of scaffolding proteins, in particular the so-called GPCR Interacting Proteins (GIPs), through the recognition of small motifs on the C-terminal regions of GPCRs. This leads to the formation of large functional protein:protein complexes involved in signaling. The role of these assemblies is particularly well-documented in the physiology and pathology of the central nervous system.<sup>124</sup> The intracellular loops of GPCRs are also involved in the interaction with G proteins. Initially demonstrated through mutagenesis and chimeric exchange experiments,<sup>125</sup> this has been further directly visualized in the cryoEM structures of the receptor:G protein complexes. Although the ICL3 loop is poorly resolved in some structures, direct contacts between the three intracellular loops and the G protein  $\alpha$  and  $\beta$  subunits have been observed in essentially all these structures, as reviewed recently.<sup>126</sup> In

the same way, hydrogen/deuterium exchange mass spectrometry experiments with the isolated  $\beta_2$ -adrenergic receptor and Gs indicated a specific involvement of the receptor intracellular loops in the interaction with the G protein, with specific residues in the second loop playing a central role in the stabilization of the complex between the  $\beta_2$ -adrenergic receptor and nucleotide free Gs.<sup>127</sup> The intracellular loops and C-terminal end are also major players in the interaction with arrestins. As stated above, the GPCR C-terminus is selectively phosphorylated by GRKs, and this PTM is the starting point for recruitment of arrestin and further activation of this protein.

Arrestins are briefly described here as, in contrast to G proteins, their structural features and conformational changes upon coupling to GPCRs may share some characteristics with IDPs. Arrestins are a family of increasing importance in GPCR signaling. This family is composed of four members (arrestin1<sup>128,129</sup>, arrestin2<sup>130</sup>, arrestin3<sup>131</sup> and arrestin4<sup>132</sup>) that share a similar 3D structure (Fig. 4). Arrestins encompass a N- and C-terminal lobe around a polar core composed of charged amino acids: Asp26, Arg169, Lys170, Asp290, Asp297, and Arg393 (numbering based on arrestin2). Arrestin inactive conformation is maintained by a salt bridge between Arg169 and Asp290,<sup>133</sup> and a ‘three-element’ (TE) interaction constituted of an  $\alpha$  helix (H1), a  $\beta$  strand (BI) in the N-terminal lobe and a  $\beta$  strand (BXX) in the C-tail of arrestin (Fig. 4).<sup>134</sup> The formation of a  $\beta$ -sheet between the two strands of the TE lays this inactive state in a ‘closed’ conformation. Upon binding of the phosphorylated GPCR C-terminal domain, arrestin undergoes a conformational change to its active state<sup>135–139</sup>: (1) a twisting of the N and C-lobes ( $\sim 20^\circ$ ); (2) a movement of the finger loop (that can be inserted in the core of the receptor, see below) and of the middle loop; (3) a release of arrestin C-tail; (4) a disruption of the polar core and (5) a conformational change of H1, BI and BII. The release of the flexible arrestin C-tail during its transition from an inactive to an active form, allows an exposition of binding sites for its partner interactions that trigger downstream signaling pathways.<sup>140</sup> More precisely, arrestin C-tail regions, which are present in the electron density of the inactive state X-ray structure (residues 354-357 and 383-393) (Fig. 4), are conversely absent in the active form X-ray structures (complex of arrestin2 cross-linked with a synthetic phosphorylated peptides mimicking the C-terminal domain of V2<sup>137</sup> or GPCR-arrestin complexes<sup>71–73,141</sup>). Indeed, in this active ‘open’ state, the disruption of TE interactions and the release of the C-tail bring some disorder and flexibility in the complex. The aforementioned exposed arrestin C-tail’s binding sites allow the interaction with downstream effectors such as clathrin and AP2 (Fig. 4), which are important for the

internalization of GPCR.<sup>140,142</sup> These binding sites are conserved motifs, where the AP2 binding site is <sup>391</sup>FARQRLK<sup>397</sup>,<sup>140,143</sup> and the clathrin binding site motif is LøXø(D/E),<sup>142</sup> (where ø corresponds to a hydrophobic residue and X to any polar residue). A second clathrin binding motif (L/I)<sub>2</sub>GXL has been found in a flexible loop of the C-terminal lobe (Fig. 4).<sup>142,144</sup> It was suggested that this second clathrin binding site, which is overlapped with a membrane binding site, would enhance the stability of the interaction between the receptor and its partner arrestin.<sup>144</sup>



**Fig. 4 Common structural organization of arrestin family** (PDB:1G4M/X-ray structure of arrestin2 inactive state of <sup>130</sup>). Arrestin family shares similar elongated shape and present two concave lobes mostly composed of antiparallel  $\beta$ -sheets: a N-terminal lobe (blue; residues 1-172) and a C-terminal lobe (light blue, residues 185-353), linked by a hinge in green (residues 173-184); flexible loops: finger loop (brown), middle loop (orange) and lariat loop (light orange). The polar core is highlighted by a red circle, mostly composed of charged amino-acid (in red, main and side chain of polar core residues) and the three-elements (TE) interaction highlighted by a grey circle involving  $\beta$ -strand 1 and 20 (BI and BXX) and  $\alpha$ -helix 1 (HI). Arrestin C-tail (pink, residues 354-418) encompasses  $\beta$ -strand 20 (BXX) (which interact with the N-terminal lobe in the ‘closed’ state). Concerning non-visual arrestin isoform, its C-tail contains AP2 binding site (<sup>391</sup>FARQRLK<sup>397</sup>; not present in the structure) and one clathrin binding site (<sup>376</sup>LIELD<sup>380</sup>; not present in the structure), the second clathrin binding site is localized in a flexible loop of C-terminal lobe (black circles).<sup>142</sup>

With regard to regions of GPCRs involved in the interaction with arrestins, in the recent cryoEM structures of the neurotensin receptor in complex with arrestin, the receptor C-terminus binds the N-terminal lobe of arrestin, a region that contains all the basic amino-acids that interact with phosphorylated residues of the receptor.<sup>71</sup> In the same way, in the complex between the muscarinic receptor and arrestin, the arrestin N-lobe engages the phosphorylated V2R C-terminal domain of the chimeric receptor, with critical phosphorylated residues of the

receptor in interaction with basic amino-acids of arrestin.<sup>72</sup> These features are fully consistent with the previous crystal structure of arrestin associated to a phosphorylated peptide from the vasopressin V2 receptor C-terminal domain.<sup>137</sup> In this structure also, the peptide phosphates directly interact with arrestin through charge-charge interactions. Of importance, in these cryo-EM structures, not only the C-terminus of the receptors is in interaction with arrestin, but also the intracellular loops participate to complex formation. In the case of the neurotensin receptor, arrestin engages the central core and third intracellular loop of the receptor in addition to its C-terminus.<sup>71</sup> In the same way, arrestin interdomain loops engage the receptor core and its second intracellular loop in the complex with the muscarinic receptor.<sup>72</sup> Such an interaction pattern is fully consistent with the model where arrestins were proposed to interact with GPCRs in a “two-site” manner, with a first binding site corresponding to phosphorylated C-terminus of the receptors and the other to the GPCR central core, including the intracellular loops.<sup>73,145</sup> Importantly, each of these interactions could be associated to a distinct functional outcome, thus expanding the repertoire for modulating GPCR-mediated signaling.

## **2.4 Therapeutic implications**

So far, most if not all therapeutic interventions targeting GPCR involved the ligand binding sites, either orthosteric or allosteric. However, the unstructured parts of GPCRs can also be responsible for disfunction of the signaling systems, as evidenced by the fact that several diseases are associated to mutations in the extra- or intracellular, unstructured parts of the receptors.<sup>146</sup> It cannot be excluded that this effect is due to the impact of such mutations on the three-dimensional arrangement of the receptor, *e.g.* by affecting the relative arrangement of the TM helices. However, these mutations could also alter the interaction pattern with specific signaling partners, as these are all regions that are responsible for protein:protein contacts in the signaling complexes. The extra-membrane regions of GPCRs are far less conserved than the transmembrane domains. Hence, this could pave the way for more specific interventions affecting protein:protein contacts in the signaling complexes. However, so far, if one excepts the pepducin approach that is based on peptides mimicking to intracellular loops of GPCRs and has been proposed for treating some physiological dysfunctions,<sup>147,148</sup> the regions lying outside the membrane have been largely unused in terms of therapeutical applications. Hence, protein:protein interaction (PPI) inhibitor strategies aimed at exploiting the structural and sequence diversity of these regions to regulate the composition and functioning of the signaling complexes could constitute an original, alternate, avenue in drug discovery.

## 2.5 Conclusion

In addition to many GPCR ligands where conformational disorder is an intrinsic feature of the binding process,<sup>122,149,150</sup> receptors themselves include a substantial number of their residues outside the membrane that present many of the characteristic of intrinsically disordered regions (IDRs). These include lack of defined structure and high conformational dynamics, interaction with multiple protein components of the signaling complexes, and post-translational modifications that specifically regulate these interactions. As such, they represent major players in the signaling processes. However, so far, their structural and functional properties are largely unexplored, in particular because of these conformational dynamics and labile structural and functional elements. Hence, future research in molecular pharmacology is also to be directed to illuminate our present understanding of the structural and functional relevance of GPCR loops and N- and C-terminal extensions in the context of signaling.

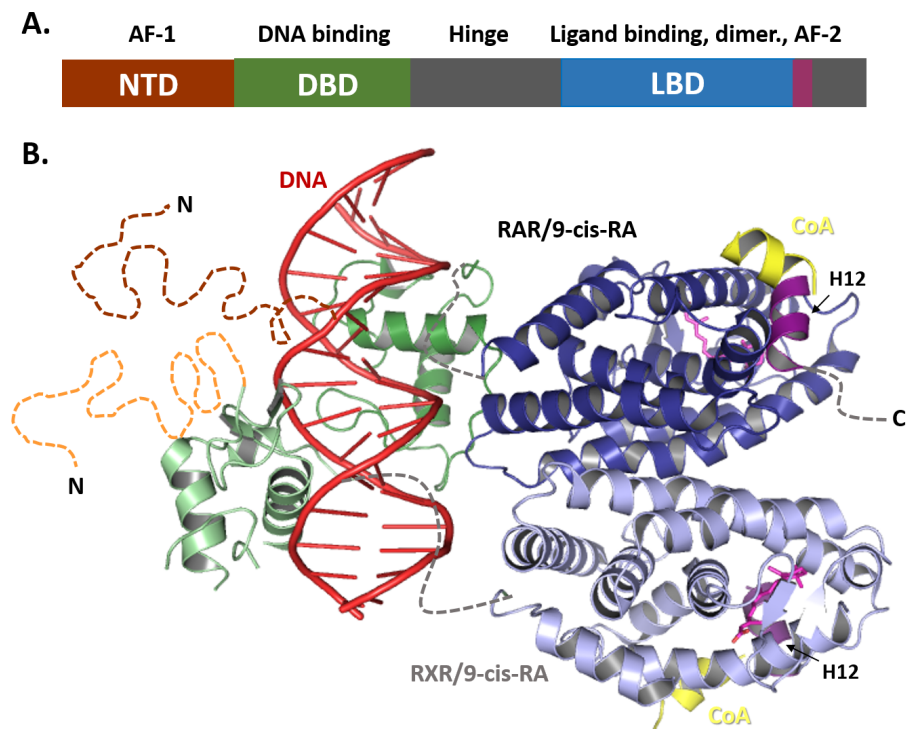
## 3. Nuclear receptors and their transcriptional coregulators

### 3.1 Biological context

Human nuclear receptors (NRs) constitute a family of 48 transcription factors that plays an essential role in regulating the expression of a multitude of genes that control most of fundamental biological processes such as development, organ homeostasis, metabolism, immune function, or reproduction.<sup>151–153</sup> This regulatory capacity of NRs occurs through their ability to recognize specific DNA sequences promoters of their target genes and their relationships with the general transcription machinery. NRs respond directly to a large variety of hormonal and metabolic substances that are hydrophobic, lipid soluble, and of small size (*e.g.* retinoic acid, estradiol, thyroid hormone, vitamin D3, etc.). They are also the target for numerous post-translational modifications including phosphorylation, SUMOylation, and acetylation.<sup>154</sup> They have a conserved modular structure with a variable unstructured N-terminal domain (NTD) that hosts the activation function 1 (AF-1), a well conserved zinc-finger DNA-binding domain (DBD) and a C-terminal moderately conserved ligand-binding domain (LBD) (Fig. 5). This multifunctional C-terminal domain (CTD) is responsible for ligand binding or dimerization and contains a ligand-dependent activation function (AF-2), which corresponds to transcriptional coregulator interaction surfaces that can be modulated by natural, pharmacological or environmental ligands.<sup>155–157</sup> The LBD also contributes to the modulation of the N-terminal AF-1 through interdomain crosstalk so that both NTD/AF-1 and



CTD/AF-2 domains can recruit a range of coregulatory proteins and act individually or in a synergistic manner.<sup>158–161</sup> Most NRs bind to DNA as homodimers or as heterodimers with the retinoid X receptor (RXR), the obligate heterodimerization partner of more than one third of the NR family members (Fig. 5).



**Fig. 5 Common organization of nuclear receptors.** (A) Functional organization of NRs. The N-terminal domain (NTD) contains a ligand-independent activation function (AF-1). The DNA-binding domain (DBD) mediates sequence-specific DNA recognition. The ligand-binding domain (LBD) mediates ligand binding, dimerization and a ligand-dependent transactivation function (AF-2). The hinge region corresponds to a linker allowing the proper orientation of the DBD and LBD within DNA-bound dimers. (B) Crystal structure of RAR $\beta$ /RXR $\alpha$  heterodimer bound to DNA.<sup>162</sup> LBDs (dark and light blue) and DBDs (dark and light green) are represented as ribbons. Helix H12 (AF-2) of each monomer is highlighted in purple. The coactivator peptides interacting with the heterodimer are drawn in yellow and the agonist ligands (9-cis-RA) in each subunit are drawn as pink sticks. Dotted lines denote regions with unresolved structures.

The activity of NRs depends on a diverse group of proteins termed “coregulators” that affect the transcriptional machinery in different ways and more than 350 coregulators have been reported to date. NRs may act either as repressors or activators of gene transcription depending on their ligation status that in turn determines the ability of DNA-bound receptors to recruit so-called “corepressors” or “coactivators” to target gene promoters. Coactivators and corepressors are in fact general terms defining the components of high molecular weight coregulator complexes.<sup>163,164</sup> They correspond to large groups of proteins with a diverse range

of protein activities and enzymatic functions whose interactions with NRs and many other transcription factors provide a complex regulatory framework for controlling the transcription of target genes. Coactivator complexes contribute to the enhancement of transcription by acetylating histones, a process that plays an important role in the opening of chromatin during transcription activation, whereas corepressors display the opposite activity by recruiting histone deacetylases. Coactivator recruitment is usually ligand dependent, whereas corepressors interact in most cases with unliganded receptors. Coactivators can be subdivided into two groups. Secondary coactivators represent a subgroup of molecules that are constituents of multisubunit coactivator complexes and that also contribute to the enhancement of NR-mediated transcription, but that do not directly contact the NRs. Primary coactivators, such as those of the SRC-1/TIF-2/RAC3 (p160) family, mediate the interaction of coactivator complexes with NRs. CBP, p300, P/CAF, and some p160 coactivators themselves are reported to act as histone acetyltransferases (HATs).<sup>165,166</sup> The corepressors NCoR and SMRT have been shown to reside in, or recruit, large complexes with deacetylase activity that maintain histones in a deacetylated form, a status that is associated with silent regions of the genome.<sup>167,168</sup> Therefore, histone acetylation and deacetylation shuffle nucleosomal targets between a condensed and relaxed chromatin configuration, the latter being requisite for transcriptional activation.

In the absence of their cognate ligand, some NRs are located in the nucleus, bind to the DNA response elements of their target genes, and recruit corepressors, whereas others are located in the cytoplasm in an inactive complex with chaperones. Ligand binding induces major structural alterations of receptor LBDs leading to (1) destabilization of corepressor or chaperone interfaces, (2) exposure of nuclear localization signals allowing nuclear translocation and DNA binding of cytoplasmic receptors, and (3) recruitment of coactivators triggering gene transcription through chromatin remodeling and activation of the general transcription machinery. Both NRs and their coregulators have been shown to frequently contain large intrinsically disordered regions that mediate their mutual interaction and play a pivotal role in transcription regulation. These unstructured proteins or domains may be functional, undergoing transitions to more ordered states or folding into stable secondary or tertiary structures upon binding to DNA response elements and protein partners. Below we review recent studies that shed light on the mechanisms governing signaling by these flexible domains and how they can be exploited for the development of new endocrine-based therapies.



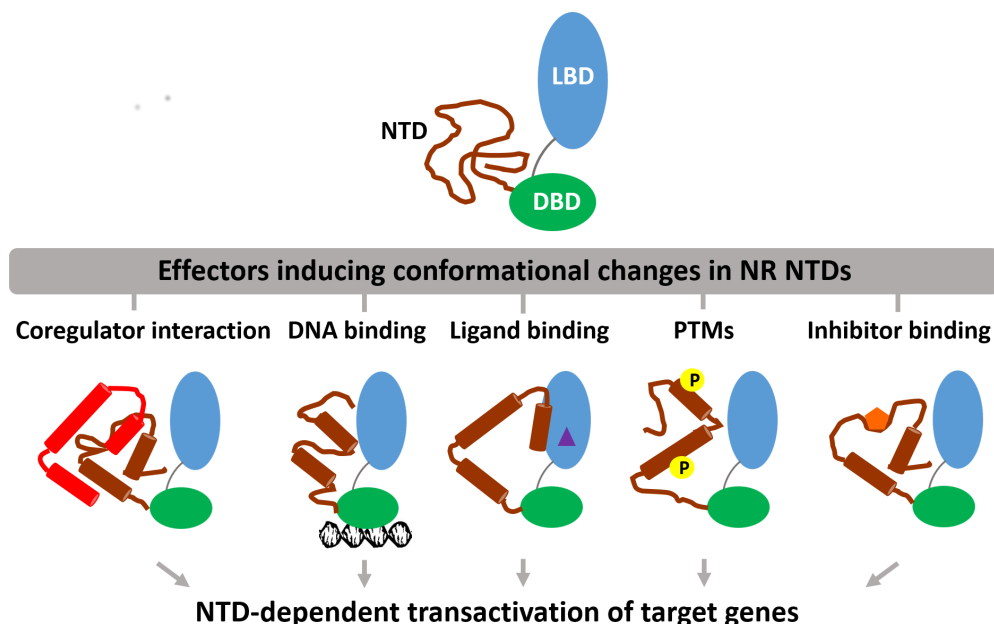
### 3.2 The NTD/AF-1 of NRs and its interactions with transcriptional coregulators

The NTD of NRs harbors an autonomous transcriptional activation function termed AF-1 that, when linked to a heterologous DNA-binding domain, can activate transcription in a constitutive manner, however, in the context of full-length receptors AF-1 is silent in the absence of ligand,<sup>169,170</sup> suggesting an allosteric communication between the N- and C-termini of NRs. When comparing NRs from different sub-families, the sequence of NTDs is highly variable, revealing a very weak evolutionary conservation. Their variable length, from a few residues like in the constitutive androstane receptor (CAR, 10 amino acids) or the vitamin D receptor (VDR, 24 amino acids) up to more than 500 residues in steroid hormone receptors such as the glucocorticoid receptor (GR, 420 amino acids), the androgen receptor (AR, 558 amino acids), the progesterone receptor (PR, 566 amino acids), and the mineralocorticoid receptor (MR, 602 amino acids), also suggests distinct transcriptional activities and biological roles.<sup>171</sup> Interestingly, the length of this domain has been positively correlated with the activity of the AF-1 for different members of the NR superfamily,<sup>172</sup> and logically, the NTD accounts for a major proportion of the total transcriptional activity of all steroid hormone receptors where AF-1 is the predominant activation domain over that of AF-2 lying in the C-terminal LBD.<sup>161,173–176</sup> Moreover, the NTD is subjected to alternative splicing and differential promoter usage so that the majority of NR isoforms differs in their N-terminal region. Notably, the production of translational isoforms differing only in the length of the disordered region was recently proposed as a mechanism by which NRs such as the GR can tune their transcriptional activity.<sup>177,178</sup> The NTD through its activation function AF-1 is involved in multiple protein-protein interactions that are critical for the modulation of the transcriptional activation of target genes in a cell-specific and promoter dependent manner.<sup>171,179–182</sup> As with the AF-2 domain, coactivators such as CREB-binding protein (CBP) or the p160 proteins SRC-1 and TIF-2 can also interact with and increase the transcriptional activity of the AF-1 domain.<sup>183–185</sup> However, unlike the LXXLL (X stands for any amino acid) binding motif for AF-2 interactions (see below for more details), no such conserved motif is known for the AF-1 binding. Note however that the NTD of AR contains the FXXLF motif (amino acids 23–27) which was shown to mediate a physical interaction between the AF-1 and the AF-2.<sup>172</sup> In the majority of NRs, the NTD is subjected to post-translational modifications, such as phosphorylation and sumoylation, that play essential roles in regulating the receptor transcriptional activities.<sup>186–188</sup>

In contrast to the DBD and the LBD bearing the C-terminal AF-2, and due to the intrinsically disordered nature of the NTD, the structural basis for AF-1 activity are not very well defined and only now begin to be revealed. Structural disorder is a key structural feature for transactivation domains in members of the NR superfamily<sup>50,176,178,187,189,190</sup> as well as many other transcription factors<sup>191</sup> so that their interactions with transcriptional coregulators are among the best characterized examples of complexes involving IDPs.<sup>192-194</sup> The structural plasticity of these domains is thought to allow (1) sampling the environment until appropriate binding partners are found and (2) specific interactions through induced conformational changes upon binding to targets, to (3) result in transient/low affinity complex formation and dynamic signaling regulation and (4) facilitate post-translational modifications.<sup>45,195</sup> Knowledge about NR NTDs was hampered by the technical challenges of characterizing IDPs.<sup>196,197</sup> Only limited structural information was available from circular dichroism (CD) spectra and chemical shifts from nuclear magnetic resonance (NMR) spectroscopy showing that isolated NTDs are largely unstructured with poorly resolved chemical shifts, and with a high random coil content and minimal secondary structural elements.<sup>182,198-200</sup> In the last decade, complementary biophysical techniques such as small-angle X-ray scattering (SAXS), fluorescence resonance energy transfer (FRET), and hydrogen exchange mass spectrometry (HDX-MS) in combination with computational approaches such as molecular dynamics (MD) simulations have confirmed the previous observations and allowed studies in the context of full-length receptors, thereby revealing unforeseen structural features of NR NTDs.

Such analyzes of full-length peroxisome proliferator-activated receptor gamma (PPAR $\gamma$ ),<sup>202</sup> PR<sup>189</sup> and retinoid X receptor alpha (RXR $\alpha$ )<sup>203,204</sup> have further revealed that the disordered nature observed for the isolated NTDs is also exhibited in the context of intact NRs, indicating that this property is not abrogated via interdomain interactions. Experiments performed in the presence of secondary structure ( $\alpha$ -helices and  $\beta$ -strands) stabilizing agents such as the osmolyte trimethylamine-N-oxide (TMAO) or the hydrophobic solvent trifluoroethanol (TFE),<sup>198,200,205-208</sup> interacting proteins,<sup>161,182,184,189,206,208-210</sup> or DNA<sup>50,199,211,212</sup> clearly demonstrated that the NTDs undergo partial folding under these circumstances and adopt some degree of  $\alpha$ -helical conformation. Furthermore, these conformational alterations were shown to have functional consequences since when residues involved in the formation of putative  $\alpha$ -helices were mutated, the ability of the AF-1 to form such secondary structures was remarkably reduced and the transactivation activity of this domain decreased, suggesting that the formation of  $\alpha$ -helices is important for AF-1 to mediate gene activation<sup>172,200,206,213,214</sup> (Fig. 6). This was best illustrated for the case of PR whose NTD contains a large content of

random coil and is capable of adopting secondary helical structure in the presence of TMAO. HDX-MS confirmed the high disorder in the NTD within the context of full-length PR, but also showed that it can adopt a more stable tertiary folding through a direct interaction with the TATA-binding protein (TBP).<sup>189</sup> Mapping experiments and point mutation analysis revealed that a central region of the NTD between residues 350 and 428 appeared to be sufficient for inducing secondary structure in the presence of TBP. These data suggested that short helical element(s) embedded within the larger disordered NTD may constitute a recognition site for TBP that can induce its folding upon binding. Beyond the formation of secondary structures upon interaction, it was further shown that while mostly disordered, the NTD of NRs might be more compact than expected for an IDP and exist in a premolten globule conformation,<sup>215,216</sup> and that long-range metastable contacts could explain this compactness. A recent study of the estrogen receptor alpha (ER $\alpha$ ) using SAXS, hydroxyl radical footprinting and MD simulations to generate a representative ensemble of structures identified a constitutive intra-domain interaction between two small 5-residue motifs centered around residues L33 and S118 in the NTD of this receptor.<sup>217</sup> In full agreement, the addition of a chemical denaturant induced a structural expansion as indicated by the significant increase of the radius of gyration of the ER $\alpha$  NTD monitored by SAXS.

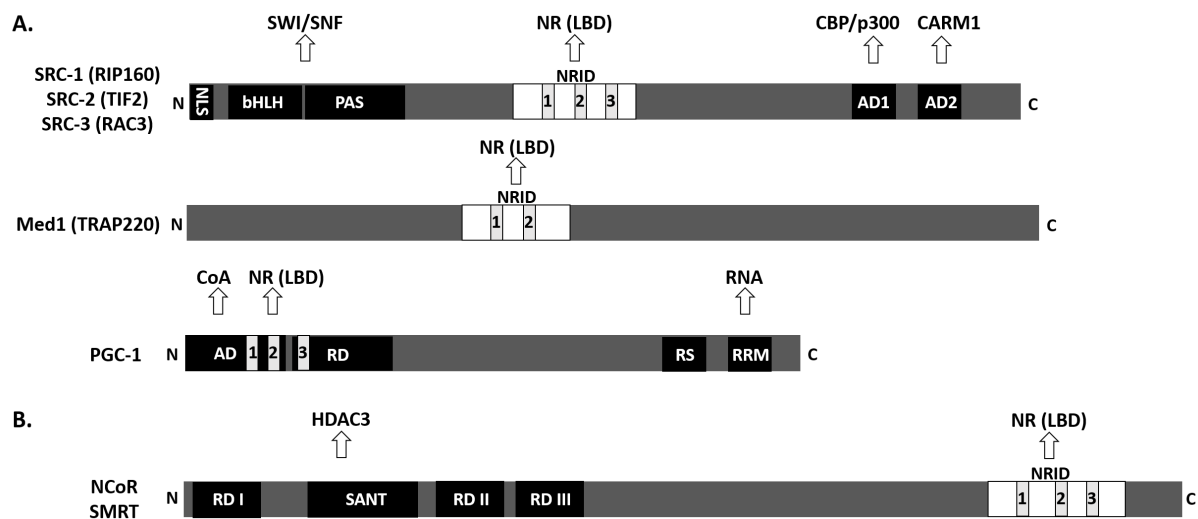


**Fig. 6 NR NTDs respond to various stimuli.** The non-conserved NTD of NRs is disordered but can acquire higher order structure in response to coregulator interaction, DNA binding, ligand binding to LBD, post-translational modifications (PTMs) and inhibitor binding. The NTD can adopt a multitude of conformations in response to these different signals, which ultimately translates into NTD-dependent transactivation of target genes specific of each receptor (adapted from <sup>201</sup>).

Post-translational modifications such as phosphorylation elicit diverse effects on the biological functions of IDPs by altering the energetics of their conformational landscape and by modulating interactions with other cellular components<sup>191,218</sup> through stabilization and/or induction of secondary structural elements<sup>219–221</sup> (Fig. 6). Accordingly, most of the functionally known sites subjected to post-translational modifications in NRs are located in their NTDs, as for example in GR whose five phosphorylatable serine residues are found in the NTD. In full agreement, the phosphorylation on S203, S211 and S226 of GR has been shown to induce conformational rearrangements leading to significantly enhanced AF-1-mediated transcriptional activity in a phosphorylation site-dependent manner.<sup>209,222,223</sup> Along this line, Peng and collaborators showed that mutation of S118, involved in a metastable contact with L33 in the NTD of ER $\alpha$  (see above), by the phosphomimetic aspartate residue (ER $\alpha$ S118D mutant) induced an expansion of the protein and increased the binding affinity to TBP, thus suggesting possible disruption of the long-range contact and a possible link between phosphorylation, conformational change of the AF-1 and transcription.<sup>217</sup> Interestingly, phosphorylation of ER $\alpha$  at S118 was previously shown to induce chemical shifts that propagate throughout the NTD and allows recruitment of the peptidyl prolyl *cis/trans* isomerase Pin1. Subsequently, the *cis-trans* isomerization of P119 caused by Pin1 was found to create additional local structural changes, as revealed by NMR, leading to the enhancement of ER $\alpha$  transcriptional function.<sup>180</sup> Another recent study on the NTD of AR revealed that its interaction with the subunit 1 (RAP74) of the general transcription factor TFIIF is mediated by the sequence 423-446 within the AF-1.<sup>224</sup> Interestingly, further stabilization of the  $\alpha$ -helical structure of the motif by using stapled peptides, together with phosphorylation of residue S424 dramatically increased the binding affinity to RAP74. Reminiscent of observations made on other IDP systems, phosphorylation may thus regulate NR functions by affecting the conformational dynamics of the NTD, leading to enhanced transcriptional activity.<sup>225,226</sup> In this respect, a mechanism for the interaction of disordered domains such as the NTD of NRs has been proposed according to which, in a first step, post-translational modifications such as phosphorylation would change their net charge thereby allowing weak polyelectrostatic interactions and, in a second step, hydrophobic interactions would direct the formation of a more stable complex where the activation domain would adopt a folded conformation for optimal interaction.<sup>227,228</sup>

### 3.3 Transcriptional coregulators and their interaction with the CTD/AF-2 of NRs

Like many proteins in signaling pathways, coregulators are large proteins belonging to the class of intrinsically disordered proteins that are functional despite a lack of stable three-dimensional structure. They act as platforms where multiple proteins attach to perform activities linked to gene transcription.<sup>229,230</sup> They interact with NRs through their Nuclear Receptor Interaction Domain (NRID) and recruit additional proteins to form either coactivator or corepressor complexes<sup>231</sup> (Fig. 7). Strikingly the intrinsic disorder of NRIDs is common to many other transcription factor interaction motifs located in these coregulator proteins.<sup>232,233</sup>



**Fig. 7 Schematic representation of main transcriptional coregulators in interaction with nuclear receptors.**

(A) SRC-1, SRC-2/TIF-2 and SRC-3/RAC3 share a common organization. They have a basic helix-loop-helix/Per-ARNT-Sim homologous domain (bHLH/PAS), and a nuclear receptor (NR) interacting domain (NRID) with three LXXLL motifs critical for ligand-dependent interaction with NRs. The two activation domains AD1 and AD2 are located in the C-terminus where AD1 is responsible for interaction with the general transcriptional machinery through CBP and p300, and AD2 interacts with the histone methyltransferases coactivator-associated arginine methyltransferase 1 (CARM1). MED1/TRAP220 contains two NR boxes in its central NRID. PGC-1 contains three LXXLL motifs responsible for specific interaction with NRs. Other domains are an activation domain (AD) involved in binding to other coactivators, a repression domain (RD), an arginine-serine-rich domain (RS) and a RNA recognition motif (RRM). (B) The two corepressors NCoR and SMRT have common functional domains: C-terminal CoRNR boxes 1, 2 and 3 included in the NRID; the three repression domains (RDI, RDII and RDIII) that harbor intrinsic repression activity when tethered to the DNA-binding domain of GAL4 and two SANT domains.

Most of the coactivator complexes recruited by NRs do so through a conserved motif with the sequence LXXLL (the so-called NR box).<sup>44</sup> These LXXLL motifs form amphipathic  $\alpha$ -helices that are responsible for direct association with NRs by recognizing a hydrophobic patch on the surface of their LBD, the structure of which is stabilized by the binding of agonist ligands. Shortly after the identification of the NR box, an analogous sequence motif (L/IXXI/VI or LXXXI/LXXXI/L) was identified in corepressor proteins.<sup>234,235</sup> This motif has been termed the CoRNR box and adopts a longer amphipathic helical conformation than the LXXLL motif. Crystal structures have demonstrated that peptides corresponding to NR or CoRNR boxes bind to the same hydrophobic groove on LBDs and therefore their recruitment is mutually exclusive.<sup>236–239</sup> The disordered NRID of coactivators and corepressors typically contains multiple NR and CoRNR boxes, respectively, connected by disordered linker regions (Fig. 7). These regions are thought to be important for NR interaction, NR selectivity, binding, and a full transcriptional response.<sup>240–242</sup> Secondary structure and disorder prediction combined with computational data and in-depth biophysical and structural characterization by Size Exclusion Chromatography - Multi-Angle Light Scattering (SEC-MALS), CD, NMR and SAXS of various coactivator and corepressor NRIDs confirmed the intrinsic disorder of these regions<sup>34,243,244</sup> (our unpublished results). In addition, these studies showed the existence of pre-formed secondary structure elements localized in evolutionary conserved regions corresponding to NR and CoRNR boxes<sup>34,243,244</sup> (our unpublished results). These types of preformed secondary structures are usually important determinants for molecular recognition in IDPs.<sup>245,246</sup> Of note, these transient secondary structures in the unbound states resemble the secondary structures in the complexes with NRs. In addition, partial structures have been observed in flanking regions of these motifs that could play a role in NR specificity<sup>237,247,248</sup> and transient intramolecular contacts could exist within these NRIDs.<sup>249</sup>

The fact that coregulators contain multiple NR interaction motifs in a relatively close proximity, most of which appearing to be functional in terms of NR binding *in vitro*, brings up the still unsolved question of whether this multiplicity reflects redundancy or confers some specificity to the interface. Different structural studies showed that NR and CoRNR boxes are not equivalent in terms of NR affinity and specificity, which provide diversity and selectivity in the spectrum of NR/coregulator interactions for NR-specific responses upon specific ligand binding. For example, the crystal structure of PPAR $\gamma$ -LBD with a PGC-1 construct bearing both NR2 and NR3 boxes showed only one LXXLL motif (exclusively NR2) bound to PPAR $\gamma$  and the remaining portion of PGC-1 disordered.<sup>213</sup> This NR2 motif is also the major PGC-1 binding site to ER and RXR.<sup>250,251</sup> On the contrary, another study showed that the

hepatocyte nuclear factor-4 alpha (HNF4 $\alpha$ ) might present a lack of strict selectivity in the interaction with the coactivator PGC-1 since the crystal structure of HNF4 $\alpha$ -LBD in complex with the NRID of PGC-1 comprising all three LXXLL motifs revealed that only one LXXLL motif is bound to the canonical binding pocket and does not correspond to a single sequence but to an averaged structure of the different LXXLL motifs.<sup>252</sup> It was also discovered that two NRs, the retinoic acid receptor alpha (RAR $\alpha$ ) and Rev-erb $\alpha$ , that exhibit strong basal repressive activity, interact strongly with CoRNR1 but very weakly with CoRNR2, thanks to the formation of an antiparallel  $\beta$ -sheet of N-terminal residues of CoRNR1 with an extended  $\beta$ -strand of RAR $\alpha$  and Rev-erb $\alpha$ .<sup>239,253</sup> In addition, the differences in binding affinity of the different coregulator motifs to NRs induce directionality in the interaction of coregulators with NR heterodimers. Both receptors usually contribute to the interaction of the heterodimer with the coregulator through different thermodynamic mechanisms and to a different extent. Using a large set of biophysical and computational methods to study the interaction of RAR $\alpha$ /RXR $\alpha$  heterodimer with the corepressor NCoR, we were able to show that although both NRs contribute to the interaction with the NRID in a cooperative manner, RAR $\alpha$  plays a dominant role over RXR $\alpha$ .<sup>249</sup> The HDX-MS analysis on the complex between the heterodimer VDR/RXR and the coactivator SRC-1 supported a model of synergistic binding of one molecule of SRC-1 spanning both NRs, facilitated in the presence of agonist ligands of each subunit.<sup>203</sup> A combined NMR and HDX-MS analysis showed that all three LXXLL motifs within the SRC-2 NRID are involved or affected in the interaction with PPAR $\gamma$ /RXR $\alpha$ . One of the LXXLL motifs would bind to the RXR $\alpha$  AF-2 surface, most likely NR2 as it has the highest affinity for RXR $\alpha$ , while a second LXXLL motif would interact more weakly with the PPAR $\gamma$  AF-2 surface, and the third LXXLL motif could interact transiently with a region somewhere on the heterodimer LBD surface. Moreover, the conclusions drawn from this study suggested that ligand binding to RXR $\alpha$  would be the primary driver of SRC-2 NRID recruitment, even if both RXR and PPAR $\gamma$  agonist ligands contribute to enhance the SRC-2 NRID interaction through different thermodynamic mechanisms.<sup>244</sup> The observation that multiple NR interaction motifs within coregulators mediate cooperative binding to NR dimers was also made for homodimers, as exemplified by the interaction of SRC-1 with PPAR $\gamma$  homodimer where two motifs of SRC-1 make simultaneous contacts with both PPAR $\gamma$  LBDs.<sup>236</sup> Moreover, a SAXS analysis showed how the binding of a single PGC-1 molecule to the coactivator groove on both LBD units of estrogen-related receptor gamma (ERR $\gamma$ ) homodimer and the induced stabilization of the dimerization interface within the homodimer are facilitated by the intrinsic disorder in the coactivator.<sup>243</sup> In other cases, the binding mode

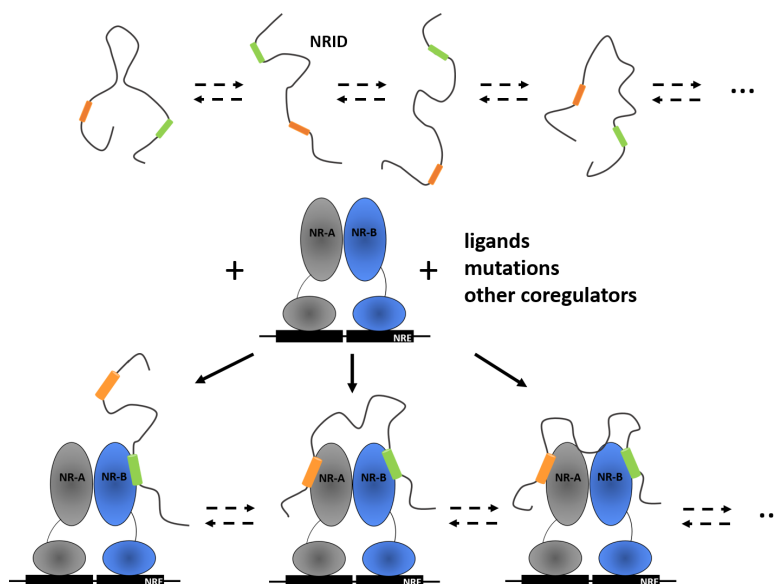


observed for different coactivators to heterodimers is in favor of a fully asymmetric interaction where binding occurs to a single LBD. SAXS studies showed that RAR $\alpha$ /RXR $\alpha$  and VDR/RXR $\alpha$  heterodimers recruit one coactivator molecule (MED1 or SRC-1 NRID) asymmetrically by only binding the partner of RXR $\alpha$ , the higher affinity of the coactivators for RAR $\alpha$  or VDR rather than RXR $\alpha$  promoting this selective binding.<sup>254</sup> A low-resolution SAXS analysis suggested also the same asymmetric binding of the entire NRID of SRC-2 to PPAR $\gamma$ /RXR $\alpha$  where the coactivator flexibly interacts only with PPAR $\gamma$  using one LXXLL motif,<sup>255</sup> even if these data appear to be contradicted by de Vera and collaborators<sup>244</sup>. Several structural studies on the interaction of coactivators with homodimers revealed also an asymmetric binding, suggesting that negative cooperativity could be an essential mechanism controlling the binding of coregulators to NR homodimers. The coactivator SRC-1 seems to bind only to one monomer in RAR and ER homodimers,<sup>255</sup> similarly as the one-monomer interaction of PGC-1 in the complex with ERR $\alpha$  and ERR $\gamma$  homodimers,<sup>256</sup> even if this last study appeared contradictory to<sup>243</sup>. A recent study<sup>249</sup> integrating information from different techniques into computational tool gave a new view on complexes between coregulator and NRs by showing that it exists an equilibrium between several conformational states of the complex formed by NCoR NRID and the heterodimer RAR $\alpha$ /RXR $\alpha$  and that this equilibrium can be tuned by cognate ligands (Fig. 8). Such an approach could permit to reconcile the different views of complexes between coregulators and NR homo- and hetero-dimers and is the most promising strategy for the structural characterization of highly dynamic proteins and complexes in solution. Even if it is possible to define 3D molecular envelopes describing the low-resolution structures of flexible macromolecular assemblies involving IDPs,<sup>243,254,255,257–259</sup> such a structural description is probably not the most appropriate for a disordered protein or a dynamic complex with multiple conformations and orientations. Ensemble analysis of explicit or coarse-grained models provide a more accurate characterization of flexible macromolecular assemblies, particularly when combined with available structure coordinates of folded-domains determined by X-ray crystallography and data affordable by NMR or by biochemical assays.<sup>34</sup>

Various HDX-MS analyses of complexes between coregulators and NRs showed significant changes in the conformational mobility of both partners during complexation. There is not only a stabilization of the conformational mobility of NR LBDs<sup>243</sup> but a disorder-to-order transition is also often observed for the coregulator.<sup>243,249,255</sup> Therefore, as for the AF-1, complex formation induces an increase in total secondary structure or a ternary folding of the coregulator often leading to its compaction<sup>260</sup> (Fig. 8). DNA may play a role in the



allosteric control of coregulator binding and dimerization in NR heterodimers. Differences in NR DNA response element sequences coming from different target gene promoters are thought to propagate an allosteric conformational change, initiating in the NR DBD and functionally transmitted to other domains of full-length NRs, affecting coregulator-NR interactions, in a ligand-specific manner.<sup>261</sup> For example, DNA binding to VDR/RXR $\alpha$  significantly impacts the conformation of the AF-2 surface in the LBDs of both components of the heterodimer,<sup>203</sup> modifying coregulator-NR interactions. DNA binding to PPAR $\gamma$ /RXR $\alpha$  stabilizes the conformation of the PPAR $\gamma$  LBD dimerization surface that interacts with the RXR $\alpha$  LBD, reduces the interaction conformational entropy of the complex to afford higher-affinity binding of SRC-2 NRID, and affects ligand potency for SRC-2 recruitment.<sup>244</sup> In addition, ligand binding to PPAR $\gamma$ /RXR $\alpha$  can allosterically affect the conformation of SRC-2 loop regions between interacting LXXLL motifs, as well as the C-terminal extended region that does not directly interact with PPAR $\gamma$ /RXR $\alpha$ . Quantitative studies on other NRs provide support for this, where the binding of the SRC-1 NRID to TR/RXR is influenced by the nature of the DNA response elements,<sup>262–264</sup> and where DNA and coregulator (SRC-2) cooperatively bind to the dimer of PR.<sup>265</sup> However, the structural mechanisms by which ligand and DNA binding to NRs may synergize to bind coregulator proteins require more studies.



**Fig. 8 Model of interaction of coregulator NRID with NR homo- and heterodimers bound to DNA.** The NRID of coregulators exists as an ensemble of conformers in equilibrium with each other, which are mainly unstructured with transient secondary structures corresponding to NR interaction motifs. In interaction with NRs, the NRID undergoes disorder-to-order transition and a dynamic complex with multiple conformations and orientations is formed and equilibrium between the different states of the complex can be influenced by ligands, NR mutations and other coregulators.

### 3.4 Importance of NR-NTDs in diseases and as pharmaceutical targets

NRs are well-validated drug targets against cancers and metabolic diseases. The most widely used drugs target the structured pocket of the LBD to modulate coregulator interaction with the activation function AF-2. A number of studies has shown that the NTD of NRs is not only critical for several aspects of NR action but also can be exploited as drug targets, thereby opening new opportunities for endocrine-based therapies.<sup>266,267</sup> Lack of significant sequence homology between NRs within the NTD also suggests that targeting this domain could give selective modulators of NRs, yielding fewer side effects. Traditional views of structural drug discovery would tend to consider such intrinsically disordered regions as unattractive targets due to their lack of structure and the concomitant lack of binding pockets for small molecules to interact. However, an analysis based on the structures of intrinsically disordered proteins when folded and bound to their partners suggests they actually have a higher proportion of potential cavities where a small molecule could bind than their folded counterparts.<sup>268</sup> In addition, their inherent flexibility may allow them to conformationally adapt to complement a small molecule. Interestingly, the development of drugs that target disordered regions to block protein-protein interactions is evolving and has provided potential for small molecules to have a sustained therapeutic effect.<sup>269,270</sup> Recent studies have shown that highly flexible and structurally dynamic NTD/AF-1 of steroid hormone receptors can be exploited as drug targets for endocrine-related cancers, thereby opening opportunities for development of novel small molecules that could block their interactions with coregulatory proteins outside the LBD/AF-2.<sup>271</sup>

Accordingly, few compounds that could find their application in the treatment of NTD-related diseases have been recently identified. The osmolyte trehalose has been shown to induce compaction or structure formation in the intrinsically disordered NTD of GR and to promote functionally ordered conformation of this domain, with potential clinical implications for human diseases involving disordered GR NTD.<sup>223</sup> A bisphenol A ester derivative (EPI-100) that specifically and covalently binds to the NTD of AR was shown to inhibit the constitutive activity of AR splice variants that lack the LBD and are suspected to be a cause of resistance in castration-recurrent prostate cancers (CRPC).<sup>224,272,273</sup> EPI-100 is the first small molecule inhibitor that directly binds to an intrinsically disordered region of a NR to enter clinical trials. Biophysical analysis revealed that EPI-001 alters folding of AR-NTD, thereby blocking the interaction of the AF-1 with specific coregulatory proteins. Other molecules, such as niphatenones and sintokamides, targeting AR NTD have also high

potential as “selective androgen receptor modulators” to overcome drug resistance.<sup>266</sup> More generally, exploration of the druggability of disordered regions of NRs and coregulators, although challenging since it precluded high throughput screening of small molecules, is highly promising.<sup>274</sup>

The AR-NTD also plays a direct role in the pathogenesis of spinal and bulbar muscular atrophy (SBMA), a late-onset neuromuscular disorder which is caused by expansion of the polyglutamine repeat (polyQ) in this domain.<sup>275</sup> The mechanistic basis of this phenomenon is a matter of intense debate, but it is widely admitted that polyQ expansion cause neurotoxicity because it decreases protein solubility and leads to the formation of aggregates.<sup>276</sup> Later on, it has been shown that polyQ repeats undergo a length-dependent conformational change producing a highly insoluble structure when their lengths reach a threshold. In the AR, biophysical analyses have revealed that the helical propensity of polyQ tracts increases upon expansion, and above 37 residues, the AR variants form fibrillar cytotoxic aggregates.<sup>277,278</sup> Using an approach combining MD simulations, NMR, CD and quantum mechanics, it has been recently shown that the helicity of the repeat correlates with its length as a result of the accumulation of unconventional interactions where glutamine side chains are involved in a hydrogen bond with the main chain carbonyl of the residue at position  $i-4$ . In addition, the study revealed that the strength of these interactions depends on the nature the acceptor residue and on its propensity to form a hydrogen bond that is strongly influenced by its side chain. Thus, the secondary structure of polyQ tracts appears to depend on their local environment.<sup>279</sup> These data provide a mechanistic basis of SBMA and a rationale for the development of therapeutic strategies to cure this disease.

### **3.5 Liquid-Liquid Phase Separation (LLPS) in NRs**

Equivalently to other transcription factors, a recent report linked the ER NTD to the liquid-liquid phase separation (LLPS) phenomenon observed for this NR and provides a potential functional role for this domain.<sup>280</sup> LLPS has emerged in the recent years as a key cellular mechanism to regulate multiple processes, including gene transcription.<sup>281</sup> The formation of liquid droplets brings together many constituents through cooperative interactions of multivalent biomolecules, such as disordered proteins and nucleic acids, constituting membrane less compartments.<sup>282</sup> Examples of such organelles in the nucleus include nucleoli, Cajal bodies and nuclear speckles.<sup>283</sup> These organelles display properties of liquid droplets, such as fusion and fission, and their formation can be formally described as a LLPS normally used soft-matter studies. To our knowledge, only the NTD of the human

RXR $\gamma$  has been described to form liquid droplets *in vitro* in the absence of other proteins.<sup>284</sup> For the previously mentioned ER case, liquid droplets were observed in the presence of the coactivator MED1 NRID, containing the LXXLL motifs. Interestingly, in the presence of the agonist (estrogen) the incorporation of the ER into the MED1 droplets is enhanced. These results suggest that transcription and incorporation into liquid droplets can be two facets of the same phenomenon.<sup>280</sup> Taking into account that NTDs and coregulators are low-complexity disordered proteins, which can form multivalent interactions, it is tempting to hypothesize that NRs can be part of liquid droplets that compartmentalize the transcription machinery and enhance its efficiency. Notice that the RNA polymerase II can also phase separate *in vivo* through its C-terminal domain.<sup>285</sup> In fact, these multiprotein condensates could explain the co-localization of multiple enhancers, enabling the simultaneous activation of multiple genes and their concerted regulation.<sup>286</sup>

## 4. Conclusions & Perspectives

The first 3D structure determination of myoglobin in 1958 by Kendrew and collaborators<sup>287</sup> in 1958 initiated the era of structural biology. A established paradigm defining that only well-structured proteins were functional had to be redefined when it was predicted that many proteins in eukaryotes were devoid of a globular form.<sup>31,288</sup> The characterization of the first partly unfolded protein under normal condition in the cell by Kriwacki and collaborators,<sup>289</sup> and the transcriptional coactivator, CPB and its paralog p300<sup>290,291</sup> were the demonstration of the relevance of IDPs.

In this chapter, the relevance of disorder is exemplified with two prototypical biological systems where the coordinated action of several players with different degrees of structuration is crucial for signaling and regulation. In these two biological machineries, GPCRs and NRs, we see the importance of disorder in orchestrating and modulating in a highly precise manner very complex biological processes. Indeed, the diversity of actions and outputs enabled by disorder is further increased thanks to PTMs and alternative splicing that widely occurs in these regions. These modifications modulate molecular recognition and specificity as well as define the cell type where their action should take place.

GPCRs and NRs are major players in cellular communication activities. As such, they represent targets of paramount interest for drug candidates, with applications in many different physiopathological processes. Mainly organized as a globular core with disordered appendices, most pharmacological and structural analyses have focused so far on the globular domains, while disordered regions are far less well characterized despite the fact that they are

pivotal in coupling the signaling partners and regulating transcription. Further characterization of these domains is thus required to understand signaling and fully exploit their therapeutic potential. However, the structural characterization is still obscured by intrinsic properties of these disordered domains, including the absence of stable secondary structure elements, high conformational dynamics, the occurrence of tightly regulated PTMs and the interaction with many different proteins. In this chapter we have highlighted the importance of a better description of the disordered parts of these complex machineries to design better modulators/inhibitors of ordered/disordered interacting surfaces to target specific pathway and decrease deleterious side effects of pharmacologic drugs.

## Acknowledgments

We thank current and former members of our research groups for invaluable contributions at many stages of this work. N.S. & P.B. acknowledge support of the Labex EpiGenMed (ANR-10-LABX-12-01), N.S. & J.L.B. support of the ANR GPCteR (ANR-17-CE11-0022-01). The CBS is a member of the French Infrastructure for Integrated Structural Biology (FRISBI) supported by the French National Research Agency (ANR-10-INSB-05).

## References

1. Potenza E, Di Domenico T, Walsh I, Tosatto SCE. MobiDB 2.0: An improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.* 2015;43(D1):D315-D320. doi:10.1093/nar/gku982
2. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradović Z. Intrinsic disorder and protein function. *Biochemistry.* 2002;41(21):6573-6582. doi:10.1021/bi012159+
3. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J Mol Biol.* 2004;337(3):635-645. doi:10.1016/j.jmb.2004.02.002
4. Pejaver V, Hsu WL, Xin F, Dunker AK, Uversky VN, Radivojac P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.* 2014;23(8):1077-1093. doi:10.1002/pro.2494
5. Iakoucheva LM, Radivojac P, Brown CJ, et al. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 2004;32(3):1037-1049. doi:10.1093/nar/gkh253
6. Cox CJ, Dutta K, Petri ET, et al. The regions of securin and cyclin B proteins recognized by the ubiquitination machinery are natively unfolded. *FEBS Lett.* 2002;527(1-3):303-308. doi:10.1016/S0014-5793(02)03246-5
7. Girdwood D, Bumpass D, Vaughan OA, et al. p300 transcriptional repression is mediated by SUMO modification. *Mol Cell.* 2003;11(4):1043-1054. doi:10.1016/S1097-2765(03)00141-2

8. Witze ES, Old WM, Resing KA, Ahn NG. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods*. 2007;4(10). doi:10.1038/1100
9. Uversky VN. Dancing protein clouds: The strange biology and chaotic physics of intrinsically disordered proteins. *J Biol Chem*. 2016;291(13):6681-6688. doi:10.1074/jbc.R115.685859
10. Uversky VN, Gillespie JR, Fink AL. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins Struct Funct Genet*. 2000;41(3):415-427. doi:10.1002/1097-0134(20001115)41:3<415::AID-PROT130>3.0.CO;2-7
11. Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK. Sequence complexity of disordered protein. *Proteins Struct Funct Genet*. 2001;42(1):38-48. doi:10.1002/1097-0134(20010101)42:1<38::AID-PROT50>3.0.CO;2-3
12. Davey NE, Van Roey K, Weatheritt RJ, et al. Attributes of short linear motifs. *Mol Biosyst*. 2012;8(1):268-281. doi:10.1039/c1mb05231d
13. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol*. 2015;16(1):18-29. doi:10.1038/nrm3920
14. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol*. 2005;6(3):197-208. doi:10.1038/nrm1589
15. Xie H, Vucetic S, Iakoucheva LM, et al. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J Proteome Res*. 2007;6(5):1882-1898. doi:10.1021/pr060392u
16. Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: Intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit*. 2005;18(5):343-384. doi:10.1002/jmr.747
17. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically Disordered Proteins in Human Diseases: Introducing the D 2 Concept . *Annu Rev Biophys*. 2008;37(1):215-246. doi:10.1146/annurev.biophys.37.032807.125924
18. Iakoucheva LM, Brown CJ, Lawson JD, Obradović Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol*. 2002;323(3):573-584. doi:10.1016/S0022-2836(02)00969-5
19. Landau K, Na I, Schenck R, Uversky V. Unfoldomics of prostate cancer: On the abundance and roles of intrinsically disordered proteins in prostate cancer. *Asian J Androl*. 2016;18(5):662-672. doi:10.4103/1008-682X.184999
20. Uversky VN. Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: Another illustration of the D 2 concept. *Expert Rev Proteomics*. 2010;7(4):543-564. doi:10.1586/epr.10.36
21. Santamaria N, Alhothali M, Alfonso MH, Breydo L, Uversky VN. Intrinsic disorder in proteins involved in amyotrophic lateral sclerosis. *Cell Mol Life Sci*. 2017;74(7):1297-1318. doi:10.1007/s00018-016-2416-6
22. Cheng Y, LeGall T, Oldfield CJ, Dunker AK, Uversky VN. Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry*. 2006;45(35):10448-10460. doi:10.1021/bi060981d
23. Du Z, Uversky VN. A comprehensive survey of the roles of highly disordered proteins in type 2 diabetes. *Int J Mol Sci*. 2017;18(10). doi:10.3390/ijms18102010
24. Dyson HJ, Wright PE. Unfolded proteins and protein folding studied by NMR. *Chem Rev*. 2004;104(8):3607-3622. doi:10.1021/cr030403s
25. Sibille N, Bernadó P. Structural characterization of intrinsically disordered proteins by the combined use of NMR and SAXS. *Biochem Soc Trans*. 2012;40(5):956-962. doi:10.1042/BST20120149
26. Delaforge E, Cordeiro TN, Bernadó P, Sibille N. Conformational characterization of intrinsically disordered proteins and its biological significance. *Mod Magn Reson*. 2018;381-399. doi:10.1007/978-3-319-28388-3\_52
27. Milles S, Salvi N, Blackledge M, Jensen MR. Characterization of intrinsically disordered proteins and their dynamic complexes: From in vitro to cell-like environments. *Prog Nucl Magn Reson Spectrosc*. 2018;109:79-100. doi:10.1016/j.pnmrs.2018.07.001

28. Dinkel H, Van Roey K, Michael S, et al. The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res.* 2014;42(D1). doi:10.1093/nar/gkt1047
29. Mi T, Merlin JC, Deverasetty S, et al. Minimoto Miner 3.0: Database expansion and significantly improved reduction of false-positive predictions from consensus sequences. *Nucleic Acids Res.* 2012;40(D1):252-260. doi:10.1093/nar/gkr1189
30. Van Der Lee R, Buljan M, Lang B, et al. Classification of intrinsically disordered regions and proteins. *Chem Rev.* 2014;114(13):6589-6631. doi:10.1021/cr400525m
31. Wright PE, Dyson HJ. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J Mol Biol.* 1999;293(2):321-331. doi:10.1006/jmbi.1999.3110
32. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol.* 2002;12(1):54-60. doi:10.1016/S0959-440X(02)00289-0
33. Spolar RS, Record MT. Coupling of local folding to site-specific binding of proteins to DNA. *Science (80- ).* 1994;263(5148):777-784. doi:10.1126/science.8303294
34. Cordeiro TN, Herranz-Trillo F, Urbanek A, et al. Small-angle scattering studies of intrinsically disordered proteins and their complexes. *Curr Opin Struct Biol.* 2017;42:15-23. doi:10.1016/j.sbi.2016.10.011
35. Bernadó P, Svergun DI. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol BioSyst.* 2012;8(1):151-167. doi:10.1039/C1MB05275F
36. Marsh JA, Forman-Kay JD. Ensemble modeling of protein disordered states: Experimental restraint contributions and validation. *Proteins Struct Funct Bioinforma.* 2012;80(2):556-572. doi:10.1002/prot.23220
37. Salmon L, Nodet G, Ozenne V, et al. NMR characterization of long-range order in intrinsically disordered proteins. *J Am Chem Soc.* 2010;132(24):8407-8418. doi:10.1021/ja101645g
38. Dunker AK, Brown CJ, Obradovic Z. Identification and functions of usefully disordered proteins. *Adv Protein Chem.* 2002;62:25-49. doi:10.1016/S0065-3233(02)62004-2
39. Robinson C V, Sali A, Baumeister W. The molecular sociology of the cell. *Nature.* 2007;450(7172):973-982. doi:10.1038/nature06523
40. Lagerström MC, Schiöth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov.* 2008;7(4):339-357. doi:10.1038/nrd2518
41. Hauser AS, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. *Nat Rev Drug Discov.* 2017;16(12):829-842. doi:10.1038/nrd.2017.178
42. Azam S, Haque ME, Jakaria M, Jo S-H, Kim I-S, Choi D-K. G-Protein-Coupled Receptors in CNS: A Potential Therapeutic Target for Intervention in Neurodegenerative Disorders and Associated Cognitive Deficits. *Cells.* 2020;9(506).
43. Ottow E, Weinmann H. Nuclear Receptors as Drug Targets: A Historical Perspective of Modern Drug Discovery. In: *Nuclear Receptors as Drug Targets.* John Wiley & Sons, Ltd; 2008:1-23. doi:10.1002/9783527623297.ch1
44. Heery DM, Kalkhoven E, Hoare S, Parker MG. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature.* 1997;387(6634):733-736. doi:10.1038/42750
45. Fuxreiter M, Tompa P, Simon I, Uversky VN, Hansen JC, Asturias FJ. Malleable machines take shape in eukaryotic transcriptional regulation. *Nat Chem Biol.* 2008;4(12):728-737. doi:10.1038/nchembio.127
46. Rasmussen SG, Choi HJ, Rosenbaum DM, et al. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature.* 2007;450(7168):383-387. doi:10.1038/nature06325
47. Palczewski K, Kumasaka T, Hori T, et al. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science (80- ).* 2000;289(5480):739-745. doi:10.1126/science.289.5480.739
48. Granier S, Kim S, Shafer AM, et al. Structure and conformational changes in the C-terminal domain of the beta2-adrenoceptor: insights from fluorescence resonance energy transfer studies. *J Biol Chem.* 2007;282(18):13895-13905. doi:10.1074/jbc.M611904200

49. Fonin A V., Darling AL, Kuznetsova IM, Turoverov KK, Uversky VN. *Multi-Functionality of Proteins Involved in GPCR and G Protein Signaling: Making Sense of Structure–Function Continuum with Intrinsic Disorder-Based Proteoforms*. Vol 76. Springer International Publishing; 2019. doi:10.1007/s00018-019-03276-1
50. Hilser VJ, Thompson EB. Structural dynamics, intrinsic disorder, and allostery in nuclear receptors as transcription factors. *J Biol Chem*. 2011. doi:10.1074/jbc.R111.278929
51. Uversky VN. Functional roles of transiently and intrinsically disordered regions within proteins. *FEBS J*. 2015;282(7):1182-1189. doi:10.1111/febs.13202
52. Pierce KL, Premont RT, Lefkowitz RJ. Seven-transmembrane receptors. *Nat Rev Mol Cell Biol*. 2002;3(9):639-650. doi:10.1038/nrm908
53. Lagerstrom MC, Schioth HB. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov*. 2008;7(4):339-357. doi:10.1038/nrd2518
54. Oldham WM, Hamm HE. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol*. 2008;9(1):60-71. doi:10.1038/nrm2299
55. Gurevich V V, Gurevich E V. GPCR Signaling Regulation: The Role of GRKs and Arrestins. *Front Pharmacol*. 2019;10:125. doi:10.3389/fphar.2019.00125
56. Komolov KE, Benovic JL. G protein-coupled receptor kinases: Past, present and future. *Cell Signal*. 2018;41:17-24. doi:10.1016/j.cellsig.2017.07.004
57. Vilardaga JP, Jean-Alphonse FG, Gardella TJ. Endosomal generation of cAMP in GPCR signaling. *Nat Chem Biol*. 2014;10(9):700-706. doi:10.1038/nchembio.1611
58. Gurevich V V., Gurevich E V. Plethora of functions packed into 45 kDa arrestins: biological implications and possible therapeutic strategies. *Cell Mol Life Sci*. 2019;76(22):4413-4421. doi:10.1007/s00018-019-03272-5
59. Grundmann M, Merten N, Malfacini D, et al. Lack of beta-arrestin signaling in the absence of active G proteins. *Nat Commun*. 2018;9(1):341. doi:10.1038/s41467-017-02661-3
60. Chaumont-Dubel S, Dupuy V, Bockaert J, Becamel C, Marin P. The 5-HT6 receptor interactome: New insight in receptor signaling and its impact on brain physiology and pathologies. *Neuropharmacology*. 2019:107839. doi:10.1016/j.neuropharm.2019.107839
61. Schioth HB, Fredriksson R. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen Comp Endocrinol*. 2005;142(1-2):94-101. doi:10.1016/j.ygcen.2004.12.018
62. Jaakola VP, Prilusky J, Sussman JL, Goldman A. G protein-coupled receptors show unusual patterns of intrinsic unfolding. *Protein Eng Des Sel*. 2005;18(2):103-110. doi:10.1093/protein/gzi004
63. Cherezov V, Rosenbaum DM, Hanson MA, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science (80- )*. 2007;318(5854):1258-1265. doi:10.1126/science.1150577
64. Shimada I, Ueda T, Kofuku Y, Eddy MT, Wuthrich K. GPCR drug discovery: integrating solution NMR data with crystal and cryo-EM structures. *Nat Rev Drug Discov*. 2019;18(1):59-82. doi:10.1038/nrd.2018.180
65. Rasmussen SG, DeVree BT, Zou Y, et al. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature*. 2011;477(7366):549-555. doi:10.1038/nature10361
66. Garcia-Nafria J, Tate CG. Cryo-Electron Microscopy: Moving Beyond X-Ray Crystal Structures for Drug Receptors and Drug Development. *Annu Rev Pharmacol Toxicol*. 2020;60:51-71. doi:10.1146/annurev-pharmtox-010919-023545
67. Liang YL, Belousoff MJ, Zhao P, et al. Toward a Structural Understanding of Class B GPCR Peptide Binding and Activation. *Mol Cell*. 2020;77(3):656-668 e5. doi:10.1016/j.molcel.2020.01.012
68. Koehl A, Hu H, Feng D, et al. Structural insights into the activation of metabotropic glutamate receptors. *Nature*. 2019;566(7742):79-84. doi:10.1038/s41586-019-0881-4
69. Qi X, Liu H, Thompson B, McDonald J, Zhang C, Li X. Cryo-EM structure of oxysterol-bound human Smoothed coupled to a heterotrimeric Gi. *Nature*. 2019;571(7764):279-283.



- doi:10.1038/s41586-019-1286-0
70. Yin W, Li Z, Jin M, et al. A complex structure of arrestin-2 bound to a G protein-coupled receptor. *Cell Res*. 2019;29(12):971-983. doi:10.1038/s41422-019-0256-2
  71. Huang W, Masureel M, Qu Q, et al. Structure of the neurotensin receptor 1 in complex with beta-arrestin 1. *Nature*. 2020;579(7798):303-308. doi:10.1038/s41586-020-1953-1
  72. Staus DP, Hu H, Robertson MJ, et al. Structure of the M2 muscarinic receptor-beta-arrestin complex in a lipid nanodisc. *Nature*. 2020;579(7798):297-302. doi:10.1038/s41586-020-1954-0
  73. Nguyen AH, Thomsen ARB, Cahill 3rd TJ, et al. Structure of an endosomal signaling GPCR-G protein-beta-arrestin megacomplex. *Nat Struct Mol Biol*. 2019;26(12):1123-1131. doi:10.1038/s41594-019-0330-y
  74. Venkatakrisnan AJ, Flock T, Prado DE, Oates ME, Gough J, Madan Babu M. Structured and disordered facets of the GPCR fold. *Curr Opin Struct Biol*. 2014;27:129-137. doi:10.1016/j.sbi.2014.08.002
  75. Hilger D, Masureel M, Kobilka BK. Structure and dynamics of GPCR signaling complexes. *Nat Struct Mol Biol*. 2018;25(1):4-12. doi:10.1038/s41594-017-0011-7
  76. Casiraghi M, Point E, Pozza A, Moncoq K, Baneres JL, Catoire LJ. NMR analysis of GPCR conformational landscapes and dynamics. *Mol Cell Endocrinol*. 2019;484:69-77. doi:10.1016/j.mce.2018.12.019
  77. Tastan O, Klein-Seetharaman J, Meirovitch H. The effect of loops on the structural organization of alpha-helical membrane proteins. *Biophys J*. 2009;96(6):2299-2312. doi:10.1016/j.bpj.2008.12.3894
  78. Woods AS. The dopamine D(4) receptor, the ultimate disordered protein. *J Recept Signal Transduct Res*. 2010;30(5):331-336. doi:10.3109/10799893.2010.513842
  79. Rahmeh R, Damian M, Cottet M, et al. Structural insights into biased G protein-coupled receptor signaling revealed by fluorescence spectroscopy. *Proc Natl Acad Sci U S A*. 2012;109(17):6733-6738. doi:10.1073/pnas.1201093109
  80. Bichet DG. [Nephrogenic diabetes insipidus]. *Nephrol Ther*. 2006;2(6):387-404. doi:10.1016/j.nephro.2006.07.010
  81. Rene P, Le Gouill C, Pogozheva ID, et al. Pharmacological chaperones restore function to MC4R mutants responsible for severe early-onset obesity. *J Pharmacol Exp Ther*. 2010;335(3):520-532. doi:10.1124/jpet.110.172098
  82. Goth CK, Petäjä-Repo U, Rosenkilde MM. G Protein-Coupled Receptors in the sweet spot: glycosylation and other post-translational modifications. *ACS Pharmacol Transl Sci*. 2020;in press.
  83. Lanctot PM, Leclerc PC, Escher E, Leduc R, Guillemette G. Role of N-glycosylation in the expression and functional properties of human AT1 receptor. *Biochemistry*. 1999;38(27):8621-8627. doi:10.1021/bi9830516
  84. Hawtin SR, Davies AR, Matthews G, Wheatley M. Identification of the glycosylation sites utilized on the V1a vasopressin receptor and assessment of their role in receptor signalling and expression. *Biochem J*. 2001;357(Pt 1):73-81. doi:10.1042/0264-6021:3570073
  85. Compton SJ, Sandhu S, Wijesuriya SJ, Hollenberg MD. Glycosylation of human proteinase-activated receptor-2 (hPAR2): role in cell surface expression and signalling. *Biochem J*. 2002;368(Pt 2):495-505. doi:10.1042/BJ20020706
  86. He J, Xu J, Castleberry AM, Lau AG, Hall RA. Glycosylation of beta(1)-adrenergic receptors regulates receptor surface expression and dimerization. *Biochem Biophys Res Commun*. 2002;297(3):565-572. doi:10.1016/s0006-291x(02)02259-3
  87. Markkanen PM, Petaja-Repo UE. N-glycan-mediated quality control in the endoplasmic reticulum is required for the expression of correctly folded delta-opioid receptors at the cell surface. *J Biol Chem*. 2008;283(43):29086-29098. doi:10.1074/jbc.M801880200
  88. Virion Z, Doly S, Saha K, et al. Sialic acid mediated mechanical activation of beta2 adrenergic receptors by bacterial pili. *Nat Commun*. 2019;10(1):4752. doi:10.1038/s41467-019-12685-6
  89. Mukai A, Yamamoto-Hino M, Komada M, Okano H, Goto S. Balanced ubiquitination

- determines cellular responsiveness to extracellular stimuli. *Cell Mol Life Sci*. 2012;69(23):4007-4016. doi:10.1007/s00018-012-1084-4
90. Malik R, Soh UJ, Trejo J, Marchese A. Novel roles for the E3 ubiquitin ligase atrophin-interacting protein 4 and signal transduction adaptor molecule 1 in G protein-coupled receptor signaling. *J Biol Chem*. 2012;287(12):9013-9027. doi:10.1074/jbc.M111.336792
  91. Chini B, Parenti M. G-protein-coupled receptors, cholesterol and palmitoylation: facts about fats. *J Mol Endocrinol*. 2009;42(5):371-379. doi:10.1677/JME-08-0114
  92. Prihandoko R, Bradley SJ, Tobin AB, Butcher AJ. Determination of GPCR Phosphorylation Status: Establishing a Phosphorylation Barcode. *Curr Protoc Pharmacol*. 2015;69:2 13 1-2 13 26. doi:10.1002/0471141755.ph0213s69
  93. Premont RT, Macrae AD, Aparicio SAJR, Kendall HE, Welch JE, Lefkowitz RJ. The GRK4 subfamily of G protein-coupled receptor kinases. Alternative splicing, gene organization, and sequence conservation. *J Biol Chem*. 1999;274(41):29381-29389. doi:10.1074/jbc.274.41.29381
  94. Taylor SS, Yang J, Wu J, Haste NM, Radzio-Andzelm E, Anand G. PKA: A portrait of protein kinase dynamics. *Biochim Biophys Acta - Proteins Proteomics*. 2004;1697(1-2):259-269. doi:10.1016/j.bbapap.2003.11.029
  95. Huse M, Kuriyan J. The conformational plasticity of protein kinases. *Cell*. 2002;109(3):275-282. doi:10.1016/S0092-8674(02)00741-9
  96. Komolov KE, Bhardwaj A, Benovic JL. Atomic structure of GRK5 reveals distinct structural features novel for G protein-coupled receptor kinases. *J Biol Chem*. 2015;290(34):20629-20647. doi:10.1074/jbc.M115.647297
  97. Komolov KE, Du Y, Duc NM, et al. Structural and Functional Analysis of a  $\beta$ 2-Adrenergic Receptor Complex with GRK5. *Cell*. 2017;169(3):407-421.e16. doi:10.1016/j.cell.2017.03.047
  98. Palczewski K, Buczylo J, Kaplan MW, Polans AS, Crabb JW. Mechanism of rhodopsin kinase activation. *J Biol Chem*. 1991;266(20):12949-12955.
  99. Chen CY, Dion SB, Kim CM, Benovic JL.  $\beta$ -Adrenergic receptor kinase. Agonist-dependent receptor binding promotes kinase activation. *J Biol Chem*. 1993;268(11):7825-7831.
  100. Kannan N, Haste N, Taylor SS, Neuwald AF. The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module (Proceedings of the National Academy of Sciences of the United States of America (2007) 104, (1272-1277) DOI:10.1073/pnas.0610251104). *Proc Natl Acad Sci U S A*. 2008;105(26):9130. doi:10.1073/pnas.0804708105
  101. Singh P, Wang B, Maeda T, Palczewski K, Tesmer JJG. Structures of rhodopsin kinase in different ligand states reveal key elements involved in G protein-coupled receptor kinase activation. *J Biol Chem*. 2008;283(20):14053-14062. doi:10.1074/jbc.M708974200
  102. Boguth CA, Singh P, Huang CC, Tesmer JJG. Molecular basis for activation of G protein-coupled receptor kinases. *EMBO J*. 2010;29(19):3249-3259. doi:10.1038/emboj.2010.206
  103. Yu QM, Cheng ZJ, Gan XQ, Bao G Bin, Li L, Pei G. The amino terminus with a conserved glutamic acid of G protein-coupled receptor kinases is indispensable for their ability to phosphorylate photoactivated rhodopsin. *J Neurochem*. 1999;73(3):1222-1227. doi:10.1046/j.1471-4159.1999.0731222.x
  104. Noble B, Kallal LA, Pausch MH, Benovic JL. Development of a yeast bioassay to characterize G protein-coupled receptor kinases: Identification of an NH<sub>2</sub>-terminal region essential for receptor phosphorylation. *J Biol Chem*. 2003;278(48):47466-47476. doi:10.1074/jbc.M308257200
  105. Pao CS, Barker BL, Benovic JL. Role of the amino terminus of G protein-coupled receptor kinase 2 in receptor phosphorylation. *Biochemistry*. 2009;48(30):7325-7333. doi:10.1021/bi900408g
  106. Huang CC, Yoshino-Koh K, Tesmer JJG. A surface of the kinase domain critical for the allosteric activation of G protein-coupled receptor kinases. *J Biol Chem*. 2009;284(25):17206-17215. doi:10.1074/jbc.M809544200
  107. Sterne-Marr R, Leahey PA, Bresee JE, et al. GRK2 activation by receptors: Role of the kinase

- large lobe and carboxyl-terminal tail. *Biochemistry*. 2009;48(20):4285-4293. doi:10.1021/bi900151g
108. Lodowski DT, Tesmer VM, Benovic JL, Tesmer JGG. The structure of G protein-coupled receptor kinase (GRK)-6 defines a second lineage of GRKs. *J Biol Chem*. 2006;281(24):16785-16793. doi:10.1074/jbc.M601327200
  109. Shiraishi Y, Natsume M, Kofuku Y, et al. Phosphorylation-induced conformation of beta2-adrenoceptor related to arrestin recruitment revealed by NMR. *Nat Commun*. 2018;9(1):194. doi:10.1038/s41467-017-02632-8
  110. Vishnivetskiy SA, Raman D, Wei J, Kennedy MJ, Hurley JB, Gurevich V V. Regulation of arrestin binding by rhodopsin phosphorylation level. *J Biol Chem*. 2007;282(44):32075-32083. doi:10.1074/jbc.M706057200
  111. Nobles KN, Xiao K, Ahn S, et al. Distinct phosphorylation sites on the beta(2)-adrenergic receptor establish a barcode that encodes differential functions of beta-arrestin. *Sci Signal*. 2011;4(185):ra51. doi:10.1126/scisignal.2001707
  112. Pack TF, Orlen MI, Ray C, Peterson SM, Caron MG. The dopamine D2 receptor can directly recruit and activate GRK2 without G protein activation. *J Biol Chem*. 2018;293(16):6161-6171. doi:10.1074/jbc.RA117.001300
  113. Prihandoko R, Alvarez-Curto E, Hudson BD, et al. Distinct Phosphorylation Clusters Determine the Signaling Outcome of Free Fatty Acid Receptor 4/G Protein-Coupled Receptor 120. *Mol Pharmacol*. 2016;89(5):505-520. doi:10.1124/mol.115.101949
  114. Bouzo-Lorenzo M, Santo-Zas I, Lodeiro M, et al. Distinct phosphorylation sites on the ghrelin receptor, GHSR1a, establish a code that determines the functions of ss-arrestins. *Sci Rep*. 2016;6:22495. doi:10.1038/srep22495
  115. Mokrosinski J, Frimurer TM, Sivertsen B, Schwartz TW, Holst B. Modulation of constitutive activity and signaling bias of the ghrelin receptor by conformational constraint in the second extracellular loop. *J Biol Chem*. 2012;287(40):33488-33502. doi:10.1074/jbc.M112.383240
  116. Lin X, Li M, Wang N, et al. Structural basis of ligand recognition and self-activation of orphan GPR52. *Nature*. 2020;579(7797):152-157. doi:10.1038/s41586-020-2019-0
  117. Rajagopalan L, Rajarathnam K. Ligand selectivity and affinity of chemokine receptor CXCR1. Role of N-terminal domain. *J Biol Chem*. 2004;279(29):30000-30008. doi:10.1074/jbc.M313883200
  118. Rajagopalan L, Rajarathnam K. Structural basis of chemokine receptor function--a model for binding affinity and ligand selectivity. *Biosci Rep*. 2006;26(5):325-339. doi:10.1007/s10540-006-9025-9
  119. Prado GN, Suetomi K, Shumate D, et al. Chemokine signaling specificity: essential role for the N-terminal domain of chemokine receptors. *Biochemistry*. 2007;46(31):8961-8968. doi:10.1021/bi7004043
  120. De Filippo E, Hinz S, Pellizzari V, et al. A2A and A2B adenosine receptors: The extracellular loop 2 determines high (A2A) or low affinity (A2B) for adenosine. *Biochem Pharmacol*. 2020;172:113718. doi:10.1016/j.bcp.2019.113718
  121. Catoire LJ, Damian M, Baaden M, Guittet E, Baneres JL. Electrostatically-driven fast association and perdeuteration allow detection of transferred cross-relaxation for G protein-coupled receptor ligands with equilibrium dissociation constants in the high-to-low nanomolar range. *J Biomol NMR*. 2011;50(3):191-195. doi:10.1007/s10858-011-9523-3
  122. Ferre G, Louet M, Saurel O, et al. Structure and dynamics of G protein-coupled receptor-bound ghrelin reveal the critical role of the octanoyl chain. *Proc Natl Acad Sci U S A*. 2019;116(35):17525-17530. doi:10.1073/pnas.1905105116
  123. Wheatley M, Wootten D, Conner MT, et al. Lifting the lid on GPCRs: the role of extracellular loops. *Br J Pharmacol*. 2012;165(6):1688-1703. doi:10.1111/j.1476-5381.2011.01629.x
  124. Bockaert J, Perroy J, Becamel C, Marin P, Fagni L. GPCR interacting proteins (GIPs) in the nervous system: Roles in physiology and pathologies. *Annu Rev Pharmacol Toxicol*. 2010;50:89-109. doi:10.1146/annurev.pharmtox.010909.105705
  125. Kim JM, Hwa J, Garriga P, Reeves PJ, RajBhandary UL, Khorana HG. Light-driven activation

- of beta 2-adrenergic receptor signaling by a chimeric rhodopsin containing the beta 2-adrenergic receptor cytoplasmic loops. *Biochemistry*. 2005;44(7):2284-2292. doi:10.1021/bi048328i
126. Garcia-Nafria J, Tate CG. Cryo-EM structures of GPCRs coupled to Gs, Gi and Go. *Mol Cell Endocrinol*. 2019;488:1-13. doi:10.1016/j.mce.2019.02.006
  127. Du Y, Duc NM, Rasmussen SGF, et al. Assembly of a GPCR-G Protein Complex. *Cell*. 2019;177(5):1232-1242 e11. doi:10.1016/j.cell.2019.04.022
  128. Granzin J, Wilden U, Choe HW, Labahn J, Krafft B, Buldt G. X-ray crystal structure of arrestin from bovine rod outer segments. *Nature*. 1998;391(6670):918-921. doi:10.1038/36147
  129. Hirsch JA, Schubert C, Gurevich V V, Sigler PB, Hughes H, Haven N. The 2.8° Crystal Structure of Visual Arrestin : The 2 . 8 Å A Model for Arrestin ' s Regulation. *Cell*. 1999;97(2):257-269.
  130. Han M, Gurevich V V., Vishnivetskiy SA, Sigler PB, Schubert C. Crystal structure of  $\beta$ -arrestin at 1.9 Å: Possible mechanism of receptor binding and membrane translocation. *Structure*. 2001;9(9):869-880. doi:10.1016/S0969-2126(01)00644-X
  131. Zhan X, Gimenez LE, Gurevich V V., Spiller BW. Crystal structure of arrestin-3 reveals the basis of the difference in receptor binding between two non-visual subtypes. *J Mol Biol*. 2011;406(3):467-478. doi:10.1016/j.jmb.2010.12.034
  132. Sutton RB, Vishnivetskiy SA, Robert J, et al. Crystal structure of cone arrestin at 2.3 Å: Evolution of receptor specificity. *J Mol Biol*. 2005;354(5):1069-1080. doi:10.1016/j.jmb.2005.10.023
  133. Martin J. Lohse and Carsten Hoffmann. *Arrestin Interactions with G Protein-Coupled Receptors*. Vol 219.; 2014. doi:10.1007/978-3-642-41199-1
  134. Sente A, Peer R, Srivastava A, et al. Molecular mechanism of modulating arrestin conformation by GPCR phosphorylation. *Nat Struct Mol Biol*. 2018;25(6):538-545. doi:10.1038/s41594-018-0071-3
  135. Yang F, Yu X, Liu C, et al. Phospho-selective mechanisms of arrestin conformations and functions revealed by unnatural amino acid incorporation and 19 F-NMR. *Nat Commun*. 2015;6(May):1-15. doi:10.1038/ncomms9202
  136. Shukla AK, Westfield GH, Xiao K, et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature*. 2014;512(7513):218-222. doi:10.1038/nature13430
  137. Shukla AK, Manglik A, Kruse AC, et al. Structure of active  $\beta$ -arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. *Nature*. 2013;497(7447):137-141. doi:10.1038/nature12120
  138. Latorraca NR, Wang JK, Bauer B, et al. Molecular mechanism of GPCR-mediated arrestin activation. *Nature*. 2018;557.
  139. Gurevich V V., Gurevich E V. The structural basis of the arrestin binding to GPCRs. *Mol Cell Endocrinol*. 2019;484(November 2018):34-41. doi:10.1016/j.mce.2019.01.019
  140. Gurevich V V., Gurevich E V., Uversky VN. Arrestins: structural disorder creates rich functionality. *Protein Cell*. 2018;9(12):986-1003. doi:10.1007/s13238-017-0501-8
  141. Kang Y, Zhou XE, Gao X, et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature*. 2015;523(7562):561-567. doi:10.1038/nature14656
  142. Kang DS, Kern RC, Puthenveedu MA, von Zastrow M, Williams JC, Benovic JL. Structure of an arrestin2-clathrin complex reveals a novel clathrin binding domain that modulates receptor trafficking. *J Biol Chem*. 2009;284(43):29860-29872. doi:10.1074/jbc.M109.023366
  143. Gurevich V V., Gurevich E V. Analyzing the roles of multi-functional proteins in cells: The case of arrestins and GRKs. *Crit Rev Biochem Mol Biol*. 2015;50(5):440-452. doi:10.3109/10409238.2015.1067185
  144. Lally CCM, Bauer B, Selent J, Sommer ME. C-edge loops of arrestin function as a membrane anchor. *Nat Commun*. 2017;8:1-12. doi:10.1038/ncomms14258
  145. Kumari P, Srivastava A, Banerjee R, et al. Functional competence of a partially engaged GPCR-beta-arrestin complex. *Nat Commun*. 2016;7:13416. doi:10.1038/ncomms13416
  146. Schoneberg T, Schulz A, Biebermann H, Hermsdorf T, Rompler H, Sangkuhl K. Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacol Ther*. 2004;104(3):173-

206. doi:10.1016/j.pharmthera.2004.08.008
147. Zhang P, Covic L, Kuliopulos A. Pepducins and Other Lipidated Peptides as Mechanistic Probes and Therapeutics. *Methods Mol Biol.* 2015;1324:191-203. doi:10.1007/978-1-4939-2806-4\_13
  148. Zhang P, Leger AJ, Baleja JD, et al. Allosteric Activation of a G Protein-coupled Receptor with Cell-penetrating Receptor Mimetics. *J Biol Chem.* 2015;290(25):15785-15798. doi:10.1074/jbc.M115.636316
  149. Vincenzi M, Costantini S, Scala S, et al. Conformational ensembles explored dynamically from disordered peptides targeting chemokine receptor CXCR4. *Int J Mol Sci.* 2015;16(6):12159-12173. doi:10.3390/ijms160612159
  150. Yedvabny E, Nerenberg PS, So C, Head-Gordon T. Disordered structural ensembles of vasopressin and oxytocin and their mutants. *J Phys Chem B.* 2015;119(3):896-905. doi:10.1021/jp505902m
  151. Germain P, Staels B, Dacquet C, Spedding M, Laudet V. Overview of nomenclature of nuclear receptors. *Pharmacol Rev.* 2006. doi:10.1124/pr.58.4.2
  152. Gustafsson JA. Historical overview of nuclear receptors. *J Steroid Biochem Mol Biol.* 2016. doi:10.1016/j.jsbmb.2015.03.004
  153. Weikum ER, Liu X, Ortlund EA. The nuclear receptor superfamily: A structural perspective. *Protein Sci.* 2018. doi:10.1002/pro.3496
  154. Anbalagan M, Huderson B, Murphy L, Rowan BG. Post-translational modifications of nuclear receptors and human disease. *Nucl Recept Signal.* 2012. doi:10.1621/nrs.10001
  155. Nagy L, Schwabe JWR. Mechanism of the nuclear receptor molecular switch. *Trends Biochem Sci.* 2004. doi:10.1016/j.tibs.2004.04.006
  156. le Maire A, Alvarez S, Shankaranarayanan P, R de Lera A, Bourguet W, Gronemeyer H. Retinoid Receptors and Therapeutic Applications of RAR/RXR Modulators. *Curr Top Med Chem.* 2012;12(6):505-527. doi:10.2174/156802612799436687
  157. Delfosse V, le Maire A, Balaguer P, Bourguet W. A structural perspective on nuclear receptors as targets of environmental compounds. *Acta Pharmacol Sin.* 2015. doi:10.1038/aps.2014.133
  158. Motlagh HN, Hilser VJ. Agonism/antagonism switching in allosteric ensembles. *Proc Natl Acad Sci U S A.* 2012. doi:10.1073/pnas.1120519109
  159. Onate SA, Boonyaratankornkit V, Spencer TE, et al. The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF2 domains of steroid receptors. *J Biol Chem.* 1998. doi:10.1074/jbc.273.20.12101
  160. Chen W, Rogatsky I, Garabedian MJ. MED14 and MED1 differentially regulate target-specific gene activation by the glucocorticoid receptor. *Mol Endocrinol.* 2006. doi:10.1210/me.2005-0318
  161. Choudhry MA, Ball A, McEwan IJ. The role of the general transcription factor IIF in androgen receptor-dependent transcription. *Mol Endocrinol.* 2006. doi:10.1210/me.2005-0486
  162. Chandra V, Wu D, Li S, Potluri N, Kim Y, Rastinejad F. The quaternary architecture of RAR $\beta$ -RXR $\alpha$  heterodimer facilitates domain-domain signal transmission. *Nat Commun.* 2017. doi:10.1038/s41467-017-00981-y
  163. Lonard DM, O'Malley BW. The Expanding Cosmos of Nuclear Receptor Coactivators. *Cell.* 2006. doi:10.1016/j.cell.2006.04.021
  164. Rosenfeld MG, Lunyak V V, Glass CK. Sensors and signals: A coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev.* 2006. doi:10.1101/gad.1424806
  165. Glass CK, Rosenfeld MG. The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* 2000. doi:10.1101/gad.14.2.121
  166. Lonard DM, O'Malley BW. Nuclear Receptor Coregulators: Judges, Juries, and Executioners of Cellular Regulation. *Mol Cell.* 2007. doi:10.1016/j.molcel.2007.08.012
  167. Lazar MA. Nuclear receptor corepressors. *Nucl Recept Signal.* 2003. doi:10.1621/nrs.01001
  168. Perissi V, Jepsen K, Glass CK, Rosenfeld MG. Deconstructing repression: Evolving models of

- co-repressor action. *Nat Rev Genet.* 2010. doi:10.1038/nrg2736
169. Tora L, Gronemeyer H, Turcotte B, Gaub MP, Chambon P. The N-terminal region of the chicken progesterone receptor specifies target gene activation. *Nature.* 1988. doi:10.1038/333185a0
  170. Miranda TB, Morris SA, Hager GL. Complex genomic interactions in the dynamic regulation of transcription by the glucocorticoid receptor. *Mol Cell Endocrinol.* 2013. doi:10.1016/j.mce.2013.03.002
  171. Lavery DN, McEwan IJ. Structure and function of steroid receptor AF1 transactivation domains: Induction of active conformations. *Biochem J.* 2005. doi:10.1042/BJ20050872
  172. He B, Bai S, Hnat AT, et al. An androgen receptor NH2-terminal conserved motif interacts with the COOH terminus of the Hsp70-interacting protein (CHIP). *J Biol Chem.* 2004. doi:10.1074/jbc.M403117200
  173. Wärnmark A, Treuter E, Wright APH, Gustafsson JÅ. Activation Functions 1 and 2 of Nuclear Receptors: Molecular Strategies for Transcriptional Activation. *Mol Endocrinol.* 2003. doi:10.1210/me.2002-0384
  174. McEwan IJ, Lavery D, Fischer K, Watt K. Natural disordered sequences in the amino terminal domain of nuclear receptors: lessons from the androgen and glucocorticoid receptors. *Nucl Recept Signal.* 2007. doi:10.1621/nrs.05001
  175. Huet G, Mérot Y, Percevault F, et al. Repression of the estrogen receptor- $\alpha$  transcriptional activity by the Rho/megakaryoblastic leukemia 1 signaling pathway. *J Biol Chem.* 2009. doi:10.1074/jbc.M109.045534
  176. Simons SS, Edwards DP, Kumar R. Minireview: Dynamic structures of nuclear hormone receptors: New promises and challenges. *Mol Endocrinol.* 2014. doi:10.1210/me.2013-1334
  177. Motlagh HN, Li J, Thompson EB, Hilser VJ. Interplay between allostery and intrinsic disorder in an ensemble. *Biochem Soc Trans.* 2012. doi:10.1042/BST20120163
  178. Li J, White JT, Saavedra H, et al. Genetically tunable frustration controls allostery in an intrinsically disordered transcription factor. *Elife.* 2017. doi:10.7554/eLife.30688
  179. Metivier R, Petit FG, Valotaire Y, Pakdel F. Function of N-terminal transactivation domain of the estrogen receptor requires a potential  $\alpha$ -helical structure and is negatively regulated by the A domain. *Mol Endocrinol.* 2000. doi:10.1210/mend.14.11.0546
  180. Rajbhandari P, Finn G, Solodin NM, et al. Regulation of Estrogen Receptor N-Terminus Conformation and Function by Peptidyl Prolyl Isomerase Pin1. *Mol Cell Biol.* 2012. doi:10.1128/mcb.06073-11
  181. Sadovsky Y, Webb P, Lopez G, et al. Transcriptional activators differ in their responses to overexpression of TATA-box-binding protein. *Mol Cell Biol.* 1995. doi:10.1128/mcb.15.3.1554
  182. Wärnmark A, Wikström A, Wright APH, Gustafsson JÅ, Härd T. The N-terminal Regions of Estrogen Receptor  $\alpha$  and  $\beta$  Are Unstructured in Vitro and Show Different TBP Binding Properties. *J Biol Chem.* 2001. doi:10.1074/jbc.M107875200
  183. Kitagawa H, Yanagisawa J, Fuse H, et al. Ligand-Selective Potentiation of Rat Mineralocorticoid Receptor Activation Function 1 by a CBP-Containing Histone Acetyltransferase Complex. *Mol Cell Biol.* 2002. doi:10.1128/mcb.22.11.3698-3706.2002
  184. Khan SH, Awasthi S, Guo C, et al. Binding of the N-terminal region of coactivator TIF2 to the intrinsically disordered AF1 domain of the glucocorticoid receptor is accompanied by conformational reorganizations. *J Biol Chem.* 2012. doi:10.1074/jbc.M112.411330
  185. Simons SS, Kumar R. Variable steroid receptor responses: Intrinsically disordered AF1 is the key. *Mol Cell Endocrinol.* 2013. doi:10.1016/j.mce.2013.06.007
  186. Gioeli D, Paschal BM. Post-translational modification of the androgen receptor. *Mol Cell Endocrinol.* 2012. doi:10.1016/j.mce.2011.07.004
  187. Kumar R, McEwan IJ. Allosteric modulators of steroid hormone receptors: Structural dynamics and gene regulation. *Endocr Rev.* 2012. doi:10.1210/er.2011-1033
  188. Kumar R, Thompson EB. Role of phosphorylation in the modulation of the glucocorticoid receptor's intrinsically disordered domain. *Biomolecules.* 2019. doi:10.3390/biom9030095

189. Kumar R, Moure CM, Khan SH, et al. Regulation of the structurally dynamic N-terminal domain of progesterone receptor by protein-induced folding. *J Biol Chem.* 2013. doi:10.1074/jbc.M113.491787
190. Kumar R, Thompson EB. Folding of the glucocorticoid receptor N-terminal transactivation function: Dynamics and regulation. *Mol Cell Endocrinol.* 2012. doi:10.1016/j.mce.2011.03.024
191. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol.* 2015. doi:10.1038/nrm3920
192. Brzovic PS, Heikaus CC, Kisselev L, et al. The Acidic Transcription Activator Gen4 Binds the Mediator Subunit Gal11/Med15 Using a Simple Protein Interface Forming a Fuzzy Complex. *Mol Cell.* 2011. doi:10.1016/j.molcel.2011.11.008
193. Di Lello P, Jenkins LMM, Jones TN, et al. Structure of the Tfb1/p53 Complex: Insights into the Interaction between the p62/Tfb1 Subunit of TFIIF and the Activation Domain of p53. *Mol Cell.* 2006. doi:10.1016/j.molcel.2006.05.007
194. Feng H, Jenkins LMM, Durell SR, et al. Structural Basis for p300 Taz2-p53 TAD1 Binding and Modulation by Phosphorylation. *Structure.* 2009. doi:10.1016/j.str.2008.12.009
195. Wright PE, Dyson HJ. Linking folding and binding. *Curr Opin Struct Biol.* 2009. doi:10.1016/j.sbi.2008.12.003
196. Oldfield CJ, Dunker AK. Intrinsically Disordered Proteins and Intrinsically Disordered Protein Regions. *Annu Rev Biochem.* 2014. doi:10.1146/annurev-biochem-072711-164947
197. Van Der Lee R, Buljan M, Lang B, et al. Classification of intrinsically disordered regions and proteins. *Chem Rev.* 2014. doi:10.1021/cr400525m
198. Kumar R, Lee JC, Bolen DW, Thompson EB. The Conformation of the Glucocorticoid Receptor AF1/taul Domain Induced by Osmolyte Binds Co-regulatory Proteins. *J Biol Chem.* 2001. doi:10.1074/jbc.M100825200
199. Bain DL, Franden MA, McManaman JL, Takimoto GS, Horwitz KB. The N-terminal region of human progesterone B-receptors: Biophysical and biochemical comparison to A-receptors. *J Biol Chem.* 2001. doi:10.1074/jbc.M102611200
200. Reid J, Kelly SM, Watt K, Price NC, McEwan IJ. Conformational analysis of the androgen receptor amino-terminal domain involved in transactivation. Influence of structure-stabilizing solutes and protein-protein interactions. *J Biol Chem.* 2002. doi:10.1074/jbc.M201003200
201. Aagaard MM, Siersbæk R, Mandrup S. Molecular basis for gene-specific transactivation by nuclear receptors. *Biochim Biophys Acta - Mol Basis Dis.* 2011. doi:10.1016/j.bbadis.2010.12.018
202. Chandra V, Huang P, Hamuro Y, et al. Structure of the intact PPAR- $\gamma$ -RXR- $\alpha$  nuclear receptor complex on DNA. *Nature.* 2008. doi:10.1038/nature07413
203. Zhang J, Chalmers MJ, Stayrook KR, et al. DNA binding alters coactivator interaction surfaces of the intact VDR-RXR complex. *Nat Struct Mol Biol.* 2011. doi:10.1038/nsmb.2046
204. Belorusova AY, Osz J, Petoukhov M V, et al. Solution Behavior of the Intrinsically Disordered N-Terminal Domain of Retinoid X Receptor  $\alpha$  in the Context of the Full-Length Protein. *Biochemistry.* 2016. doi:10.1021/acs.biochem.5b01122
205. Baskakov I V, Kumar R, Srinivasan G, Ji YS, Wayne Bolen D, Thompson EB. Trimethylamine N-oxide-induced cooperative folding of an intrinsically unfolded transcription-activating fragment of human glucocorticoid receptor. *J Biol Chem.* 1999. doi:10.1074/jbc.274.16.10693
206. Fischer K, Kelly SM, Watt K, Price NC, McEwan IJ. Conformation of the mineralocorticoid receptor N-terminal domain: Evidence for induced and stable structure. *Mol Endocrinol.* 2010. doi:10.1210/me.2010-0005
207. Kumar R, Serrette JM, Khan SH, Miller AL, Thompson EB. Effects of different osmolytes on the induced folding of the N-terminal activation domain (AF1) of the glucocorticoid receptor. *Arch Biochem Biophys.* 2007. doi:10.1016/j.abb.2007.06.019
208. Kumar R, Betney R, Li J, Thompson EB, McEwan IJ. Induced alpha-Helix Structure in AF1 of the Androgen Receptor upon Binding Transcription Factor TFIIF. *Biochemistry.* 2004. doi:10.1021/bi035934p
209. Kumar R, Volk DE, Li J, Lee JC, Gorenstein DG, Thompson EB. TATA box binding protein



- induces structure in the recombinant glucocorticoid receptor AF1 domain. *Proc Natl Acad Sci U S A*. 2004. doi:10.1073/pnas.0407160101
210. Wardell SE, Kwok SC, Sherman L, Hodges RS, Edwards DP. Regulation of the Amino-Terminal Transcription Activation Domain of Progesterone Receptor by a Cofactor-Induced Protein Folding Mechanism. *Mol Cell Biol*. 2005. doi:10.1128/mcb.25.20.8792-8808.2005
  211. Wood JR. Allosteric Modulation of Estrogen Receptor Conformation by Different Estrogen Response Elements. *Mol Endocrinol*. 2001. doi:10.1210/me.15.7.1114
  212. Brodie J, McEwan IJ. Intra-domain communication between the N-terminal and DNA-binding domains of the androgen receptor: Modulation of androgen response element DNA binding. *J Mol Endocrinol*. 2005. doi:10.1677/jme.1.01723
  213. Li W, Cavasotto CN, Cardozo T, et al. Androgen receptor mutations identified in prostate cancer and androgen insensitivity syndrome display aberrant ARTt-27 coactivator function. *Mol Endocrinol*. 2005. doi:10.1210/me.2005-0134
  214. Betney R, McEwan IJ. Role of conserved hydrophobic amino acids in androgen receptor AF-1 function. *J Mol Endocrinol*. 2003. doi:10.1677/jme.0.0310427
  215. Lavery DN, McEwan IJ. The human androgen receptor AF1 transactivation domain: Interactions with transcription factor IIF and molten-globule-like structural characteristics. *Biochem Soc Trans*. 2006. doi:10.1042/BST0341054
  216. Lavery DN, McEwan IJ. Structural characterization of the native NH2-terminal transactivation domain of the human androgen receptor: A collapsed disordered conformation underlies structural plasticity and protein-induced folding. *Biochemistry*. 2008. doi:10.1021/bi702221e
  217. Peng Y, Cao S, Kiselar J, et al. A Metastable Contact and Structural Disorder in the Estrogen Receptor Transactivation Domain. *Structure*. 2019. doi:10.1016/j.str.2018.10.026
  218. Bah A, Forman-Kay JD. Modulation of intrinsically disordered protein function by post-translational modifications. *J Biol Chem*. 2016. doi:10.1074/jbc.R115.695056
  219. Baker JMR, Hudson RP, Kanelis V, et al. CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. *Nat Struct Mol Biol*. 2007. doi:10.1038/nsmb1278
  220. He Y, Chen Y, Mooney SM, et al. Phosphorylation-induced conformational ensemble switching in an intrinsically disordered cancer/testis antigen. *J Biol Chem*. 2015. doi:10.1074/jbc.M115.658583
  221. Bah A, Vernon RM, Siddiqui Z, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature*. 2015. doi:10.1038/nature13999
  222. Garza AMS, Khan SH, Kumar R. Site-Specific Phosphorylation Induces Functionally Active Conformation in the Intrinsically Disordered N-Terminal Activation Function (AF1) Domain of the Glucocorticoid Receptor. *Mol Cell Biol*. 2010. doi:10.1128/mcb.00552-09
  223. Khan SH, McLaughlin WA, Kumar R. Site-specific phosphorylation regulates the structure and function of an intrinsically disordered domain of the glucocorticoid receptor. *Sci Rep*. 2017. doi:10.1038/s41598-017-15549-5
  224. De Mol E, Szulc E, Di Sanza C, et al. Regulation of Androgen Receptor Activity by Transient Interactions of Its Transactivation Domain with General Transcription Regulators. *Structure*. 2018. doi:10.1016/j.str.2017.11.007
  225. Kulkarni P, Jolly MK, Jia D, et al. Phosphorylation-induced conformational dynamics in an intrinsically disordered protein and potential role in phenotypic heterogeneity. *Proc Natl Acad Sci U S A*. 2017. doi:10.1073/pnas.1700082114
  226. Dunker AK, Uversky VN. Signal transduction via unstructured protein conduits. *Nat Chem Biol*. 2008. doi:10.1038/nchembio0408-229
  227. Borg M, Mittag T, Pawson T, Tyers M, Forman-Kay JD, Hue SC. Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitive. *Proc Natl Acad Sci U S A*. 2007. doi:10.1073/pnas.0702580104
  228. Mittag T, Kay LE, Forman-Kaya JD. Protein dynamics and conformational disorder in molecular recognition. *J Mol Recognit*. 2010. doi:10.1002/jmr.961
  229. Csizmok V, Follis AV, Kriwacki RW, Forman-Kay JD. Dynamic Protein Interaction Networks and New Structural Paradigms in Signaling. *Chem Rev*. 2016.

- doi:10.1021/acs.chemrev.5b00548
230. Hegyi H, Schad E, Tompa P. Structural disorder promotes assembly of protein complexes. *BMC Struct Biol.* 2007. doi:10.1186/1472-6807-7-65
  231. Millard CJ, Watson PJ, Fairall L, Schwabe JWR. An evolving understanding of nuclear receptor coregulator proteins. *J Mol Endocrinol.* 2013. doi:10.1530/JME-13-0227
  232. Ahmad KF, Melnick A, Lax S, et al. Mechanism of SMRT Corepressor Recruitment by the BCL6 BTB Domain. *Mol Cell.* 2003. doi:10.1016/S1097-2765(03)00454-4
  233. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. *Biochemistry.* 2006. doi:10.1021/bi0602718
  234. Hu X, Lazar MA. The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. *Nature.* 1999. doi:10.1038/47069
  235. Nagy L, Kao HY, Love JD, et al. Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes Dev.* 1999;13(24):3209-3216. doi:10.1101/gad.13.24.3209
  236. Nolte RT, Wisely GB, Westin S, et al. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor- $\gamma$ . *Nature.* 1998. doi:10.1038/25931
  237. Xu HE, Stanley TB, Montana VG, et al. Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPAR $\alpha$ . *Nature.* 2002. doi:10.1038/415813a
  238. Wang CY, Li CW, Chen JD, Welsh WJ. Structural model reveals key interactions in the assembly of the pregnane X receptor/corepressor complex. *Mol Pharmacol.* 2006. doi:10.1124/mol.106.022368
  239. Le Maire A, Teyssier C, Erb C, et al. A unique secondary-structure switch controls constitutive gene repression by retinoic acid receptor. *Nat Struct Mol Biol.* 2010. doi:10.1038/nsmb.1855
  240. Chang C, Norris JD, Grøn H, et al. Dissection of the LXXLL Nuclear Receptor-Coactivator Interaction Motif Using Combinatorial Peptide Libraries: Discovery of Peptide Antagonists of Estrogen Receptors  $\alpha$  and  $\beta$ . *Mol Cell Biol.* 1999. doi:10.1128/mcb.19.12.8226
  241. Darimont BD, Wagner RL, Apriletti JW, et al. Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev.* 1998. doi:10.1101/gad.12.21.3343
  242. McInerney EM, Rose DW, Flynn SE, et al. Determinants of coactivator LXXLL motif specificity in nuclear receptor transcriptional activation. *Genes Dev.* 1998. doi:10.1101/gad.12.21.3357
  243. Devarakonda S, Gupta K, Chalmers MJ, et al. Disorder-to-order transition underlies the structural basis for the assembly of a transcriptionally active PGC-1 $\alpha$ /ERR $\gamma$  complex. *Proc Natl Acad Sci U S A.* 2011. doi:10.1073/pnas.1113813108
  244. de Vera IMS, Zheng J, Novick S, et al. Synergistic Regulation of Coregulator/Nuclear Receptor Interaction by Ligand and DNA. *Structure.* 2017. doi:10.1016/j.str.2017.07.019
  245. Mohan A, Oldfield CJ, Radivojac P, et al. Analysis of Molecular Recognition Features (MoRFs). *J Mol Biol.* 2006. doi:10.1016/j.jmb.2006.07.087
  246. Iešmantavičius V, Dogan J, Jemth P, Teilum K, Kjaergaard M. Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew Chemie - Int Ed.* 2014. doi:10.1002/anie.201307712
  247. Hu X, Li Y, Lazar MA. Determinants of CoRNR-Dependent Repression Complex Assembly on Nuclear Hormone Receptors. *Mol Cell Biol.* 2001. doi:10.1128/mcb.21.5.1747-1758.2001
  248. Perissi V, Staszewski LM, McInerney EM, et al. Molecular determinants of nuclear receptor-corepressor interaction. *Genes Dev.* 1999. doi:10.1101/gad.13.24.3198
  249. Cordeiro TN, Sibille N, Germain P, et al. Interplay of Protein Disorder in Retinoic Acid Receptor Heterodimer and Its Corepressor Regulates Gene Expression. *Structure.* 2019. doi:10.1016/j.str.2019.05.001
  250. Delerive P, Wu Y, Burris TP, Chin WW, Suen CS. PGC-1 functions as a transcriptional coactivator for the retinoid X receptors. *J Biol Chem.* 2002. doi:10.1074/jbc.M109409200
  251. Bourdoncle A, Labesse G, Margueron R, Castet A, Cavailles V, Royer CA. The nuclear receptor coactivator PGC-1 $\alpha$  exhibits modes of interaction with the estrogen receptor distinct from those of SRC-1. *J Mol Biol.* 2005. doi:10.1016/j.jmb.2005.01.048
  252. Rha GB, Wu G, Shoelson SE, Chi YC. Multiple binding modes between HNF4 $\alpha$  and the

- LXXLL motifs of PGC-1 $\alpha$  lead to full activation. *J Biol Chem*. 2009. doi:10.1074/jbc.M109.052506
253. Phelan CA, Gampe RT, Lambert MH, et al. Structure of Rev-erb $\alpha$  bound to N-CoR reveals a unique mechanism of nuclear receptor-co-repressor interaction. *Nat Struct Mol Biol*. 2010. doi:10.1038/nsmb.1860
  254. Rochel N, Ciesielski F, Godet J, et al. Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. *Nat Struct Mol Biol*. 2011. doi:10.1038/nsmb.2054
  255. Osz J, Pethoukhov M V., Sirigu S, Svergun DI, Moras D, Rochel N. Solution structures of PPAR  $\gamma$  2/RXR  $\alpha$  complexes. *PPAR Res*. 2012. doi:10.1155/2012/701412
  256. Takacs M, Petoukhov M V., Atkinson RA, et al. The Asymmetric Binding of PGC-1 $\alpha$  to the ERR $\alpha$  and ERR $\gamma$  Nuclear Receptor Homodimers Involves a Similar Recognition Mechanism. *PLoS One*. 2013. doi:10.1371/journal.pone.0067810
  257. Longhi S, Receveur-Bréchet V, Karlin D, et al. The C-terminal domain of the measles virus nucleoprotein is intrinsically disordered and folds upon binding to the C-terminal moiety of the phosphoprotein. *J Biol Chem*. 2003. doi:10.1074/jbc.M300518200
  258. Marsh JA, Dancheck B, Ragusa MJ, Allaire M, Forman-Kay JD, Peti W. Structural diversity in free and bound states of intrinsically disordered protein phosphatase 1 regulators. *Structure*. 2010. doi:10.1016/j.str.2010.05.015
  259. Gosselin P, Oulhen N, Jam M, et al. The translational repressor 4E-BP called to order by eIF4E: New structural insights by SAXS. *Nucleic Acids Res*. 2011. doi:10.1093/nar/gkq1306
  260. Pavlin MR, Brunzelle JS, Fernandez EJ. Agonist ligands mediate the transcriptional response of nuclear receptor heterodimers through distinct stoichiometric assemblies with coactivators. *J Biol Chem*. 2014. doi:10.1074/jbc.M114.575423
  261. Meijsing SH, Pufall MA, So AY, Bates DL, Chen L, Yamamoto KR. DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science (80- )*. 2009. doi:10.1126/science.1164265
  262. Takeshita A, Yen PM, Ikeda M, et al. Thyroid hormone response elements differentially modulate the interactions of thyroid hormone receptors with two receptor binding domains in the steroid receptor coactivator-1. *J Biol Chem*. 1998. doi:10.1074/jbc.273.34.21554
  263. Putchá BDK, Fernandez EJ. Direct interdomain interactions can mediate allostereism in the thyroid receptor. *J Biol Chem*. 2009. doi:10.1074/jbc.M109.026682
  264. Velasco LFR, Togashi M, Walfish PG, et al. Thyroid hormone response element organization dictates the composition of active receptor. *J Biol Chem*. 2007. doi:10.1074/jbc.M610700200
  265. Heneghan AF, Connaghan-Jones KD, Miura MT, Bain DL. Coactivator assembly at the promoter: Efficient recruitment of SRC2 is coupled to cooperative DNA binding by the progesterone receptor. *Biochemistry*. 2007. doi:10.1021/bi700850v
  266. Monaghan AE, McEwan IJ. A sting in the tail: The N-terminal domain of the androgen receptor as a drug target. *Asian J Androl*. 2016. doi:10.4103/1008-682X.181081
  267. Shamilov R, Aneskievich BJ. Intrinsic Disorder in Nuclear Receptor Amino Termini: From Investigational Challenge to Therapeutic Opportunity. *Nucl Recept Res*. 2019. doi:10.32527/2019/101417
  268. Zhang Z, Boskovic Z, Hussain MM, et al. Chemical perturbation of an intrinsically disordered region of TFIID distinguishes two modes of transcription initiation. *Elife*. 2015. doi:10.7554/eLife.07777
  269. Dunker AK, Uversky VN. Drugs for “protein clouds”: Targeting intrinsically disordered transcription factors. *Curr Opin Pharmacol*. 2010. doi:10.1016/j.coph.2010.09.005
  270. Santofimia-Castaño P, Rizzuti B, Xia Y, et al. Targeting intrinsically disordered proteins involved in cancer. *Cell Mol Life Sci*. 2019. doi:10.1007/s00018-019-03347-3
  271. Kumar R. Steroid hormone receptors and prostate cancer: Role of structural dynamics in therapeutic targeting. *Asian J Androl*. 2016. doi:10.4103/1008-682X.183380
  272. Myung JK, Banuelos CA, Fernandez JG, et al. An androgen receptor N-terminal domain antagonist for treating prostate cancer. *J Clin Invest*. 2013. doi:10.1172/JCI66398

273. Andersen RJ, Mawji NR, Wang J, et al. Regression of Castrate-Recurrent Prostate Cancer by a Small-Molecule Inhibitor of the Amino-Terminus Domain of the Androgen Receptor. *Cancer Cell*. 2010. doi:10.1016/j.ccr.2010.04.027
274. Bushweller JH. Targeting transcription factors in cancer — from undruggable to reality. *Nat Rev Cancer*. 2019. doi:10.1038/s41568-019-0196-7
275. Spada ARL, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*. 1991. doi:10.1038/352077a0
276. Chen S, Ferrone FA, Wetzel R. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc Natl Acad Sci U S A*. 2002. doi:10.1073/pnas.182276099
277. Eftekharzadeh B, Piai A, Chiesa G, et al. Sequence Context Influences the Structure and Aggregation Behavior of a PolyQ Tract. *Biophys J*. 2016. doi:10.1016/j.bpj.2016.04.022
278. Davies P, Watt K, Kelly SM, Clark C, Price NC, McEwan IJ. Consequences of poly-glutamine repeat length for the conformation and folding of the androgen receptor amino-terminal domain. *J Mol Endocrinol*. 2008. doi:10.1677/JME-08-0042
279. Escobedo A, Topal B, Kunze MBA, et al. Side chain to main chain hydrogen bonds stabilize a polyglutamine helix in a transcription factor. *Nat Commun*. 2019. doi:10.1038/s41467-019-09923-2
280. Boija A, Klein IA, Sabari BR, et al. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*. 2018. doi:10.1016/j.cell.2018.10.042
281. Brangwynne CP, Eckmann CR, Courson DS, et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science (80- )*. 2009. doi:10.1126/science.1172046
282. Banjade S, Wu Q, Mittal A, Peeples WB, Pappu R V, Rosen MK. Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc Natl Acad Sci U S A*. 2015. doi:10.1073/pnas.1508778112
283. Zhu L, Brangwynne CP. Nuclear bodies: The emerging biophysics of nucleoplasmic phases. *Curr Opin Cell Biol*. 2015. doi:10.1016/j.ceb.2015.04.003
284. Sołtys K, Ożyhar A. Ordered structure-forming properties of the intrinsically disordered AB region of hRXR $\gamma$  and its ability to promote liquid-liquid phase separation. *J Steroid Biochem Mol Biol*. 2020. doi:10.1016/j.jsbmb.2019.105571
285. Boehning M, Dugast-Darzacq C, Rankovic M, et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol*. 2018. doi:10.1038/s41594-018-0112-y
286. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A Phase Separation Model for Transcriptional Control. *Cell*. 2017. doi:10.1016/j.cell.2017.02.007
287. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*. 1958;181(4610):662-666. doi:10.1038/181662a0
288. Dunker AK, Garner E, Guillot S, et al. Protein disorder and the evolution of molecular recognition: theory, predictions and observations. *Pac Symp Biocomput*. 1998:473-484.
289. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc Natl Acad Sci U S A*. 1996;93(21):11504-11509. doi:10.1073/pnas.93.21.11504
290. Shikama N, Ivon J, La Thangue N. The p300/CBP family: Integrating signals with transcription factors and chromatin. *Trends Cell Biol*. 1997;7(6):230-236. doi:10.1016/S0962-8924(97)01048-9
291. Janknecht R, Hunter T. Transcriptional control: Versatile molecular glue. *Curr Biol*. 1996;6(8):951-954. doi:10.1016/S0960-9822(02)00636-X