A combined computational and structural model of the full-length human prolactin receptor

Katrine Bugge[1], Elena Papaleo[1], Gitte W. Haxholm[1], Jonathan T. S. Hopper[2], Carol V. Robinson[2], Johan G. Olsen[1], Kresten Lindorff-Larsen[1] & Birthe B. Kragelund[1]

The prolactin receptor is an archetype member of the class I cytokine receptor family, comprising receptors with fundamental functions in biology as well as key drug targets. Structurally, each of these receptors represent an intriguing diversity, providing an exceptionally challenging target for structural biology. Here, we access the molecular architecture of the monomeric human prolactin receptor by combining experimental and computational efforts. We solve the NMR structure of its transmembrane domain in micelles and collect structural data on overlapping fragments of the receptor with small-angle X-ray scattering, native mass spectrometry and NMR spectroscopy. Along with previously published data, these are integrated by molecular modelling to generate a full receptor structure. The result provides the first full view of a class I cytokine receptor, exemplifying the architecture of more than 40 different receptor chains, and reveals that the extracellular domain is merely the tip of a molecular iceberg.

[1] Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark. [2] Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, South Parks Road, Oxford OX1 3QZ, UK. Correspondence and requests for materials should be addressed to B.B.K. (email: bbk@bio.ku.dk).

The prolactin receptor (PRLR) and its primary ligand prolactin (PRL) constitute a complex receptor system, linked to more than 300 biological functions ranging from reproduction and cell differentiation to immune responses[1,2]. It is best known for its role in mammary gland development and lactation[3] as well as the pathology hyperprolactinemia[4], but has also been linked to reproductive disorders[5] as well as breast[6,7] and prostate[8] tumorigenesis, and has therefore attracted significant pharmaceutical interest.

The PRLR belongs to the hematopoietic cytokine receptor superfamily, which consists of more than 40 members[9]. It is considered an archetype of the homodimeric group 1 of the family[9], constituting the simplest cytokine receptors and including, for example, the growth hormone receptor (GHR), the erythropoietin receptor (EPOR) and the thrombopoietin receptor[9]. All these receptors lack intrinsic kinase activity, making them dependent on associated kinases such as the Janus kinases (JAKs) to mediate signalling[9]. They are single-pass transmembrane (TM) proteins with similar overall topologies: (1) a folded extracellular domain (ECD) responsible for ligand binding; (2) a TM domain (TMD) connecting the extracellular- and intracellular parts and (3) an intrinsically disordered intracellular domain (ICD) orchestrating down-stream signalling[9,10]. Their ECDs fold into two fibronectin type III domains, named D1 (membrane-distal) and D2 (membrane-proximal). The latter contains a conserved WS-motif[11] that for the PRLR acts as a molecular switch during activation[12]. The ICDs have low-sequence conservation, except for two regions named Box1 and Box2. Box1 is a membrane-proximal proline-rich motif responsible for constitutive association of JAKs[13], while the function of Box2 remains unclear, although some studies have suggested it to be involved in JAK2 association[13–15]. Recently, studies have revealed that homodimerization of group I cytokine receptors may occur in the absence of hormone, and is insufficient for receptor activation[16–18]. Binding of a hormone to the ECDs leads to the formation of an asymmetric ternary complex consisting of one hormone and two receptor chains[19–21].

Parts of the PRL/PRLR receptor system have been structurally characterized including structures of the PRL[22], the 1:1 and 1:2 complexes of PRL:PRLR-ECD[16–18] and the unliganded human (h)PRLR-ECD-D2 (ref. 12), as well as a recent characterization of the hPRLR-ICD as being intrinsically disordered throughout its length[10]. However, despite substantial efforts, important structural and functional aspects of the PRLR remain uncharacterized. As no structure is available of neither the unliganded PRLR-ECD nor of the TMD, the overall structure and mechanism of signal transfer remains elusive. Signal transduction across the membrane following hormone binding is suggested to occur by subunit rearrangements or propagation of local structural changes, providing the TMDs with a key role. Furthermore, it has been shown for several class 1 cytokine receptors that the ligand-independent homodimerization takes place in the TMDs[16–18]. Despite these significant roles, the TMDs remain largely understudied, although cellular reports on their characteristics and roles in signal transduction are now emerging. A recent study of the GHR proposed a model for signal transduction where the homodimer TMD conformation switches upon receptor activation from a parallel to a left-hand crossover structure, thereby separating the TMDs at the C-terminal end and eventually bringing the two JAK2s into productive positions[17]. However, analogous alanine-insertion studies of the GHR[23] and the PRLR[24] suggested that the activation mechanism of the two related receptors differs in a poorly understood manner.

Although high-resolution structures of multi-pass trans-membrane proteins in connection with extra-membranous globular domains have been solved (for example, refs 25,26), no such structures are available of single-pass transmembrane proteins, which have fewer stabilizing contacts in their membrane-embedded region. The structural characterization of the full PRLR is further hampered by the challenges associated with studies of a protein consisting of three structurally diverse domains: a soluble, folded domain, a membrane-embedded domain and an intrinsically disordered domain. Currently, no single high-resolution method is capable of describing the structural characteristics of a protein of this size and complexity. To overcome these inherent limitations in sample preparation and individual structure-determination techniques, we here apply a multidisciplinary divide-and-conquer approach in which we combine data acquired with different techniques on overlapping domain variants of the hPRLR with molecular modelling[9]. In the centre of this endeavour lies the determination of the hPRLR-TMD structure, which we here solve in 1,2-dihexanoyl-sn-glycero-phosphocholine (DHPC) micelles using nuclear magnetic resonance (NMR) spectroscopy. Combining this with previously published and new biophysical data of the soluble domains, we generate a structural model of the monomeric hPRLR, suggested to represent up to 70% of the PRLRs in the cell membrane[15]. This hPRLR structure provides the first full molecular architecture of a class I cytokine receptor, exemplifying more than 40 different receptor chains, and reveals that the extracellular domain is merely the tip of a molecular iceberg.

Results

Structure of hPRLR-TMD in micelles. As the missing piece in the structural description of the hPRLR, we determined the structure of the TMD in a membrane-mimicking environment using solution state NMR spectroscopy. To ensure overlap with previously determined structures, we used a hPRLR-TMD construct with a five-residue overlap between hPRLR-D2 (ref. 12) and hPRLR-TMD in its N-terminus (F206-D210), and a five-residue overlap with hPRLR-ICD[10] in its C-terminus (G236-V240). The resulting 37-residue hPRLR-TMD harboured residues F206-V240, as well as an N-terminal G-S sequence.

hPRLR-TMD was expressed and fast-track purified to >95% purity (Supplementary Fig. 1) using a novel method[27] and screened for suitable detergents and temperatures for the structural studies (Supplementary Fig. 2). A high concentration of DHPC (TMD:DHPC 1:700) at 37 °C provided narrow linewidths and the expected number of peaks and hence was selected for the structural studies. An SDS-PAGE of hPRLR-TMD (Supplementary Fig. 1) revealed a clear monomer band below 14 kDa, with a faint dimer band just above 14 kDa. The deviation from the average molecular weight of hPRLR-TMD (3,959.78 Da) can be explained by the commonly encountered gel-shifting phenomenon observed for membrane proteins[28]. The oligomeric state of hPRLR-TMD reconstituted in DHPC (1:700) was further evaluated by native mass spectrometry, which is sensitive to oligomerization of membrane proteins[29] (Fig. 1a). The peaks identified at 1,980.56 m/z in the $+2$ charge state and at 3,958.16 m/z in the $+1$ charge state are consistent with the average molecular weight of one monomer (3,959.78 Da). No peaks representing higher oligomers of hPRLR-TMD were observed. Thus, hPRLR-TMD appeared monomeric under these conditions.

We proceeded to assign the chemical shifts of hPRLR-TMD in DHPC micelles by standard heteronuclear NMR methods. Manual assignments resulted in 99.0% completeness for backbone ^{15}N, ^{13}C and ^{1}H resonances, and 90.3% for side chain proton resonances. The secondary structure was evaluated by secondary

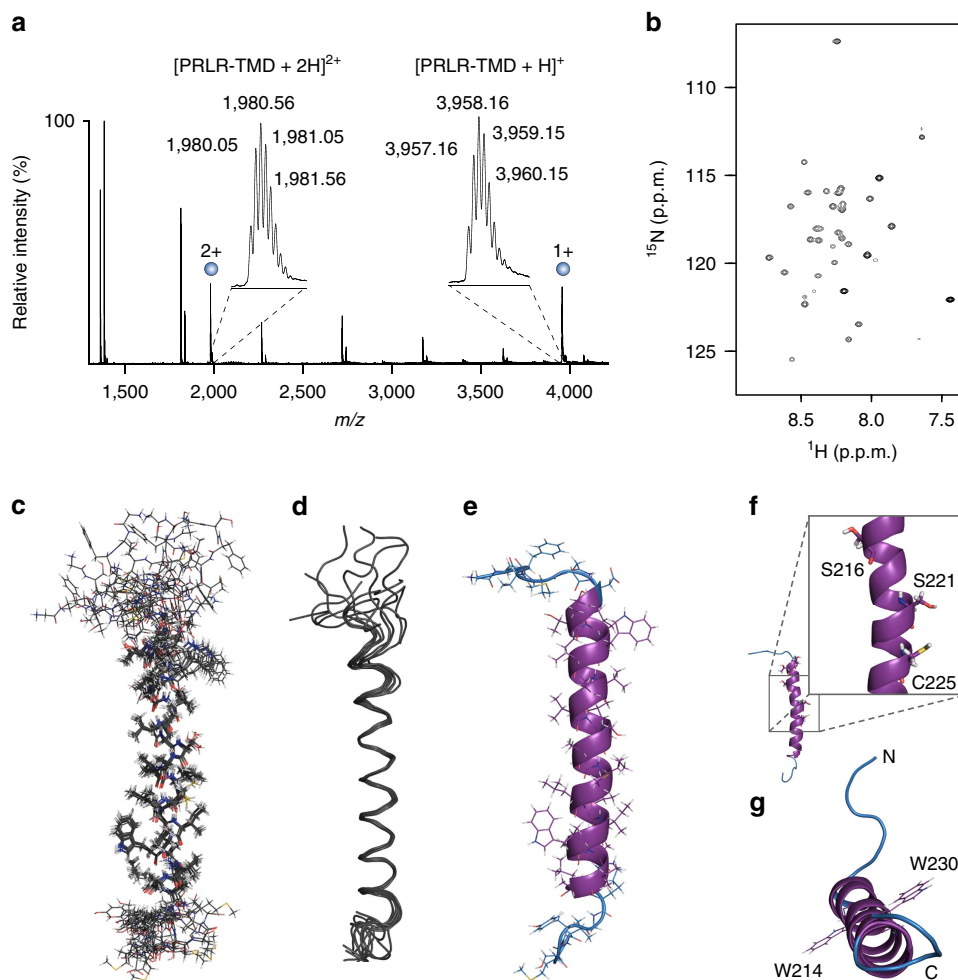**Figure 1 | Structure of hPRLR-TMD in DHPC micelles.** (**a**) Native mass spectrum of hPRLR-TMD in 700× molar excess of DHPC. The peaks at 1,980.56 m/z (+2) and 3,958.16 m/z (+1) are consistent with one hPRLR-TMD monomer (3,959.78 Da), while no peaks representing higher oligomers of hPRLR-TMD were observed. (**b**) ^{1}H-^{15}N-HSQC spectrum of 1 mM hPRLR-TMD in 700 mM DHPC, 50 mM NaCl, 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7.2, 37 °C. Superimposition of the 10 lowest energy hPRLR-TMD structures in (**c**) stick and (**d**) ribbon representations. (**e**) Lowest energy structure of hPRLR-TMD in DHPC micelles with backbone atoms in cartoon representation. The DHPC embedded region is shown in purple. (**f**) Cartoon representation of hPRLR-TMD with polar side chains shown as sticks. (**g**) Top view of hPRLR-TMD highlighting the two symmetrically placed tryptophans.

chemical shift (SCS) values of C$^\alpha$, C$^\beta$, C′ and H$^\alpha$ calculated from published random coil values[30], and with the motif identification from chemical shifts (MICS) programme[31] using C$^\alpha$, C$^\beta$, C′, N^H, H^N and H$^\alpha$ chemical shifts. Together these analyses suggested α-helical conformation for residues D210-A233 (Supplementary Figs 3a and 4a), additionally supported by diagnostic α-helical nuclear Overhauser effects (NOEs) (Supplementary Fig. 3a). The SCSs further suggested that residues L234–G236 continued in a transient α-helical turn (Supplementary Figs 3a and 4a). The SCSs and NOE patterns of residues F206–N209 and Y237–V240 did not support highly populated secondary structures for these residues. Analysis of possible N- and C-terminal α-helical capping (N-cap and C-cap) motifs using MICS (Supplementary Fig. 4c,d) suggested N209 or D210 (0.114 and 0.118, respectively) to potentially form transient N-cap structures.

The above results were supported by ^{15}N-T_2-relaxation times measured on hPRLR-TMD in DHPC micelles (Supplementary Fig. 3c). Little variation was observed in backbone dynamics for residues T212–L234 with ^{15}N-T_2-relaxation times just below 50 ms. Towards the N- and C-terminus, the relaxation times gradually increased, suggesting faster dynamics and correlating well with the lack of secondary structure in these regions. Similar relaxation measurements conducted on hPRLR-ICD$_{G236–Q396}$

(including residues S238–V240 of hPRLR-TMD) (Fig. 2a–c) supported these findings and further showed fast backbone dynamics of the ICD.

The structure of hPRLR-TMD in DHPC micelles was determined using 754 NMR-derived restraints (Table 1), including short- and medium-range NOE-, dihedral angle- and hydrogen bonding restraints (Supplementary Fig. 3a,b). Dihedral angle restraints were estimated by TALOS[32] and φ-angles further refined from ^{3}J(H^N–H$^\alpha$) coupling constants[33] (Supplementary Fig. 3b). We used the presence of small coupling constants and characteristic α-helical NOE patterns as basis for including hydrogen bonds between H$_i$(N) and (C)O$_{i-4}$ as restraints (Supplementary Fig. 3b, grey circles). A final set of 200 structures was calculated with Aria/crystallography & NMR system (CNS),[34,35] and the 10 structures with the lowest energy conformations and without significant violations were selected to represent the structure of the monomeric hPRLR-TMD (Fig. 1c,d and Table 1). The lowest energy structure formed a single α-helix from residues D210–A233 (Fig. 1e) with a length of 36 Å and two symmetrically placed tryptophans pointing in opposite directions on each side of the α-helix (Fig. 1g). A slight bend around S221 was observed, with a bending angle of 6°. The α-helix contained mainly hydrophobic residues, but harboured a few residues in the
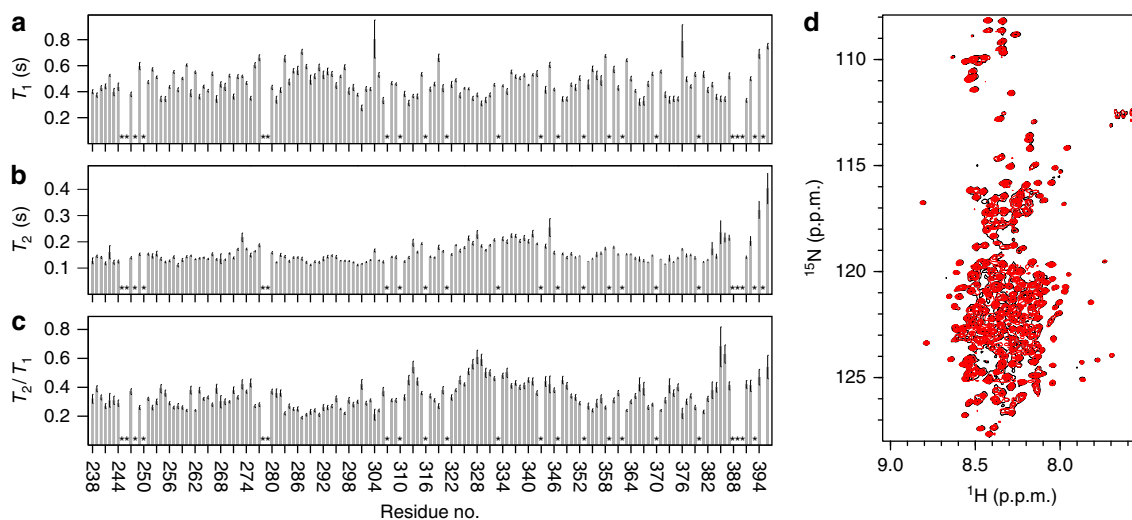
Figure 2 | Intrinsic structural properties of hPRLR-ICD. (**a**) T_1- and (**b**) T_2-relaxation times for hPRLR-ICD$_{G236-Q396}$ backbone amides plotted against residue number. (**c**) T_2/T_1 for hPRLR-ICD$_{G236-Q396}$ backbone amides plotted against residue number. The vertical black lines represent the estimated error of each fit. '*' indicates residues for which the relaxation time was not determined. (**d**) Concentration dependence of the spectral properties of hPRLR-ICD$_{G236-H598}$. Overlay of ^{1}H-^{15}N-HSQC spectra of 20 μM (red) and 400 μM (black) hPRLR-ICD$_{G236-H598}$. The spectra were recorded under identical conditions, except that the low-concentration sample was recorded with 152 transients, while the high concentration was recorded with 16. No changes in chemical shifts were detected as a consequence of the 20 times dilution, highlighting the monomeric properties of hPRLR-ICD$_{G236-H598}$.

Table 1 | NMR restraints per structure and statistics for the hPRLR-TMD structure ensemble.

	hPRLR-TMD
NMR distance and dihedral constraints	
Distance constraints	
Total NOE	677
Intra-residue	396
Inter-residue	281
Sequential ($|i - j| = 1$)	139
Medium-range ($|i - j| < 4$)	142
Long-range ($|i - j| > 5$)	0
Intermolecular	0
Hydrogen bonds	13
Total dihedral angle restraints	
φ	47
ψ	26
Structure statistics*	
Violations (mean and s.d.)	
Distance constraints (Å)	0.020 +/− 0.002
Dihedral angle constraints (°)	1.84 +/− 0.001
Max. dihedral angle violation (°)	5.00
Max. distance constraint violation (Å)	0.301
Deviations from idealized geometry	
Bond lengths (Å)	0.002 +/− 0.000
Bond angles (°)	0.368 +/− 0.004
Impropers (°)	0.298 +/− 0.009
Average pairwise r.m.s. deviation† (Å)	
Heavy	0.40 +/− 0.09
Backbone	0.30 +/− 0.11

*Statistics were calculated and averaged over an ensemble of the 10 lowest energy structures out of 200 calculated structures.
†Pairwise r.m.s. deviation was calculated among 10 structures, and includes only the structured region of the protein.

N-terminal part with side chains capable of forming hydrogen bonds: S216, S221 and C225 (Fig. 1f). S221 and C225 were located on the same side of the helix, while S216 was located half a turn away (Fig. 1f).

Envelope structure of the unliganded hPRLR-ECD. To generate a full structural model of the monomeric hPRLR, we needed a representation of the unliganded ECD. However, all known structures of the PRLR-ECD are crystal structures solved in complex with PRL, placental lactogen or PRL-based antagonists[19,20,36–38]. We previously attempted to solve the structure of the unliganded hPRLR-ECD in solution, but found it incompatible with acquisition of high-quality NMR data[12]. Instead, the structure of hPRLR-ECD-D2 was solved, revealing that the overall conformation of the unliganded ECD-D2 was similar to that of the liganded state, but with significant local differences in strand length and in the WS-motif[12]. Thus, to obtain structural information on the full ECD in the unliganded state, we measured solution small-angle X-ray scattering (SAXS) profiles of the unliganded hPRLR-ECD, shown to be monomeric by size-exclusion chromatography[22] (Figs 3b and 4b). The resulting SAXS curve was fitted to the coordinates of a liganded hPRLR-ECD structure[37] using CRYSOL[39]. The theoretical scattering curve of this ECD model fits in detail with the experimentally determined envelope ($\chi^2 = 1.2$) and a pseudo-atom model was constructed using DAMMIN (Fig. 3a,b). The angle between D1 and D2 in the ECD model was, within the resolution measurable by SAXS analysis, similar to both the liganded and unliganded GHR and EPOR structures (PDB entries 2AEW, 1A22, 1ERN and 4Y5Y). Together these findings suggested that no substantial structural rearrangement occurs upon ligand binding to hPRLR-ECD, and provided a model of the unliganded hPRLR-ECD.

Ensemble description of the hPRLR-ICD. We recently showed the long-form hPRLR-ICD to be intrinsically disordered with five transiently populated α-helices[10]. Here, by NMR, we further investigated the intrinsic propensity of the ICD to homodimerize and found no evidence to support dimerization of the unmodified ICD chain (Fig. 2d). Thus, we used Flexible Meccano[40] (FM) to generate a model of the hPRLR-ICD$_{G236-H598}$ as an ensemble of monomeric, rapidly inter-converting conformers. For validation purposes, NMR diffusion experiments were applied to measure
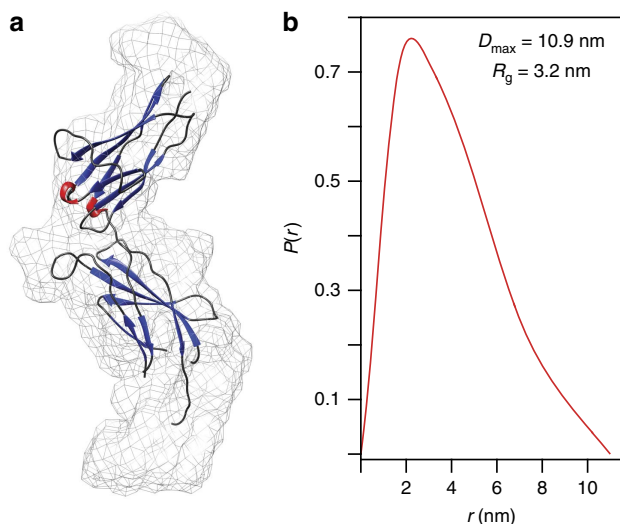
Figure 3 | Monomeric, unliganded hPRLR-ECD envelope generated from SAXS data. (**a**) Superimposition of the SAXS ECD envelope (mesh) and a crystal structure of the liganded hPRLR-ECD (PDB entry 3D48, β-sheets are blue, α-helices are red, loops are grey) generated with UCSF CHIMERA[66]. (**b**) Pair distance distribution function of the SAXS data collected on the unliganded, monomeric ECD.

the hydrodynamic radius (R_H) of hPRLR-ICD to $74\,Å \pm 1\,Å$ (Fig. 4e). The calculated average R_H of the FM ensemble was $58 \pm 6\,Å$, suggesting it to be slightly more compact than proposed by the experimentally determined R_H (discussed below).

Membrane embedment and interactions. Although the TMD is presumed to be the only truly membrane-embedded part of the hPRLR, the ECD and the ICD are tethered by the TMD to the outer or inner leaflets of the lipid bilayer, respectively. Potential interactions between the hPRLR and the membrane bilayer should therefore not be ignored. We recently established that hPRLR-ICD interacts with lipids characteristic of the inner membrane leaflet[10], but it is unknown if—and to what extent—the ECD interacts with the membrane as suggested from theoretical considerations[41].

Using an analogous approach as for hPRLR-ICD[10], putative interactions between hPRLR-D2 and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), the most abundant lipid of the mammalian outer membrane[42], were probed using NMR by titrating small-unilamellar vesicles (SUVs) into a solution of ^{15}N-hPRLR-ECD-D2. However, in contrast to the ICD, we observed no effect on the chemical shifts or peak intensities on addition of $16 \times$ molar excess of POPC (Fig. 4a), suggesting no significant affinity between the ECD-D2 and the major lipid constituent of the membrane bilayer.

Embedment of hPRLR-TMD in DHPC was examined with a series of experiments. First, the proximity and interaction between water and backbone amides were investigated using amide–water NOEs (Supplementary Fig. 5a). Not surprisingly, NOEs were readily detected between water and the amides in the flexible N- and C-terminal regions (F206-D210, K235-V240), whereas no detectable NOEs were identified from the majority of the amides in the TMD α-helix. However, we did observe NOEs between water and the amides of T212, V213, S216, V217 and S221, some of which are positioned close to the centre of the TMD. However, these NOE peaks could originate from the proximate hydroxyls of T212, S216 and S221. To further understand the TMD embedment, we therefore performed

hydrogen–deuterium (H–D) exchange experiments at different levels of D_2O following establishment of a quasi-stationary state (Supplementary Figs 5b and 6). These data supported the pattern revealed from the amide–water NOE data, suggesting some degree of water contact at the N-terminal part of the α-helix, potentially explained by this being more polar than the C-terminal (Supplementary Fig. 5c). The chemical shifts, NOEs, coupling constants and ^{15}N-T_2-relaxation times clearly established structure in the N-terminal end of the α-helix. Thus, the non-uniform exchange pattern and the water NOEs are likely not a result of extreme dynamics, but instead of the N-terminal polarity in combination with the properties of the detergent.

DHPC is, as a short-tail version of POPC, in theory only capable of assembling a $\sim 16\,Å$ wide hydrocarbon bilayer[43], compared to the $29\,Å$ of a POPC bilayer[44], rendering a spherical micelle-embedment model unlikely. Thus, in light of the above data, it seems likely that DHPC patches the TMD surface with a prolate ellipsoid monolayer (Supplementary Fig. 5d) as suggested also for the OmpX[43]. Despite this, the hPRLR-TMD region embedded in DHPC fits remarkably well with the region predicted to be within the native bilayer (T211–L234)[45], possibly caused by anchoring at two charged residues; D210 and K235. Hence, even with the shortcomings of DHPC, the data collectively suggest that the membrane embedment range was well-simulated and constituted T211–L234.

In conclusion, the ICD has previously been shown to interact with membrane bilayer constituents, whereas the ECD-D2 appears intrinsically void of significant affinity for POPC. The data collectively supported that T211–L234 of the TMD were embedded in the membrane mimetics, while the extracellular F206–D210 and intracellular K235–V240 protruded at each end.

Data integration to generate a full structural model. We combined the experimentally based ensemble of the ICD and the unliganded ECD with the structure of the TMD to generate a complete structural model of the hPRLR (Fig. 4). The overlapping region between the hPRLR-ECD and hPRLR-TMD (F206-D210) formed little regular structure, as evidenced by the SCSs from the solution structure of the ECD-D2 (ref. 12), and the coupling constants, relaxation rates and SCSs of the TMD (Supplementary Fig. 3a–c). Thus, residues P203–N209 most likely constitute a linker region without regular secondary structure between β-strand G of the ECD-D2 and the TMD α-helix, which we term juxtamembrane-linker 1 (JML1). Proceeding the TMD helix, L234–G236 continued in a transient helical turn, while we observed no regular secondary structure in the overlap region between the TMD and the ICD (G236–V240) neither in the data from this study (Supplementary Fig. 3a–c and Fig. 2a) nor from the previous characterization of the ICD[10]. We thus refer to residues Y237–C242 as juxtamembrane-linker 2 (JML2). Hence, we find that the overlap regions between the three isolated domains lacked regular secondary structure and therefore were suitable as assembly sites. We oriented the ECD perpendicular to the membrane surface, as no interactions between hPRLR-ECD-D2 and POPC were observed (Fig. 4a), while the TMD was oriented in accordance with the DHPC-embedment data presented above.

The resulting structure of the monomeric hPRLR provides the first view of the molecular architecture of a full class I cytokine receptor (Figs 4 and 5). Moving from N- to C-terminus, it consists of the ECD (Q1-I202), JML1 (P203-N209), TMD (D210-G236), JML2 (Y237-C242) and ICD (I243-H598). The hPRLR structure is $\sim 345\,Å$ on the vertical axis from N to C terminus, of which the ECD constitute $\sim 20\%$, the TMD $\sim 10\%$ and the ICD $\sim 70\%$.
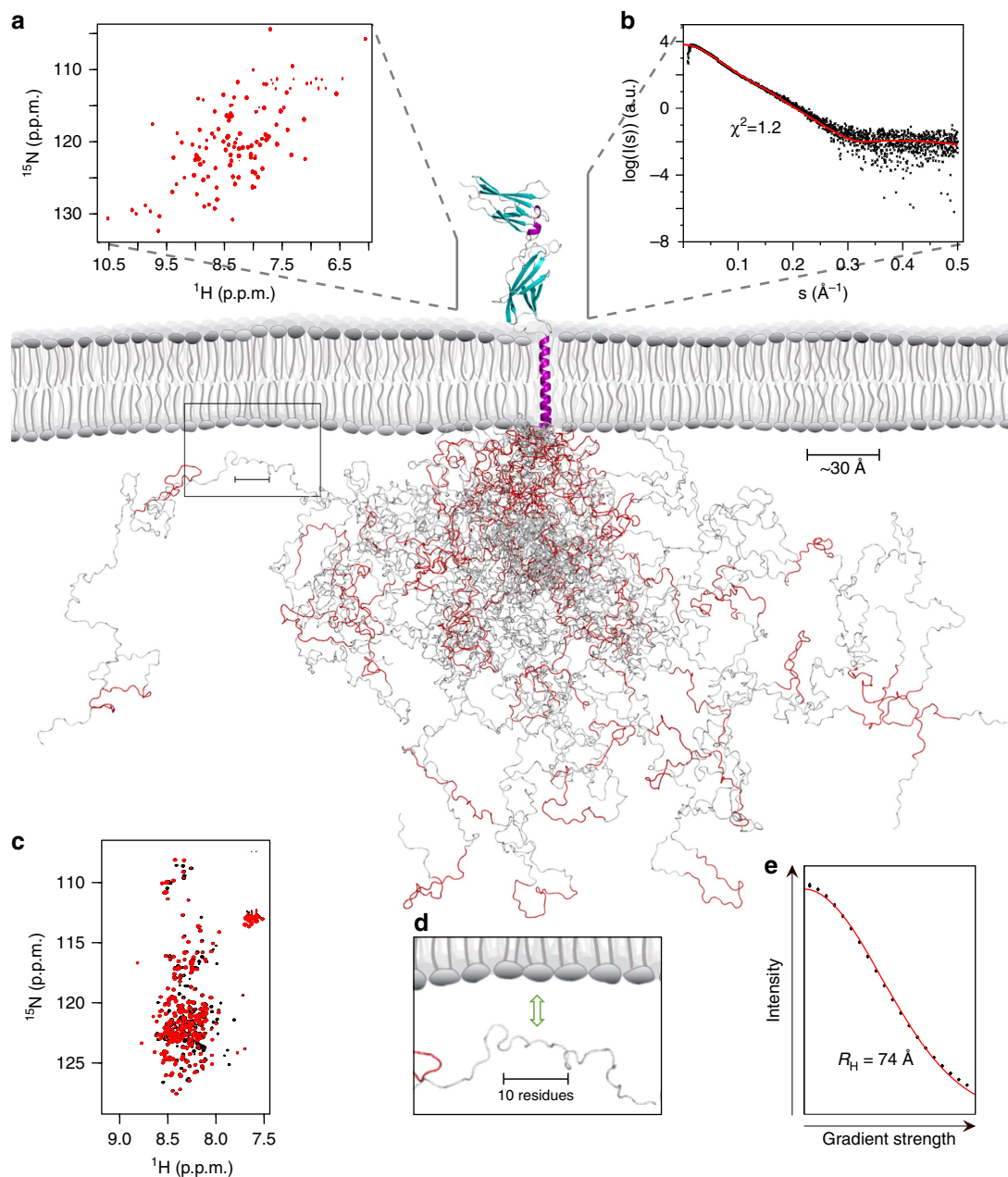
Figure 4 | Structural model of the full hPRLR. The ICD is for clarity represented by 25 models out of the 1,000-member FM ensemble. α-helices are purple, β-sheets cyan and transient α-helices are red. (**a**) ^{1}H-^{15}N-HSQC spectrum of hPRLR-D2 in the presence (red) and absence (black) of POPC SUVs. (**b**) SAXS-scattering curve of monomeric hPRLR-ECD (black) fitted with the coordinates of the liganded hPRLR-ECD (PDB entry 3D48) using CRYSOL (red). (**c**) ^{1}H-^{15}N-HSQC spectrum of hPRLR-ICD in the presence (red) and absence (black) of POPC/POPS SUVs. (**d**) Zoom in on one ICD model. The ICD has an affinity for the inner membrane leaflet[10] indicating that at certain times, dependent on for example the phosphorylation status of the receptor, much of the ICD will be located mainly at the membrane interface, in close vicinity to for example membrane anchored kinases. (**e**) Determination of R_H of hPRLR-ICD by pulse-field gradient NMR. The NMR signal decay of hPRLR-ICD (black) is shown as a function of field strength, together with the corresponding fit (red).

Discussion

In this work we have presented a structural model of the unliganded, monomeric hPRLR-ECD generated from SAXS data. The structure of the liganded ECD[37] fits in detail into the unliganded ECD SAXS envelope, revealing that the relative orientation of D1 and D2 is preserved (Fig. 3a). Thus, in line with findings of time-resolved Förster resonance energy transfer (FRET) studies[46], no substantial conformational changes appear to take place in the PRLR-ECD upon hormone binding. This observation is consistent with findings for the GHR; the crystal structure of the unliganded GHR-ECD[23] shows only minor changes compared to the ligand bound state, pointing towards a signal transduction mechanism that rely on subunit reorientation, as suggested for the GHR[17]. However, there are also studies pointing at differences in the signal transduction mechanism between the GHR and the PRLR. For example, although the ICDs of the receptors share properties such as intrinsic disorder and conserved motifs (Box1 and Box2), they also display interesting differences in their pattern of transient structures and number and position of lipid interaction domains (LIDs)[10]. Further, in
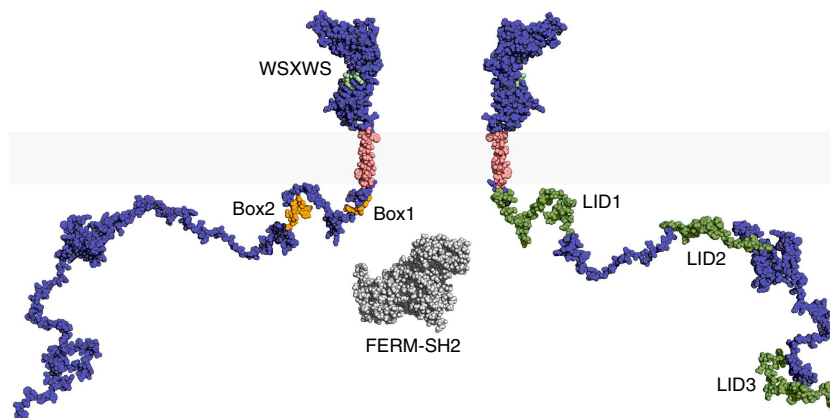
Figure 5 | Molecular architecture of the hPRLR with important regions highlighted. The full hPRLR structural model is shown as spheres from two different angles, with the ICD represented by 1 out of the 1,000 models. The membrane-embedded part of the receptor is coloured pink, while the water-soluble domains are blue. The WSXWS motif is shown in mint green, Box1 and Box2 are orange and the three LIDs are green. For size comparison, the FERM-SH2 domains from TYK2 (PDB entry 4PO6 (ref. 56)) are shown in grey.

alanine-insertion studies of the GHR[23] and the PRLR[24], insertions in the JML2 of the GHR affected receptor activity[23], while analogous insertions in the PRLR had no effect[24]. For the JMLs to function as hinges in the subunit reorientation mechanism they need to have some degree of rigidity. The lack of effect on insertions in the hPRLR-JML2 may be understood from the structural model presented here and previously published data[10], showing that the hPRLR-JML2 is without rigid structure, whereas the corresponding region in the GHR showed propensity for transient α-helical structures. Thus, if JML2 in the PRLR should function as a hinge during activation, the required rigidity must be imposed by other components than the intrinsic structure alone. Overall, these findings suggest that the PRLR and the GHR may share similar activation mechanisms on the extracellular side of the receptor, but vary intracellularly in line with their different biological functions.

The structure of the monomeric hPRLR-TMD presented in this work revealed the extent of the α-helical secondary structure and a characteristic tryptophan symmetry. Tryptophans are preferably positioned towards the ends of bitropic TMD α-helices[47], receive the greatest energy reward for partitioning into the lipid head group region[48], and may act as interfacial anchors that regulate the helix tilt angle[49]. The placement of W214 and W230 in the second helix turn from the N- and C-termini of the TMD, respectively, fits well with a role of these in membrane anchoring. If the PRLR exerts its functions through a similar mechanism as suggested for the GHR[17], the hPRLR-TMD should be able to switch between two different dimer conformations corresponding to the inactive and active states. With this in mind, W214 and W230 may be important not only in terms of controlling the crossing angle of the lowest energy monomer and dimer, but also in rendering a second dimerization interface, with a different crossing angle, less favourable.

Although the TMDs of class I cytokine receptors have been shown to be responsible for receptor homodimerization[16–18], no classical dimerization motifs[50] are present in the hPRLR-TMD and the monomeric form was readily obtained in this work even at high concentrations. Thus, hPRLR-TMD$_{206-240}$ appears to have a weak inherent propensity to homodimerize. In fact, the lack of classical dimerization motifs is shared by the related hEPOR-TMD, and its structures have also only been solved in monomeric forms[51,52]. This weak inherent dimerization propensity suggests that other components than those of the hPRLR-TMD$_{206-240}$/DHPC system may be important to drive TMD dimerization. Potential factors include the adjacent domains, membrane interaction partners such as specific lipid components (for example, phosphoinositides, cholesterols etc.) or accessory bound proteins (for example, JAK2). The polar residues in the core of the hPRLR-TMD α-helix (S216, S221 and C225) may be of possible relevance to homodimerization. In TMDs, such side chains are often involved in inter-monomeric hydrogen bonds, improving the stability and specificity of α-helical associations with one of the most common participants being serine[53]. Interestingly, the least common amino acid type in TM regions is cysteine[54], suggesting that C225, placed deep within the TMD, might play a special role. A cellular mutagenesis study on the hPRLR has shown that substituting C225 with a serine decreased ligand-independent dimerization by ∼30% (ref. 18), suggesting that C225 takes part in TMD dimerization, without being vital. In hPRLR-TMD, S221 and C225 are positioned in two adjacent α-helical turns, while S216 is positioned half a turn away (Fig. 1f), suggesting that the hPRLR has the potential to form two different dimerization interfaces involving serines. These two could be interchangeable through rotation of the TMDs during receptor activation, perhaps with C225 as pivot point.

The structures of the monomeric TMDs from the related human and mouse (m) EPOR were recently solved in *n*-dodecylphosphocholine (DPC) micelles on the basis of dihedral angles obtained from TALOS using C$^{\alpha}$ chemical shifts only, a modest number of NOEs and backbone hydrogen bond restraints[51,52]. Due to a lack of restraints on the side chains of these structures, we regard them mainly as backbone structures, represented by a straight α-helix spanning residues L226-W258 for both species. Surprisingly, these α-helices continue nine residues into the JM-region without any apparent flexibility, having important implications for the mode of hormone-binding-induced signal propagation from the TMD to Box1. We note that no non-sequential NOEs for H249-L253 (hEPOR-TMD) or around H249 (mEPOR-TMD) or other data appear to support this α-helix extension[51,52]. In contrast, relaxation data for both EPOR-TMDs reveal increased internal dynamics from H249–P254, supporting a less-structured conformation not captured by the structures[51,52]. Thus, if these regions instead are interpreted as forming transient helical structures, these related structures have a similar overall topology as hPRLR-TMD.

Inherent limitations in sample preparation and individual techniques for structure determination make atomic-resolution studies of the structurally diverse hPRLR challenging. We therefore applied an approach in which data were integrated from multiple structural disciplines combining new SAXS, MS

and NMR data with previously published data and molecular modelling to acquire the molecular architecture of the full hPRLR (Figs 4 and 5). This model provides the first molecular architecture of a full-length class I cytokine receptor, revealing the relative sizes of the individual domains. The hPRLR model consists of the soluble, globular ECD (Q1–I202), the unstructured JML1 (P203–N209), the membrane-embedded TMD (D210–G236), the unstructured JML2 (Y237–C242) and the intrinsically disordered ICD (I243–H598). The ICD was in a previous study shown to contain five transient α-helices (Fig. 4, red stretches) and three non-cooperative LIDs along the chain[10], which have not been directly included in the model, but are shown in Fig. 5 (green).

The full hPRLR structure reveals that the ECD, which until recently was the only structurally characterized part of the class I cytokine receptors, merely constitutes the tip of an extensive molecular iceberg. Previously, visualizations of this receptor family have typically shown a disproportionally dominating ECD compared to the TMD and ICD, likely reflecting the overall information content available for each domain. From our structure it is now evident that the ICD is more than twice as extended in the direction of the membrane normal compared to the ECD and the TMD combined (Fig. 4). The structural flexibility and large capture radius of the ICD allows it to reach and interact with a variety of interaction partners, possibly also other receptors' ICDs. Interaction sites in disordered regions typically only constitute a few residues, so-called small linear motifs (SLiMs), and hence the ICD is geared to simultaneously interact with many kinases, phosphatases and other proteins. Box1 is an example of such a SLiM in the PRLR-ICD, shown to interact with JAK2 (ref. 13), most likely through its FERM (4.1, Ezrin, Radixin, Moesin) domain[55]. To illustrate the relative sizes of these proteins, Box1 along with the FERM-SH2 domains from TYK2 (ref. 56), a JAK2 homologue, are highlighted in Fig. 5.

The generated ensemble model of the ICD represents its unbound intrinsic structure, represented by an ensemble of 25 conformers in Fig. 4. However, in the cell the structural ensemble of the ICD may be different in several ways. First, of all the three LIDs[10] have not been restricted to interact with the membrane in this model. Since experimental measurements of the R_H in the presence of SUVs would be dominated by signals from the SUVs, this was not attempted, but it is likely that the LID-mediated membrane contacts would result in a smaller R_H of the ICD. Thus, we regard the experimentally determined R_H obtained in the absence of SUVs as an upper limit capture radius for the ICD. Second, constitutively bound kinases have deliberately been omitted, primarily because their mutual binding sites have not been adequately described, if at all known. Last, post-translational modifications of the ICD, in particular phosphorylations, may have the potential to change the structural ensemble of the ICD[57] and thus its compactness. Importantly, there are no indications to suggest the ICD to be folded, and the ICD does not intrinsically dimerize (Fig. 2d). Hence, other proteins or modifications such as acetylation and phosphorylations would be needed to promote ICD dimerization, in essence completely analogous to the hormone-induced dimerization of the ECD-D2 domains.

The presented structural model of the hPRLR provides important new insights on the full structure of class I cytokine receptors, and provides a framework for understanding the mechanisms related to these receptors, as for example, derived from cellular studies. However, being based on a divide-and-conquer approach, it still poses unresolved questions. First, due to the methodological restrictions on structural characterization of a PRLR variant that includes both the ECD and TMD, the exact relative orientation of these domains remains unestablished. However, in the alanine-insertion study by Liu and

Brooks[24], insertions in PRLR-JML1 did not affect receptor functionality, suggesting that the relative orientation between the ECD and the TMD is not essential. Second, it calls for reservations that the TMD structure was solved in detergent rather than a bilayer. Consequently, the angle between the TMD and the bilayer plane (helix tilt angle) remains speculative. Furthermore, although the structural envelope was obtained in the present work and the unliganded ECD-D2 structure is available[12], a high-resolution structure of the unliganded ECD is missing. Last, the model does not include any potential interactions with the glycosaminoglycan layer, suggested to implicate the WS-motif[11], or effects from post-translational modifications.

In conclusion, our head to toe structural model of the hPRLR provides a starting point for future refinements and may help design strategies for novel structural and functional studies. Importantly, it exemplifies the architecture of the many biologically fundamental receptors of the class 1 cytokine receptor family, and specifies a scaffold onto which a new view on cellular signalling can be built.

Methods

Materials. DHPC, DPC, 1-myristoyl-2-hydroxy-sn-glycero-3-phospho-(1′-rac-glycerol) (LMPG), 1-palmitoyl-2-hydroxy-sn-glycero-3-phospho-(1′-rac-glycerol) (LPPG), 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (sodium salt) (POPS) were purchased from Avanti Polar Lipids (Alabaster, AL). 2,2-didecylpropane-1,3-bis-β-D-maltopyranoside (MNG-3) and Amphipol A8–35 were purchased from Anatrace (Maumee, OH). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) and N-lauroylsarcosine sodium salt (sarkosyl) were purchased from Sigma-Aldrich (St Louis, MO). Sodium dodecyl sulphate (SDS) was purchased from Avantor Performance Materials (Deventer, NL). Dihexanoyl-d22-phosphatidylcholine-d9 (d₉DHPC) was purchased from FB Reagents (Boston, MA).

Protein expression and purification. Complementary DNA encoding the full-length long isoform of the human PRLR was amplified from IRAKp961G13133Q (RZPD German Resource Center for Genome Research).

A detailed description of the expression and purification of hPRLR-D2$_{M99–D210}$ can be found in ref. 12, but in brief, the sequence of hPRLR-D2 (M99-D210) was sub-cloned into a pET11a vector (Novagen), transformed into competent E. coli BL21(DE3) cells, and expressed in ^{15}N-labelled M9-media (3 g per l KH₂PO₄, 7.5 g l^{-1} Na₂HPO₄ · H₂O, 5 g l^{-1} NaCl, 1 mM MgSO₄, 1 ml M2 trace solution, 4 g l^{-1} glucose, 1.5 g l^{-1} (^{15}NH₄)₂SO₄, (ISOTEC)) added 100 µg ml^{-1} ampicillin. Following collection, cells were resuspended and sonicated on ice to bring hPRLR-D2 in solution. After ammonium sulphate (AMS) precipitation (75% (w/v), stirring for 2 h at 0 °C), the precipitate was dissolved in 30 mM NH₄HCO₃, 100 mM NaCl and 1 mM dithiothreitol (DTT) (pH 8.0) and applied to a Sephadex G50-Fine (GE Healthcare) column (24 × 370 mm, 167 ml). hPRLR-D2 fractions were pooled, concentrated and buffer exchanged.

The sequence of hPRLR-ECD (Q1-D210) was sub-cloned into a pET11a vector (Novagen), transformed into competent E. coli BL21(DE3) cells, and grown in LB media (10 g per l peptone, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl, pH 7.4) at 37 °C and 180 r.p.m. Cells were induced at OD₆₀₀ = 0.8 with 1 mM isopropyl β-D-1-thiogalactopyranoside for 4 h and subsequently collected by centrifugation (5,000g, 15 min, 4 °C). The cell pellet was resuspended in lysis buffer (40 ml l^{-1} culture 25% (w/v) sucrose, 5 mM EDTA, 1 × PBS buffer (pH 7.4), 1% (w/v) Triton X-100), sonicated on ice (2 × 3 min with 3 min rest between rounds at 50% amplitude), and the inclusion bodies (IBs) collected by centrifugation (20,000g, 25 min, 4 °C). This cycle was repeated a total of three times. The IBs were denatured in 6 M Urea, 10 mM β-mercaptoethanol, 50 mM Tris–HCl (pH 9.0), followed by refolding by simple dialysis into 50 mM Tris–HCl (pH 9.0), 150 mM NaCl, 10 mM cysteamin, 1 mM cystamin (3 × 4 l), protein concentration < 0.1 mg ml^{-1}. The refolded fraction was precipitated by addition of AMS to 75% (w/v), and left for 2 h with gentle stirring at room temperature. The precipitate was dissolved in MilliQ water and gel-filtered on a Sephadex G50-Fine column (24 × 370 mm, 167 ml) (GE Healthcare) in 0.1 M NaCl, 1 mM DTT, 30 mM NH₄HCO₃ (pH 9.0). Peak fractions were dialyzed against 10 mM Na₂HPO₄ (pH 7.4).

A detailed description of the expression and purification of hPRLR-TMD$_{F206–V240}$ can be found in ref. 10, but in brief, the sequence of hPRLR-TMD (F206–V240) was sub-cloned into a pGEX-4T-1 vector (Amersham, GE Healthcare), containing an N-terminal glutathione S-transferase-carrier protein and a thrombin cleavage site. The plasmid DNA was transformed into competent E. coli BL21(DE3) cells and expressed in either unlabelled, ^{15}N or ^{13}C-, ^{15}N-labelled M9 media (if ^{13}C-labelled: ^{13}C-D-Glucose, ISOTEC, if ^{15}N-labelled: (^{15}NH₄)₂SO₄, ISOTEC) added 100 µg ml^{-1} ampicillin. The collected cells were resuspended in lysis buffer (40 ml l^{-1} culture

25% (w/v) sucrose, 5 mM EDTA, 1 × PBS buffer (pH 7.4), 1 mM PMSF), and sonicated on ice. Subsequently the IBs were collected by centrifugation (20,000g, 25 min, 4 °C). This cycle of resuspension, sonication and centrifugation was repeated three times. The resulting IBs were resuspended in 50 mM Tris–HCl buffer, collected by centrifugation (20,000g, 20 min, 4 °C), solubilized in 12 ml (per l culture) 1.5% (w/v) sarkosyl, 100 mM DTT, 20 mM Tris–HCl buffer (pH 7.4) and incubated at room temperature with gentle agitation for 3 h. Insoluble material was removed by centrifugation (12,000g, 20 min, 4 °C). The supernatant was dialyzed against 0.5% (w/v) sarkosyl, 10 mM NaCl, 50 mM Tris–HCl buffer (pH 7.4) to remove DTT and cleaved with thrombin to release the glutathione S-transferase carrier protein. After cleavage the solution was lyophilized, followed by resuspension in milliQ water (200 µl ml^{-1} of original solution). This solution was divided into batches of 50 µl, each of which was added to 750 µl of a 1:2 chloroform:methanol solution and mixed well. The solution was centrifuged (14,000g, 2 min, 4 °C), resulting in separation in three layers. The top aqueous layer was carefully removed. Subsequently, 500 µl of MeOH was added to the remaining solution, followed by thorough mixing. The mixture was incubated on ice for 20 min, followed by centrifugation (16,000g, 40 min, 4 °C). The supernatant containing the target protein was transferred to a glass vial, and the organic solvent evaporated under a stream of N$_2$.

A detailed description of the expression and purification of hPRLR-ICD$_{G236–H598}$ and hPRLR-ICD$_{G236–Q396}$ may be found in ref. 27, but in brief, G236–H598 and G236–Q396 of hPRLR were sub-cloned into pET11a vectors (Novagen). The plasmids were transformed into competent E. coli BL21(DE3) cells and protein expressed in either unlabelled, ^{15}N or ^{13}C-, ^{15}N-labelled M9 media added 100 µg ml^{-1} ampicillin. Cells were collected by centrifugation (20 min, 5,000g, 4 °C) and stored at −20 °C until thawed on ice and resuspended in 40 ml sonication buffer (20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ pH 8, 300 mM NaCl, 0.08% (w/v) Triton X-100 and one complete EDTA-free protease inhibitor cocktail tablet (Roche Diagnostics GmbH)). The cells were sonicated on ice (4 × 30 s with 30 s rest between rounds at 100% amplitude), centrifuged (25 min, 20,000g, 4 °C), and the supernatants used for purification.

After sonication the supernatant containing hPRLR-ICD$_{G236–H598}$ was heated for 5 min at 95 °C, incubated on ice for 10 min, and centrifuged to remove precipitate (10 min, 20,000g, 4 °C). The supernatant containing hPRLR-ICD$_{G236–H598}$ was added 10 mM DTT, precipitated to 35% with AMS on ice, gently stirred and incubated on ice for 2 h before centrifugation (20 min, 20,000g, 4 °C). The pellet was resuspended in 20 ml 20 mM Tris–HCl (pH 8.0) and dialyzed against 1 l 20 mM Tris–HCl (pH 8.0) at 4 °C before applying the sample to a 5 ml HiTrap Q Sepharose Fast Flow (QFF) (Amersham Bioscience) column pre-equilibrated with 5 column volumes solution buffer (20 mM Tris–HCl (pH 8.0), 0.1 mM DTT). hPRLR-ICD$_{G236–H598}$ was eluted over a linear gradient of 20 column volumes with elution buffer (20 mM Tris–HCl (pH 8.0), 1 M NaCl, 0.1 mM DTT).

In case of hPRLR-ICD$_{G236–Q396}$, DNA was precipitated with a final concentration of 0.1 % (v/v) protamine sulphate added on ice, gently stirred and incubated on ice for 10 min before centrifugation (50 min, 37,000g, 4 °C). The supernatant was heated for 5 min at 95 °C, incubated on ice for 10 min, and centrifuged to remove the precipitate (10 min, 20,000g, 4 °C). The supernatant containing hPRLR-ICD$_{G236–Q396}$ was precipitated to 35% with AMS on ice, gently stirred and incubated on ice for 2 h before centrifugation (20 min, 20,000g, 4 °C). The pellet was resuspended in 20 ml 20 mM Tris–HCl (pH 8.0) and dialyzed against 1 l 20 mM Tris–HCl (pH 8.0) at 4 °C before applying the sample to a 5 ml HiTrap QFF column using a protocol similar to that of hPRLR-ICD$_{G236–H598}$.

Screening of membrane mimetics. In case of SDS, DPC, SDS:DPC, sarkosyl, DHPC, LMPG and LPPG ∼150 nmol ^{15}N-hPRLR-TMD was reconstituted in the membrane mimetics of choice in 50 mM NaCl, 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ (pH 7.2), followed by thorough buffer exchange in a −3 kDa cutoff spinfilter with 50 mM NaCl, 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ to remove residuals. When testing the two amphipols A8–35 and MNG-3, 150 nmol ^{15}N-hPRLR-TMD was solubilized in 0.3% (w/v) sarkosyl mixed with either 800 nmol MNG-3 or 600 nmol A8–35. Sarkosyl was subsequently removed by thorough buffer exchange in a −3 kDa cutoff spinfilter with 50 mM NaCl, 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ (pH 7.2). HFIP was tested by adding 60% HFIP/40% H$_2$O directly to ∼150 nmol ^{15}N-hPRLR-TMD followed by thorough mixing. DHPC/DMPC-, DHPC/POPC- and DHPC/POPC/POPS (3:1) bicelles were prepared by solubilizing ^{15}N-hPRLR-TMD in DHPC, followed by thorough buffer exchange in a −3 kDa cutoff spinfilter with 50 mM NaCl, 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$. Subsequently, ((n_{DHPC} − n_{CMC})/0.3) mol solubilized lipids were added to the ^{15}N-hPRLR-TMD/DHPC mixture to give $q = 0.3$ (long-to-short-chain lipid ratio). The TMD/DHPC/lipid mixtures were then subjected to freeze/thaw cycles until the solution cleared. The samples were prepared at 5% or 15% (w/v) lipid content. All samples were added 10% (v/v) D$_2$O, 2 mM tris(2-carboxyethyl)phosphine (TCEP), 1 mM 2,2-dimethyl-2-silapentane-5-sulphonic acid (DSS), 0.05% NaN$_3$ and 50 mM NaCl, 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ to a final volume of 370 µl followed by pH-adjustment to 7.2 before being transferred to a Shigemi NMR tube. ^{1}H-^{15}N-HSQC spectra of hPRLR-TMD reconstituted in the various membrane mimetics were recorded on a Varian INOVA 800-MHz (^{1}H) spectrometer at temperatures between 25 and 42 °C.

Native mass spectrometry. A sample containing 25 µM hPRLR-TMD, 700 × molar excess DHPC and 200 mM ammonium acetate buffer was prepared for native MS by extensive buffer exchange in −3 kDa spinfilter with 200 mM ammonium acetate buffer (pH 7). Samples were loaded into gold coated nano-electrospray emitters prepared in-house as previously described[58]. A range of dilutions (using milliQ water) were prepared for MS analysis, ranging from 10 to 100 µM, maintaining the 1:700 hPRLR-TMD:DHPC ratio.

NMR spectroscopy on hPRLR-TMD. All NMR spectra were recorded on Varian INOVA 750- or 800-MHz (^{1}H) spectrometers with room temperature probes or a Bruker 900-MHz (^{1}H) with a 5 mM CPTCI probe. Free induction decays were transformed and visualized in NMRPipe[59] and analyzed using the CcpNmr Analysis software[60]. Unless otherwise specified, all hPRLR-TMD samples were recorded at 37 °C and contained 10% (v/v) D$_2$O, 4 mM TCEP, 2 mM DSS, 0.05% (v/v) NaN$_3$, 50 mM NaCl and 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ (pH 7.2). Proton chemical shifts were referenced internally to DSS at 0.00 p.p.m., with heteronuclei referenced by relative gyromagnetic ratios.

For assignments of backbone nuclei, heteronuclear NMR spectra were recorded on a sample containing 0.8 mM ^{13}C-,^{15}N-hPRLR-TMD in 560 mM DHPC. Backbone assignments were performed manually from the analyses of ^{1}H-^{15}N-HSQC, HNCACB, CBCA(CO)NH and HNCO spectra acquired with non-uniform sampling[61]. The backbone chemical shifts were used in TALOS[32] to estimate the dihedral angle restraints, and in the motif identification from chemical shifts (MICS) programme[31] to identify possible motifs. NOE assignments were performed manually from analysis of a ^{15}N-NOESY-HSQC spectrum (mixing time of 100 ms) and ^{13}C-NOESY-HSQC spectra of the aliphatic region (mixing time of 150 ms) and the aromatic region (mixing time of 150 ms), acquired on a sample containing 1 mM ^{13}C-,^{15}N-hPRLR-TMD in 700 mM d$_9$DHPC.

A 3D HNHA spectrum[33] was recorded on 0.7 mM ^{15}N-hPRLR-TMD in 490 mM DHPC and ^{3}J(H^N-H$^\alpha$) couplings constants were extracted from the relative intensity of H$^\alpha$ and H^N peaks using the CcpNmr Analysis software[60]. The ^{3}J(H^N-H$^\alpha$) coupling constants were utilized to estimate the backbone dihedral ϕ-angles from the Karplus relationship with coefficient values of 6.51, −1.75 and 1.60 for A, B, and C, respectively[33]. For amides with coupling constants below <5 Hz along with NOE patterns characteristic of α-helical conformation, hydrogen bond restraints were created between H$_i$(N) and (C)O$_{i-4}$.

A series of ^{1}H-^{15}N-HSQC spectra were recorded on a sample containing 0.7 mM ^{15}N-hPRLR-TMD in 490 mM DHPC to analyze the decay of the transverse relaxation (T_2). T_2 relaxation times were calculated from standard HSQC spectra recorded at 800 MHz using seven different relaxation delays between 10 and 130 ms. The relaxation decays were fitted to single exponentials and relaxation times calculated using the CcpNmr Analysis software[60].

Hydrogen-to-deuterium (H–D) exchange experiments were performed on samples containing 0.4 mM ^{15}N-hPRLR-TMD in 300 mM DHPC on the basis of the principles of Veglia et al.[62] Samples were lyophilized followed by resolubilization in 10, 30, 40, 50, 70 or 90% (v/v) D$_2$O. For each sample, a ^{1}H-^{15}N-HSQC spectrum was acquired after an incubation period of 1 h. Since the peak intensities did not change significantly between 1 and 5 h, a 1-h incubation period was deemed sufficient for reaching a quasi-stationary state.

Structure calculations of hPRLR-TMD. The assigned NOE peaks, dihedral angles and hydrogen bonding restraints obtained for hPRLR-TMD were applied in a standard simulated annealing protocol using Aria2 (version 2.3.2)[34] and CNS[35]. NOE peak intensities were calibrated and converted to inter-proton distances by Aria2 during each iteration step using a distance cutoff of 6 Å. Each run consisted of eight iterations, with 20 structures calculated in each of the first seven. The seven structures with lowest global energy were used as starting structures in the subsequent iteration. The structure calculations were evaluated using the CcpNmr Analysis software[60] and the CING suite[63]. Iteratively, assignments were checked manually, modified if needed, and structures recalculated. In the final iteration, 200 structures were generated of which the 10 lowest energy structures without significant violations (Table 1) were selected to represent the monomeric structure of hPRLR-TMD in DHPC micelles. The structures were visualized in PyMOL (DeLano Scientific). Ramachandran-plot statistics for the structure ensemble (residues 209–235) were calculated with PROCHECK[64] and are as follows: most favored (95.2%), additionally allowed (4.4%), generously allowed (0.4%) and disallowed (0.0%).

Lipid interaction studies of hPRLR-D2. ^{1}H-^{15}N-HSQC spectra were acquired on samples containing 50 µM ^{15}N-hPRLR-ECD-D2 in 10 mM Na$_2$HPO$_4$, 10 mM TCEP, 2 mM DSS and 10% D$_2$O (pH 7.4) and no POPC SUVs or POPC SUVs at a final concentration of 8 mM. Chemical shift differences and intensity ratios were compared using the CcpNmr Analysis software[60] to investigate possible interactions between hPRLR-ECD-D2 and POPC.

SAXS and envelope generation. SAXS data on the monomeric hPRLR-ECD$_{Q1–D210}$ were recorded at the HESYLAB synchrotron in Hamburg, beam line X33. Scattering was recorded at three different protein concentrations (1.15, 2.24 and 4.38 mg ml^{-1} in 10 mM Na$_2$HPO$_4$ (pH 7.4)). The three scattering curves were recorded in succession flanked by recordings of the buffer background. A high-quality scattering curve was constructed by merging the low-concentration

data for low-scattering angles, intermediate concentration for intermediate-scattering angles and data from the high concentration experiment for the highest angle part of the scattering curve. The background sample consisted of the last (pure protein sample) dialysis buffer after dialysis was completed. Data were processed using the ATSAS package[65]. The theoretical scattering curve of the ECD model (on the basis of PDB entry 3D48 (ref. 37)) was fitted to the experimentally determined envelope ($\chi^2 = 1.22$) (superposition performed in CRYSOL[39], part of the ATSAS package[65]) and the structure docked as a rigid body into the 3D density map by using the fit-in-map function from the UCSF CHIMERA[66].

Relaxation measurements on hPRLR-ICD$_{G236-Q396}$. Two series of ^{1}H-^{15}N-HSQC spectra were recorded on ^{15}N-hPRLR-ICD$_{G236-Q396}$ to analyze the T_1 and T_2 relaxation times. ^{1}H-^{15}N-HSQC spectra were recorded at 750 MHz (^{1}H) and 4°C with delay times between 10 and 1,000 ms (T_1) and 10–250 ms (T_2) with two triplicate measurements for each series. The relaxation decays are fitted to single exponentials and relaxation times determined using the CcpNmr Analysis software[60].

Oligomeric state and hydrodynamic radius of hPRLR-ICD$_{G236-H598}$. ^{1}H-^{15}N-HSQC spectra were recorded on ^{15}N-hPRLR-ICD$_{G236-H598}$ at 5°C in 10% (v/v) D$_2$O, 8 mM TCEP, 0.5 mM DSS and 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ (pH 7.3), at high (400 μM) and low (20 μM) concentrations. Chemical shifts were compared using the CcpNmr Analysis software[60] to investigate the possibility of oligomerization.

The hydrodynamic radius of hPRLR-ICD$_{G236-H598}$ was determined by PGSLED NMR diffusion experiments using the pulse sequence of ref. 67. The experiments were performed on 800 μM ^{15}N-hPRLR-ICD in 20 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, 8 mM TCEP, 0.5 mM DSS in 90% (v/v) D$_2$O (pH 7). As reference, 1.5 mM α-cyclodextrin under identical buffer and experimental conditions was used. All spectra were recorded at 5°C on a Varian Inova 750 MHz (^{1}H) spectrometer. R_H of hPRLR-ICD was calculated from the relative diffusion decays of hPRLR-ICD and α-cyclodextrin[67], which has a R$_H$ of 7.52 Å (ref. 68).

Generation of FM ensemble. An ensemble of 1,000 models of the intrinsically disordered hPRLR-ICD region (G236–H598) was generated with Flexible Meccano[40] using default options and without any restraints. We used the HYDROPRO 10 (ref. 69) to predict the hydrodynamic properties (in particular the hydrodynamic radius R_H) of each of the members of the ensemble. These values were averaged (as $< R_H^{-1} > ^{-1}$) to estimate the value of R_H our hPRLR-ICD ensemble. The software was used with default options, and in short describes the hydrodynamic properties of proteins by modelling the protein as a set of overlapping spheres that in turn results in a shell-model of the protein[69].

Structural model of the full hPRLR. The full structural model was assembled through joining the individual domains at the overlapping sequences of the structures. The overlapping regions were aligned and 1,000 model templates were constructed from the three domains by building the longest model through multiple independent cycles of refinement in Modeller 9.15 (ref. 70). As starting structures we used (1) a model of the unliganded ECD, which we built from the X-ray structure of hPRLR-ECD in complex with prolactin (PDB entry 3MZG[37], removing the prolactin before the run) using Modeller 9.15., and which was validated by comparing calculated and experimental SAXS curves using CRYSOL[39], (2) the NMR structure of hPRLR-TMD in micelles and (3) the FM ensemble of hPRLR-ICD, which was validated by comparing calculated and experimental R_H values as described above. Each of the 1,000 assembled models were subsequently refined by a Modeller routine through multiple cycles of conjugate gradient optimization (up to 100 steps each) optimizing the model with the variable target function method, followed by molecular dynamics with simulated annealing and a final optimization with conjugate gradients using the 'refine.slow' option of Modeller[70]. The ECD was oriented perpendicular to the membrane surface, as no affinity for POPC lipids was observed, while the TMD was embedded in the sketch membrane in accordance with the DHPC-embedment data. Finally, we discarded all conformations of the full-length hPRLR ensemble model where the ICD folded back into regions occupied by the bilayer.

References

1. Bole-Feysot, C., Goffin, V., Edery, M., Binart, N. & Kelly, P. A. Prolactin (PRL) and its receptor: actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. *Endocr. Rev.* **19,** 225–268 (1998).
2. Ben-Jonathan, N., Hugo, E. R., Brandebourg, T. D. & LaPensee, C. R. Focus on prolactin as a metabolic hormone. *Trends Endocrinol. Metab.* **17,** 110–116 (2006).
3. Ormandy, C. J. et al. Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. *Genes Dev.* **11,** 167–178 (1997).
4. Gillam, M. P., Molitch, M. E., Lombardi, G. & Colao, A. Advances in the treatment of prolactinomas. *Endocr. Rev.* **27,** 485–534 (2006).
5. Bernard, V., Young, J., Chanson, P. & Binart, N. New insights in prolactin: pathological implications. *Nat. Rev. Endocrinol* **11,** 265–275 (2015).
6. Carver, K. C., Arendt, L. M. & Schuler, L. A. Complex prolactin crosstalk in breast cancer: new therapeutic implications. *Mol. Cell. Endocrinol.* **307,** 1–7 (2009).
7. Clevenger, C. V., Gadd, S. L. & Zheng, J. New mechanisms for PRLr action in breast cancer. *Trends Endocrinol. Metab.* **20,** 223–229 (2009).
8. Sackmann-Sala, L. & Goffin, V. in *Recent Advances in Prolactin Research*. (ed. Diakonova, M.), 221–242. (Springer, 2015).
9. Liongue, C. & Ward, A. C. Evolution of class I cytokine receptors. *BMC Evol. Biol.* **7,** 1–15 (2007).
10. Haxholm, G. W. et al. Intrinsically disordered cytoplasmic domains of two cytokine receptors mediate conserved interactions with membranes. *Biochem. J.* **468,** 495–506 (2015).
11. Olsen, J. G. & Kragelund, B. B. Who climbs the tryptophan ladder? On the structure and function of the WSXWS motif in cytokine receptors and thrombospondin repeats. *Cytokine Growth Factor Rev.* **25,** 337–341 (2014).
12. Dagil, R. et al. The WSXWS motif in cytokine receptors is a molecular switch involved in receptor activation: Insight from structures of the prolactin receptor. *Structure* **20,** 270–282 (2012).
13. Lebrun, J.-J., Ali, S., Ullrich, A. & Kelly, P. A. Proline-rich sequence-mediated Jak2 association to the prolactin receptor is required but not sufficient for signal transduction. *J. Biol. Chem.* **270,** 10664–10670 (1995).
14. Goupille, O., Daniel, N., Bignon, C., Jolivet, G. & Djiane, J. Prolactin signal transduction to milk protein genes: carboxy-terminal part of the prolactin receptor and its tyrosine phosphorylation are not obligatory for JAK2 and STAT5 activation. *Mol. Cell. Endocrinol.* **127,** 155–169 (1997).
15. DaSilva, L., Howard, O. M., Rui, H., Kirken, R. A. & Farrar, W. L. Growth signaling and JAK2 association mediated by membrane-proximal cytoplasmic regions of prolactin receptors. *J. Biol. Chem.* **269,** 18267–18270 (1994).
16. Gadd, S. L. & Clevenger, C. V. Ligand-independent dimerization of the human prolactin receptor isoforms: functional implications. *Mol. Endocrinol.* **20,** 2734–2746 (2006).
17. Brooks, A. J. et al. Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* **344,** 1249783–1249783 (2014).
18. Qazi, A. M., Tsai-Morris, C.-H. & Dufau, M. L. Ligand-independent homo- and heterodimerization of human prolactin receptor variants: inhibitory action of the short forms by heterodimerization. *Mol. Endocrinol.* **20,** 1912–1923 (2006).
19. Broutin, I. et al. Crystal structure of an affinity-matured prolactin complexed to its dimerized receptor reveals the topology of hormone binding site 2. *J. Biol. Chem.* **285,** 8422–8433 (2010).
20. van Agthoven, J. et al. Structural characterization of the stem-stem dimerization interface between prolactin receptor chains complexed with the natural hormone. *J. Mol. Biol.* **404,** 112–126 (2010).
21. Jomain, J.-B. et al. Structural and thermodynamic bases for the design of pure prolactin receptor antagonists: X-ray structure of Del1-9-G129R-hPRL. *J. Biol. Chem.* **282,** 33118–33131 (2007).
22. Teilum, K. et al. Solution structure of human prolactin. *J. Mol. Biol.* **351,** 810–823 (2005).
23. Brown, R. J. et al. Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nat. Struct. Mol. Biol.* **12,** 814–821 (2005).
24. Liu, W. & Brooks, C. Functional impact of manipulation on the relative orientation of human prolactin receptor domains. *Biochemistry* **50,** 5333–5344 (2011).
25. Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. *Nature* **405,** 647–655 (2000).
26. Cherezov, V. et al. High-resolution crystal structure of an engineered human b2-adrenergic G protein-coupled receptor. *Science* **318,** 1258–1265 (2007).
27. Bugge, K., Steinocher, H., Brooks, A. J., Lindorff-Larsen, K. & Kragelund, B. B. Exploiting hydrophobicity for efficient production of transmembrane helices for structure determination by NMR spectroscopy. *Anal. Chem.* **87,** 9126–9131 (2015).
28. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl Acad. Sci. U.S.A.* **106,** 1760–1765 (2009).
29. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **8,** 639–651 (2013).
30. Schwarzinger, S. et al. Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* **123,** 2970–2978 (2001).
31. Shen, Y. & Bax, A. Identification of helix capping and beta-turn motifs from NMR chemical shifts. *J. Biomol. NMR* **52,** 211–232 (2012).
32. Shen, Y. & Bax, A. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* **38,** 289–302 (2007).

33. Vuister, G. W. & Bax, A. Quantitative J correlation—a new approach for measuring homonuclear 3-bond J(H(N)H(Alpha) coupling-constants in 15N-enriched proteins. *J. Am. Chem. Soc.* **115,** 7772–7777 (1993).

34. Rieping, W., Bardiaux, B., Bernard, A., Malliavin, T. E. & Nilges, M. ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23,** 381–382 (2007).

35. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921 (1998).

36. Svensson, L. A. *et al.* Crystal structure of a prolactin receptor antagonist bound to the extracellular domain of the prolactin receptor. *J. Biol. Chem.* **283,** 19085–19094 (2008).

37. Kulkarni, M. V. *et al.* Two independent histidines, one in human prolactin and one in its receptor, are critical for pH-dependent receptor recognition and activation. *J. Biol. Chem.* **285,** 38524–38533 (2010).

38. Kossiakoff, A. A. *et al.* Ternary complex between placental lactogen and the extracellular domain of the prolactin receptor. *Nat. Struct. Biol* **7,** 808–815 (2000).

39. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr* **28,** 768–773 (1995).

40. Ozenne, V. *et al.* Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. *Bioinformatics* **28,** 1463–1470 (2012).

41. Weidemann, T., Höfinger, S., Müller, K. & Auer, M. Beyond dimerization: a membrane-dependent activation model for interleukin-4 receptor-mediated signalling. *J. Mol. Biol.* **366,** 1365–1373 (2007).

42. van Meer, G. & de Kroon, A. I. P. M. Lipid map of the mammalian cell. *J. Cell Sci.* **124,** 5–8 (2011).

43. Fernández, C., Hilty, C., Wider, G. & Wüthrich, K. Lipid-protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy. *Proc. Natl Acad. Sci. U.S.A.* **99,** 13533–13537 (2002).

44. Kučerka, N., Nieh, M. P. & Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta—Biomembr.* **1808,** 2761–2771 2011.

45. The Uniprot Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **43,** D204–D212 (2015).

46. Tallet, E. *et al.* Investigation of prolactin receptor activation and blockade using time-resolved fluorescence resonance energy transfer. *Front. Endocrinol. (Lausanne)* **2,** 1–17 (2011).

47. Arkin, I. T. & Brunger, A. T. Statistical analysis of predicted transmembrane alpha-helices. *Biochim. Biophys. Acta—Protein Struct. Mol. Enzymol.* **1429,** 113–128 (1998).

48. Momany, C. *et al.* Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature* **3,** 842–848 (1996).

49. de Planque, M. R. R. & Killian, J. A. Protein–lipid interactions studied with designed transmembrane peptides: role of hydrophobic matching and interfacial anchoring. *Mol. Membr. Biol.* **20,** 271–284 (2003).

50. Kim, S. *et al.* Transmembrane glycine zippers: physiological and pathological roles in membrane proteins. *Proc. Natl Acad. Sci. U.S.A.* **102,** 14278–14283 (2005).

51. Li, Q., Wong, Y. L., Huang, Q. & Kang, C. Structural insight into the transmembrane domain and the juxtamembrane region of the erythropoietin receptor in micelles. *Biophys. J.* **107,** 2325–2336 (2014).

52. Li, Q., Lei Wong, Y., Yueqi Lee, M., Li, Y. & Kang, C. Solution structure of the transmembrane domain of the mouse erythropoietin receptor in detergent micelles. *Sci. Rep.* **5,** 1–10 (2015).

53. Adamian, L. & Liang, J. Interhelical hydrogen bonds and spatial motifs in membrane proteins: polar clamps and serine zippers. *Proteins Struct. Funct. Genet.* **47,** 209–218 (2002).

54. Ulmschneider, M. B. & Sansom, M. S. Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta* **1512,** 1–14 (2001).

55. He, K. *et al.* Janus kinase 2 determinants for growth hormone receptor association, surface assembly, and signaling. *Mol. Endocrinol.* **17,** 2211–2227 (2003).

56. Wallweber, H. J. A., Tam, C., Franke, Y., Starovasnik, M. A. & Lupardus, P. J. Structural basis of recognition of interferon-α receptor by tyrosine kinase 2. *Nat. Struct. Mol. Biol.* **21,** 443–448 (2014).

57. Bah, A. *et al.* Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519,** 106–109 (2015).

58. Hernández, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2,** 715–726 (2007).

59. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293 (1995).

60. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins Struct. Funct. Bioinform.* **59,** 687–696 (2005).

61. Orekhov, V. Y. & Jaravine, V. A. Analysis of non-uniformly sampled spectra with multi-dimensional decomposition. *Prog. Nucl. Magn. Reson. Spectrosc.* **59,** 271–292 (2011).

62. Veglia, G., Zeri, A. C., Ma, C. & Opella, S. J. Deuterium/hydrogen exchange factors measured by solution nuclear magnetic resonance spectroscopy as indicators of the structure and topology of membrane proteins. *Biophys. J.* **82,** 2176–2183 (2002).

63. Doreleijers, J. F. *et al.* CING: an integrated residue-based structure validation program suite. *J. Biomol. NMR* **54,** 267–283 (2012).

64. Oxenoid, K. & Chou, J. J. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc. Natl Acad. Sci. U.S.A.* **102,** 10870–10875 (2005).

65. Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **39,** 277–286 (2006).

66. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

67. Jones, J. A., Wilkins, D. K., Smith, L. J. & Dobson, C. M. Characterisation of protein unfolding by NMR diffusion measurements. *J. Biomol. NMR* **10,** 199–203 (1997).

68. Danielsson, J., Jarvet, J., Damberg, P. & Gräslund, A. Two-site binding of β-cyclodextrin to the Alzheimer Aβ(1-40) peptide measured with combined PFG-NMR diffusion and induced chemical shifts. *Biochemistry* **43,** 6261–6269 (2004).

69. Ortega, A., Amorós, D. & García De La Torre, J. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101,** 892–898 (2011).

70. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815 (1993).

Acknowledgements

Author contributions

The NMR analyses was designed by K.B., G.W.H and B.B.K. K.B. recorded and analyzed all NMR data on the ECD and TMD, G.W.H carried out NMR analyses on the ICD. E.P. and K.L-L. designed and carried out the modelling, J.T.S.H. and C.V.R. designed and carried out the native MS, J.G.O. analyzed the SAXS data. K.B., K.L.-L. and B.B.K. designed the outline of the work and wrote the manuscript with input from all authors.

Additional information