20 YEARS OF LEPTIN

Insights into signaling assemblies of the leptin receptor

Frank Peelman*, Lennart Zabeau*, Kedar Moharana1, Savvas N Savvides1 and Jan Tavernier

Department of Medical Protein Research, Faculty of Medicine and Health Sciences, Flanders Institute for Biotechnology (VIB), Ghent University, A. Baertsoenkaai 3, 9000 Ghent, Belgium

1Unit for Structural Biology, Laboratory for Protein Biochemistry and Biomolecular Engineering (L-ProBE), Ghent University, K.L. Ledeganckstraat 35, 9000 Ghent, Belgium

*(F Peelman and L Zabeau contributed equally to this work)

Abstract

Leptin plays a central role in the control of body weight and energy homeostasis, but is a pleiotropic cytokine with activities on many peripheral cell types. In this review, we discuss the interaction of leptin with its receptor, and focus on the structural and mechanistic aspects of the extracellular aspects of leptin receptor (LR) activation. We provide an extensive overview of all structural information that has been obtained for leptin and its receptor via X-ray crystallography, electron microscopy, small-angle X-ray scattering, homology modeling, and mutagenesis studies. The available knowledge is integrated into putative models toward a recapitulation of the LR activation mechanism.

Key Words

- leptin
- leptin receptor
- structure
- modeling

Introduction

Leptin and the leptin receptor (LR) are essential components in the complex genetic wiring diagram underlying energy homeostasis and body weight. The hormone is now known to participate in a wide range of biological functions including innate and adaptive immunity (La Cava & Matarese 2004), reproduction (Chehab et al. 1996), and bone formation (Ducy et al. 2000).

Zhang et al. (1994) identified leptin as the product of the obese (ob) gene via positional cloning strategy. Study on its mechanism of action at the molecular level was initially hampered by the lack of any obvious sequence similarity with other human proteins. Madej et al. (1995) therefore used a threading approach to predict a four-α helix bundle cytokine structure for leptin, and suggested a JAK/STAT signaling pathway for leptin’s action. This prediction soon turned out to be correct when Tartaglia et al. (1995) revealed the identity of the LR via expression cloning. The LR showed the highest similarity to the interleukin 6 (IL6) signaling receptor chain glycoprotein 130 (gp130), the granulocyte colony-stimulating factor (G-CSF) receptor, and the leukemia inhibitory factor (LIF) receptor, and indeed uses JAK2 and STAT3 for its principle signaling pathway. The crystal structure of a W100E mutant of human leptin confirmed the four-helix bundle structure and the structural similarity with and IL6 (Zhang et al. 1997).

The crystal structure of leptin W100E: similarities with IL6 and G-CSF

Mature leptin is a non-glycosylated 16 kDa protein of 146 amino acids (Fig. 1A). Human leptin has two exposed tryptophan residues and tends to aggregate. Zhang et al. (1997) therefore created a leptin W100E mutant with full biological activity, but increased solubility and solved its...
Figure 1
Leptin sequence, structure, and position of binding-sites. (A) Pairwise sequence alignment of human (purple) and murine (green) leptin shows high sequence conservation (84.9% identity). Differences are shown in red. Secondary structures (α-helices and loops), as well as the intramolecular disulfide bridge, are shown based on the crystal structure of human leptin. (B) X-ray crystallographic structure of human leptin W100E (PDB ID:1AX8), shown as cartoon with rainbow color scheme. Residues of binding sites I (C), II (D), and III (E) are shown as spheres.
crystal structure at 2.4 Å resolution. Leptin adopts a typical four-helical bundle cytokine structure with four anti-parallel α helices (A, B, C, and D) in an up-up-down-down arrangement (Fig. 1B). The C-terminal residues of helix D adopt a 3–10 helix structure. Two long crossover loops AB and CD connect the parallel helices, while the antiparallel helices B and C are connected by a short BC loop (Fig. 1). The crystal structure lacks residues T27 to G38 in the AB loop, which were not visible in the density map. The long CD loop contains an extra helix E with both 3–10 helix structure and α-helix structure. Leptin has two conserved cysteine residues (C96 in the CD loop and C146 as the C-terminal residue), which form a solvent-exposed disulfide bridge that tethers the C-terminal part of helix D to the CD loop. This disulfide bridge is essential for structural stability and thus biological activity (Rock et al. 1996a, Haglund et al. 2012).

The structure of leptin is most similar to the structures of G-CSF and to the IL6 family of cytokines, including IL6, LIF, and ciliary neurotrophic factor (CNTF). Despite limited sequence similarity, the inter-helical angles and most characteristics of the crossover loops of leptin W100E are similar in these related structures. Together with these structural relatives, leptin is classified as a member of the long-chain cytokine family (http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.b.dj.b.b.html). However, the helices of leptin are on average one or two turns shorter. As a consequence, leptin has the shortest helical bundle structure in this family.

The leptin receptor

The LR has multiple isoforms

Structurally, the LR can be classified as a class I cytokine receptor (Fig. 2A). This family consists of single-membrane spanning receptors hallmarked by the presence of one or more cytokine receptor homology (CRH) domains with WSXWS motif. Additional extracellular features can include immunoglobulin-like (IGD) and fibronectin type III (FN III) domains. Some class I cytokine receptors are activated upon clustering of a single receptor, while others use a second (shared) chain. All receptors are devoid of intrinsic kinase activity and use JAK kinases for intracellular signaling.

Up to now, six LR isoforms produced by alternative splicing or ectodomain shedding have been identified: LRα–LRf. All these isoforms, except LRε, have an identical extracellular and transmembrane domain, but differ in the length of their intracellular tail (Fig. 2). LRb, or also referred to as LR long form (LRlo), is 1162 residues long with an intracellular domain of 302 Aa and is believed to be the only isoform capable of JAK/STAT signaling. This isoform is highly expressed in the specific nuclei of the hypothalamus, where it is involved in the regulation of body weight (Mercer et al. 1996, Schwartz et al. 1996, Fei et al. 1997). However, LRb expression can also be detected in a broad range of other cell types, in line with the pleiotropic effects of leptin. LRα (or LR short form, LRsh), LRε, LRd, and LRf have only short (30–40 residues) intracellular tails and unique C-termini. LRα has the broadest and the most pronounced expression pattern (Fei et al. 1997). The precise physiological role of these isoforms remains elusive, but high levels of expression of LRα and LRε in the choroid plexus and brain microvessels might suggest that they play a role in transport of leptin across the blood–brain barrier (Hileman et al. 2000). These isoforms may also play a role in renal clearance of leptin (Tartaglia et al. 1995). Initially discovered as a defective mutant (Lee et al. 1996), LRε is a soluble variant in mice, which is directly secreted into the bloodstream. In humans, a similar soluble LR fragment is generated by proteolytic (by metalloprotease 10 and 17) ectodomain shedding (Wauman et al. 2011) and modulates the bioavailability of leptin (Ge et al. 2002).

The extracellular segment of the LR is organized into six domains

The extracellular part of the LR has a complex domain topology with six separate domains: an N-terminal domain (NTD) of undefined structure and function, a first CRH domain (CRH1), an IGD, a second CRH domain (CRH2), and two membrane-proximal FN III domains (Fig. 2A).

The NTD of unknown function and the CRH1 domain are not required for LR signaling (Fong et al. 1998, Zabeau et al. 2004). In contrast, signaling is completely abolished by deletion of any of the four other domains. The CRH2 domain is the only high-affinity binding domain for leptin, as none of the other domains shows any detectable leptin binding (Fong et al. 1998, Zabeau et al. 2004).

By analogy to the leptin structural similarity with G-CSF and IL6 family cytokines, the LR shares its highest sequence similarity with the G-CSF receptor and the gp130 family receptors, including gp130, the LIF, and oncostatin M (OSM) receptors (Zabeau et al. 2003). This relationship is also reflected in the overall architecture of the extracellular domain (ECD) of these receptors (Fig. 2B): the receptors for leptin, G-CSF, LIF, OSM, and gp130 all contain an IGD followed by a CRH domain and followed by FN III domains. Like the LR, the LIF and OSM receptors have
an additional N-terminal CRH module, which is truncated in the OSM receptor. While gp130 and the G-CSF, LIF, and OSM receptors feature three membrane-proximal FN III domains, the LR only has two such membrane-proximal domains, and stands out as a structural outlier.

**Disulfide network and glycosylation of the ECD**

Haniu et al. determined the disulfide structure and glycosylation status of the human LR ECD expressed in Chinese hamster ovary (CHO) cells via mass spectrometry. For the 28 cysteines in the human LR ECD, nine disulfide bridges could be identified (Haniu et al. 1998; Fig. 2A). The CRH2 crystal structure suggests two additional possible disulfide bridges (Carpenter et al. 2012). The LR is heavily glycosylated, causing a 30–70 kDa increase in the apparent molecular weight on SDS–PAGE (Haniu et al. 1998, Kamikubo et al. 2008). Glycosylation accounts for 36% of the total LR molecular mass when expressed in CHO cells (Haniu et al. 1998). 18 of the 20 NXS glycosylation sites are occupied.

**Figure 2**

Schematic structure of leptin receptor (LR). (A) LR can be divided into three parts: extracellular (LR_{ecto}), transmembrane (LR_{tm}), and cytoplasmic tail (LR_{cyto}). The extracellular part LR_{ecto} is composed of a N-terminal domain (NTD), two cytokine receptor homology (CRH) domains (CRH1 and CRH2), an immunoglobulin-like domain (IGD), and two membrane-proximal fibronectin type III (FN III) domains. Numbers indicate the predicted domain borders. N-linked glycosylation sites (red), position of free cysteines (purple), and disulfide clusters (green) are shown on the right. (B) Modular structure of gp130 family of cytokine receptors. (C) LR isoforms: at least six different LR isoforms are found in human, namely LRα to LRf. LRα is the longest one with fully functional intracellular parts and thus is the only functional isotype. Short forms (LRα, LRc, LRd, and LRf) have unique C-terminal stretch of amino acids. LRε is a soluble variant.
motifs are glycosylated and N-glycosylation is predominant, although some O-glycosylation is also present (Haniu et al. 1998). Distribution of these experimentally determined N-glycosylation sites across different domains are as follows: six sites in the NTD, three in CRH1, two in IGD, two in CRH2, and five in the membrane-proximal FN III domains (Fig. 2A). N-glycosylation of the LR affects its interaction with leptin, as deglycosylation of recombinant ECD of the LR with N-glycanase F reduces leptin binding by 80% (Kamikubo et al. 2008). Interestingly, this effect is not mediated by the glycosylation of the CRH2 domain, but probably by glycosylation of the IGD or membrane-proximal FN III domains. Removal of the terminal sialic acids by treatment with sialidase A does not affect leptin binding (Kamikubo et al. 2008).

### Binding and assembly characteristics of the LR

#### Affinity and stoichiometry of the leptin–LR interaction

The binding of leptin to the soluble recombinant ECD of the LR has been characterized with different techniques, including isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), and competitive displacement of labeled leptin. The results of several binding studies are summarized in Table 1. The $K_D$ value for binding to the soluble ECD ranges from 0.2 to 15 nM.

While this broad range can partially be explained by different experimental methods, the differences in the ECD preparations used in the experiments may also affect the affinity. A fusion of the ECD to an immunoglobulin Fc domain that forms disulfide-linked dimers may have higher affinity than the monomeric ECDs. Possibly, forcing dimerization of the ECD by coupling it to an Fc domain increases the affinity for leptin by facilitating the formation of a dimeric or oligomeric receptor complex. The different sources of the recombinant material can also contribute, as for example glycosylation affects leptin binding (Kamikubo et al. 2008).

The average $K_D$ values (0.2–1.2 nM) for binding of leptin to the LR on the cell surface seem to be somewhat lower than the averages found for binding to the soluble ECD. It is possible that the affinity of leptin binding to its receptor on the membrane is increased, for example by formation of oligomeric signaling complexes on the membrane, but this hypothesis certainly requires further experimental validation (Sandowski et al. 2002). The EC50 value of human LR activation is quite similar and ranges

### Table 1 Characteristics of binding of leptin to different LR fragments

<table>
<thead>
<tr>
<th>LR fragment, source/cell type</th>
<th>$K_D$ or IC50 (nM)</th>
<th>Method</th>
<th>Leptin/LR species</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracellular domain (ECD), soluble</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF9 insect cells</td>
<td>10.5</td>
<td>SPR</td>
<td>Human</td>
<td>Rock et al. (1996b)</td>
</tr>
<tr>
<td>SF9 insect cells</td>
<td>15.4</td>
<td>SPR</td>
<td>Human</td>
<td>Mancour et al. (2012)</td>
</tr>
<tr>
<td>HEK293T (+ kifunensine)</td>
<td>13</td>
<td>ITC</td>
<td>Mouse</td>
<td>Moharana et al. (2014)</td>
</tr>
<tr>
<td>NS0 mouse myeloma cells</td>
<td>0.2–0.5</td>
<td>SPR ECD-Fc-fusion</td>
<td>Human</td>
<td>Mistrik et al. (2004)</td>
</tr>
<tr>
<td>NS0 mouse myeloma cells</td>
<td>0.1–0.5</td>
<td>SPR ECD-Fc-fusion</td>
<td>Mouse</td>
<td>Mistrik et al. (2004)</td>
</tr>
<tr>
<td>NS0 mouse myeloma cells</td>
<td>1.2</td>
<td>Leptin-biotin binding ECD-Fc-fusion</td>
<td>Human</td>
<td>Kamikubo et al. (2008)</td>
</tr>
<tr>
<td>COS-7</td>
<td>0.2</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Mouse</td>
<td>Liu et al. (1997)</td>
</tr>
<tr>
<td>Residues 330–839, soluble S2 insect cells</td>
<td>3.5</td>
<td>Leptin-biotin binding</td>
<td>Human</td>
<td>Kamikubo et al. (2008)</td>
</tr>
<tr>
<td>CRH2, soluble</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli S2 insect cells</td>
<td>15.4</td>
<td>SPR</td>
<td>Human</td>
<td>Sandowski et al. (2002)</td>
</tr>
<tr>
<td>S2 insect cells</td>
<td>2.5</td>
<td>Leptin-biotin binding</td>
<td>Human</td>
<td>Kamikubo et al. (2008)</td>
</tr>
<tr>
<td>CRH2, cell surface anchored</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COS-1</td>
<td>1.3</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Human</td>
<td>Fong et al. (1998)</td>
</tr>
<tr>
<td>Full-length LR on the cell surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAF3 cell homogenate</td>
<td>0.83</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Human</td>
<td>Sandowski et al. (2002)</td>
</tr>
<tr>
<td>COS-1</td>
<td>0.6</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Human</td>
<td>Fong et al. (1998)</td>
</tr>
<tr>
<td>COS-7</td>
<td>0.2</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Human</td>
<td>Liu et al. (1997)</td>
</tr>
<tr>
<td>COS-7</td>
<td>0.2</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Mouse</td>
<td>Liu et al. (1997)</td>
</tr>
<tr>
<td>COS-7</td>
<td>0.7</td>
<td>Competition with leptin-AP</td>
<td>Mouse</td>
<td>Tartaglia et al. (1995)</td>
</tr>
<tr>
<td>COS-7</td>
<td>0.7</td>
<td>Competition with leptin-AP</td>
<td>Human</td>
<td>Tartaglia et al. (1995)</td>
</tr>
<tr>
<td>HEK293</td>
<td>0.52</td>
<td>Time-resolved fluorescence</td>
<td>Human</td>
<td>Vauthier et al. (2013)</td>
</tr>
<tr>
<td>HEK293</td>
<td>1.4</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Human</td>
<td>Vauthier et al. (2013)</td>
</tr>
<tr>
<td>HEK293</td>
<td>2.6</td>
<td>Competition with leptin-AP</td>
<td>Mouse</td>
<td>Vauthier et al. (2013)</td>
</tr>
<tr>
<td>CHO</td>
<td>0.2–0.3</td>
<td>Competition with $^{125}$I-leptin</td>
<td>Mouse</td>
<td>Uotani et al. (1999)</td>
</tr>
</tbody>
</table>
from 0.2 to 1.5 nM in different studies, which is comparable with circulating concentrations of leptin.

LR binding only strictly requires the CRH2 domain, and binding of leptin to this domain shows a comparable $K_D$ as binding to the complete ECD, showing that CRH2 is the major affinity determinant (Fong et al. 1998, Sandowski et al. 2002). When expressed in Escherichia coli and refolded, the human LR CRH2 domain binds leptin with a $K_D$ of 15 nM and in 1:1 stoichiometry (Sandowski et al. 2002). Similar results were obtained with purified proteins from other species (reviewed in Gertler (2006)).

All studies agree that leptin binds to the LR ECD in a 1:1 stoichiometry (Devos et al. 1997, Mistrik et al. 2004, Mancour et al. 2012, Moharana et al. 2014). This 1:1 complex leads to further formation of 2:2 complexes at higher concentrations of ligand and receptor (Mancour et al. 2012, Moharana et al. 2014).

**The LR can assemble as a pre-formed cluster at the cell surface independent of leptin binding**

Numerous cytokine receptors can form inactive, pre-formed receptor complexes on the cellular surface in the absence of their ligands. Examples include the erythropoietin (Epo) receptor (Livnah et al. 1999, Remy et al. 1999, Constantinescu et al. 2001), growth hormone (GH) receptor (Gent et al. 2002, Brown et al. 2005), and the IL6 receptor (Schuster et al. 2003). Earlier studies suggested ligand-independent dimerization or oligomerization of the LR ECD in solution (Devos et al. 1997, Zabeau et al. 2005), but recent evidence obtained by multi-angle laser light scattering (MALLS) and small-angle X-ray scattering (SAXS) experiments has now indicated that it is monomeric in the absence of leptin (Mancour et al. 2012, Moharana et al. 2014). On the membrane, however, the LR can assemble as preformed dimers or oligomers, as evidenced by a high basal signal in the absence of leptin in both fluorescence resonance energy transfer (FRET) and bioluminescence resonance energy transfer (BRET) between differently tagged LRs and by co-immunoprecipitation of tagged LRs (Nakashima et al. 1997, White & Tartaglia 1999, Couturier & Jockers 2003, Biener et al. 2005). Different isoforms of the LR can heterodimerize with each other both in the presence and absence of leptin (White & Tartaglia 1999, Bacart et al. 2010). Leptin treatment induces an increase in BRET and FRET signals, which are probably a consequence of reorganization within the pre-formed complexes and/or de novo oligomerization (Couturier & Jockers 2003, Biener et al. 2005).

**A JAK/STAT complementation assay suggests a higher order leptin/LR complex**

The LR uses JAK2/STAT3 activation as its main signaling pathway. The JAK2 kinases associate with a well-conserved membrane-proximal proline-rich box1 and a less well-defined box2 motif (Bahrenberg et al. 2002, Kloek et al. 2002). Leptin stimulation brings two LRs together in an orientation that allows cross-phosphorylation of activation loops of their associated JAK2. Activated JAK2 phosphorylates all three cytoplasmic receptor tyrosine residues of the receptor and these phosphotyrosines recruit STAT3 (Vaisse et al. 1996). We used a JAK/STAT complementation strategy to show that the LR can form clusters with more than two receptors per activated complex (Zabeau et al. 2004). In this study, we used two signaling-deficient LR variants: a first receptor is deficient in STAT3 binding as its three cytoplasmic tyrosines are mutated, the second receptor lacks box1 and is deficient in JAK2 activation. As expected, these receptors do not show leptin-induced STAT3 activation. However, the leptin-induced STAT3 activation is restored when they are co-expressed. As JAK2/STAT3 signaling requires two JAK2 kinases and the presence of tyrosines in the cytoplasmic tail, this complementation strongly suggests that the LR is able to form complexes with at least three receptor chains.

**Leptin uses two or three binding sites for binding and signaling via the LR**

The possible structural and evolutionary relationship of leptin and the LR with the G-CSF and IL6 cytokines and corresponding receptors allowed us to use these receptor systems as a guide to study the leptin-LR system (Fig. 3).

The IL6 forms a hexameric 2:2:2 complex with two IL6 molecules, two gp130 receptor chains and two IL6 receptor α chains (Boulanger et al. 2003; Fig. 3A). IL6 uses three binding sites to achieve this hexameric complex: a first binding epitope, termed as site I, is formed by residues in helix D. Site II is formed by residues of the antiparallel helices A and C and binding site III is found around the N-terminal of helix D. For assembly of the hexameric complex, two IL6 molecules first use binding site I to interact with high-affinity with the CRH domain of two IL6 receptor α chains. The IL6 molecules subsequently interact with the CRH domain of gp130 via their binding site II. Two such trimeric IL6-IL6 receptor α-gp130 complexes form a hexamer by two symmetric interactions of IL6-binding site III in one trimer with the IGD of gp130 in the other trimer. A herpes virus IL6 analog (vIL-6) does
not require the IL6 receptor α chains to form a 2:2 vIL-6–gp130 complex retaining only the site II and III interactions (Chow et al. 2001).

G-CSF employs similar principles to evoke a tetrameric 2:2 complex with the G-CSF receptor via two binding epitopes, sites II and III (Tamada et al. 2006; Fig. 3B). A first G-CSF molecule binds to the CRH domain of a first G-SCF receptor via binding site II. Two such dimers form a symmetrical 2:2 tetramer by interactions of binding site III in one dimer with the IGD of the G-CSF receptor in the other dimer.

Based on structural superpositions, we identified and mutated residues in leptin that correspond to binding sites I, II, and III, and tested their effect on binding to CRH2 and on activation using a STAT3-based luciferase reporter readout (Peelman et al. 2004). Mutations in the predicted binding site II in helices A and C strongly affected binding to the CRH2 domain. This is completely in line with the role of binding site II in high-affinity binding to the CRH domain in all related cytokines. The residues defining binding site II are indicated in Fig. 1D. At the edge of binding site II, mutation of Q75 and R20 drastically affects leptin binding and signaling. However, as these residues interact with each other, we could not exclude a possible impact on the structure and/or stability of the protein.

Shpilman et al. (2011) used a random mutagenesis approach to identify a D23L mutation at the C-terminus of helix A, with a more than 60-fold increased affinity of leptin for CRH2.

Leptin mutations that superposed onto binding site III in G-CSF and IL6 strongly affected the maximal STAT3 luciferase induction, without affecting the EC50 value or binding to the CRH2 domain (Fig. 1E). A combined S120A-T121A mutation completely inhibited LR activation. In analogy with the IL6 and G-CSF receptor systems, we proposed that binding site III interacts with the IGD of the LR (Peelman et al. 2004).

Most potential binding site I mutations in the D helix hardly affected LR signaling. Nevertheless, an F41S mutation in the leptin AB loop that superposes on the position of the bindings site I did have a limited effect on the maximal luciferase activation value. However, F41 was also identified as a potential site III residue, in a study that compared the leptin–LR system with the vIL-6–gp130 complex (Niv-Spector et al. 2005a). vIL-6 uses both the region around the N-terminus of helix D and a region in its A-B loop for receptor interaction. A hydrophobic strand in the A-B loop of vIL-6 interacts with a hydrophobic strand that precedes the IGD (Chow et al. 2001). Hydrophobic cluster analysis predicted a corresponding hydrophobic strand at position 36-VTGLDFI-42 in leptin and a putative strand 325-VFTT-328 at the N-terminus of the IGD. Mutations in the leptin 39-LDFI-42 sequence completely inhibited LR activation (Niv-Spector et al. 2005b; Fig. 1E).

At present, it is still unclear whether leptin uses the area around S120-T121, the 39-LFDI-42 sequence or both as a binding site III for its interaction with the IGD. Below, we outline how leptin-binding site II and both potential binding site III may interact with the LR.

**Structural insights into the LR and models for the LR ECDs and insights into their mechanism**

The CRH2 domain structure and its binding to leptin site II

The crystal structure of the LR CRH2 domain was determined at 1.95 Å resolution, in complex with a Fab fragment of a MAB, which competes for leptin binding to CRH2, and hence acts as a LR antagonist (Carpenter et al. 2012). This is the first and hitherto the only high-resolution structure for (a part of) the LR. The CRH2 domain has a typical CRH domain structure, consisting of two FN III-like subdomains, which each have an IGD beta-sandwich fold with seven strands in two layers.
The topology of the CRH2 domain and numbering of the β-strands are shown in Fig. 4. As in most CRH domains, the C-terminal domain contains a typical WSXWS motif, 622-WSNWS-626, in which the tryptophan residues W622 and W625 become sandwiched in a π-stack between K614, R612 and R573 (Fig. 4). In the leptin CRH2 domain, this π-stack is extended by W583. As an unusual feature in CRH2, N624 in the WSXWS motif is glycosylated (Haniu et al. 1998). The NTD of CRH2 is stabilized by three disulfide bridges, two of which were previously detected by mass spectrometry analysis (Haniu et al. 1998). C613 and C604 in the C-terminal domain seem to correspond to free cysteines, although C604 is bound to a free cysteine used in the refolding buffer (Carpenter et al. 2012). GRAMM docking of leptin to the CRH2 crystal structures provides a model that is similar to the G-CSF-G-CSF receptor CRH complex, and that is in line with previous models and mutagenesis data for the mouse, human, and chicken leptin–CRH2 complex (Iserentant et al. 2005, Niv-Spector et al. 2005a, Carpenter et al. 2012). The EF loop in the CRH2 domain is crucial for leptin binding and contains four consecutive hydrophobic residues: 503-IFLL-506. Mutation of these residues strongly affects the binding of leptin and increases the EC50 value for receptor activation (Iserentant et al. 2005, Niv-Spector et al. 2005a). In the leptin structure, binding site II contains a hydrophobic cleft between helices A and C, consisting of residues L13, L86, V89, and F92 (Iserentant et al. 2005). In the docking model, L505 and L506 fit into the hydrophobic cleft and interact with L13 and L86 (Iserentant et al. 2005, Carpenter et al. 2012). Furthermore, Y472 also contributes to interaction with L86, V89 and V6 in the hydrophobic cleft in leptin. The JK loop in the second domain of CRH2 is predicted to interact with leptin helix A. In this loop, mutations of E565 and N566 affect receptor activation, while the F563 mutation affects binding to leptin. In the crystal structure, the JK loop assumes two different conformations, possibly in line with conformational flexibility. Carpenter et al. (2012) therefore suggested that leptin binding to CRH2 may involve an induced fit mechanism.

The IGD and its interaction with leptin-binding site III

The IGD of the LR shows only 20% or less sequence identity with the corresponding gp130 or G-CSF receptor IGD domains. Nevertheless, these domains likely fulfill similar roles in all three receptor systems. The IGD has no detectable binding affinity for leptin, but is crucial for LR activation. Deletion of this domain in the membrane-anchored receptor completely abolishes signaling (Fong et al. 1998, Zabeau et al. 2004). This is reminiscent of the role for this domain in the gp130 and G-CSF receptor systems, where it interacts with binding site III of the ligand already bound to another receptor chain. The presence of a binding site III analog in leptin suggests that the LR uses a similar site III–IGD interaction. We built homology models for the LR IGD and identified the residues that superpose on the binding site III binding regions in the gp130 and G-CSF receptors (Fig. 5A). Mutagenesis analysis identified the corresponding cluster of residues L370, A407, Y409, H417, and H418 in a conserved surface patch in the β-sheet formed by β-strands 3, 6, and 7 as the center of the leptin-binding site in this domain (Peelman et al. 2006a; Fig. 5A). This surface area on the IGD shows a positive electrostatic surface potential, which may interact with the surface with negative electrostatic surface potential around S120-T121 at the tip of helix D of leptin.

The CRH1 domain and the enigmatic NTD

The CRH1 domain and the preceding N-terminal sequence remain as orphan domains, as deletion of both only modestly affects receptor activation in transfection experiments (Fong et al. 1998, Zabeau et al. 2004). However seemingly in contrast, a Q269P mutation in CRH1 residue is the cause of the obese Fa/Fa phenotype of fatty Zucker rats (Takaya et al. 1996). A homology model for the CRH1 domain shows that Q269 is close to the WSXWS motif of the second FN III-like domain of CRH1, and the Q269P mutation may affect the structure of this domain (Fig. 5B).
The naturally occurring single-nucleotide polymorphism Q223R caused obesity in Brazilian multiethnic subjects (Duarte et al. 2006) and increased susceptibility toward protozoan infections in children (Duggal et al. 2011). A CRH1 homology model shows that Q223 is a solvent exposed in the N-terminal barrel of CRH1, but does not provide an immediate explanation for the effect of the mutation (Fig. 5B).

The LR sequence 325-VFTT-328 was predicted as an N-terminal strand of the IGD, possibly interacting with the 39-LDFI-42-binding site III. Mutations of 325-VFTT-328 into 325-AAA-328 strongly affected LR signaling (Niv-Spector et al. 2005a). However, modeling of the CRH1 domain points out that 325-VFTT-328 is in fact the last strand of CRH1, and probably crucial for its structure (Fig. 5B). Deletion of the entire CRH1 domain, including 325-VFTT-328, hardly affects LR signaling, ruling out a strict requirement for 325-VFTT-328 interaction with binding site III (Zabeau et al. 2004). Like the LR, the LIF receptor has an extra CRH1 module preceding its IGD. In the crystal structure of the LIF receptor, ECD in complex with LIF and in a model for the signaling complex, this CRH1 domain is not involved in interactions with the ligand or gp130 (Huyton et al. 2007). Interestingly, PSI-BLAST analysis of the N-terminal sequence of the LR aligns the uncharacterized N-terminal LR domain with the IGD of gp130. This suggests that CRH1 may be preceded by another IGD, which is also supported by the disulfide bridging pattern (Haniu et al. 1998). However, the function of both N-terminal LR domains remains unclear.

**The membrane-proximal FN III domains**

The membrane-proximal FN III domains also lack any binding affinity for leptin, but are indispensable for LR activation (Fong et al. 1998, Zabeau et al. 2005). A LR deletion variant with an extracellular part consisting of only these domains shows constitutive leptin-independent signaling. This illustrates that these domains can orientate the cytoplasmic tails in a manner that favors signaling (Zabeau et al. 2005).

The crystal structures of the full ECD and the FN III domains of gp130 (Xu et al. 2010) provide a template for modeling of the LR FN III domains. Gp130 has three membrane-proximal FN III domains, while the LR has two, and these LR FN III domains align best with the first two FN III domains of gp130. These two gp130 domains are bent toward each other with tight interdomain interactions. Modeling of the LR using these gp130 domains as a template indicates that the LR FN III domains may adopt a similar mutual orientation, as several interaction determinants seem to be conserved between gp130 and the LR (Fig. 6). Combined mutation of the only two free FN III cysteine
residues (C672 and C751) in the mouse LR completely abolished signaling (Zabeau et al. 2005). Superposition of the CRH2 crystal structure and FN III models on the gp130 structure suggests that C672 comes sufficiently close in C604 in the CRH2 domain to allow the formation of an extra intramolecular disulfide bridge (Carpenter et al. 2012; Fig. 6). A C604S LR mutant can only be activated at very high leptin doses, suggesting a role for C604 in the LR activation process (Moharana et al. 2014). However, it is unclear whether such a disulfide bridge occurs, as it was not detected by mass spectrometry, but it may for example be induced upon leptin binding. A role for free cysteines is supported by iodoacetamide treatment, which specifically blocks free cysteine residues. Treatment of the LR ECD with iodoacetamide does not affect the formation of 1:1 leptin–LR complexes, but blocks the formation of 2:2 complexes (Moharana et al. 2014). Possibly, formation of a C604–C672 disulfide bridge is required for formation of the 2:2 complexes, but this requires further experimental validation (Moharana et al. 2014).

**Homology models for the activated extracellular LR complex**

A wide array of mutagenesis and functional data have suggested that the interaction of the LR with leptin adopts the principles observed in the G-CSF and IL6 receptor families. Structures of the activated IL6 receptor complex, the G-CSF receptor complex, and of LIF bound to gp130 and to the IGD of the LIF receptor can therefore be used as a template for modeling an activated LR complex, where leptin interacts with CRH2 of a first-LR chain via site II and with the IGD of a second receptor via site III. Using the leptin structure and models for the IGD plus CRH2 domain, we tried to assemble a leptin–receptor 2:2 complex by superposition onto the different possible templates. These models reveal that the CRH2 domains may come closer to each other than found in the IL6 and G-CSF complexes, as a consequence of the shorter leptin helices (Fig. 7A). Two regions on leptin were independently proposed as part of binding site III: the area around the tip of the D helix, including residues S120 and T121 (Peelman et al. 2004), and the 39-LDFI-42 sequence in the AB loop (Niv-Spector et al. 2005b). Directing both areas simultaneously toward the same IGD model is hardly possible with an isolated leptin and IGD model, and becomes impossible in an oligomeric model based on the IL6, G-CSF, or LIF receptor complexes. Thus the question as to how both regions can contribute to binding site III remains open.

Figure 7
Models for leptin–LR interaction and topology of binding site III. (A) A model for a tetrameric 2:2 leptin–LR complex. Only the CRH2 and Ig-like domain of the LR are modeled. Residues S120 and T121 (green spheres) are close to the IGD, while residues 39-LDFI-42 (yellow spheres) do not contact the IGD. (B) A model for a hexameric 2:4 leptin–LR complex. S120 and T121 (green spheres) are close to the Ig-like domain, as in the tetrameric model of Fig. 7A. Residues 39-LDFI-42 (yellow spheres) interact with the CRH2 domain of an additional LR chain. (C) Binding site III residues mapped on the leptin structure. Residues corresponding to mutations that strongly affect LR activation are colored green, red, or yellow. Mutations with a moderate effect on LR activation are colored cyan blue. Residues S117 and 120-STE-122 seem to form a first-site IIIa, residues R128 and 39-LDFI-42 form a second-site IIlb.

Three alternative models can be proposed, which all deviate substantially from what is seen in the receptor complex structures for G-CSF, IL6, and LIF. A first model was suggested by Niv-Spector et al. (2005b) who suggested that the leptin 39-LDFI-42 motif interacts with a strand 325-VFTT-328 at the N-terminus of the IGD. However, modeling shows that this strand is part of the CRH1 structure. In this first alternative model, leptin uses a first
site III around S120-T121 to interact with the IGD, and 39-LDFI-42 to interact with the CRH1 domain. However, this seems to contradict the modest effect of CRH1 domain deletion on LR activation (Zabeau et al. 2004).

A second model is again based on an alternative role of the 39-LDFI-42 motif. We initially identified F41 as part of binding site I (Peelman et al. 2004), as superposition of leptin on IL6 brings the 39-LDFI-42 motif in the position of binding site I. Based on this idea, we proposed that this site may be used to attract another signaling chain. As no accessory receptor chain is known for LR signaling until now, we proposed a hexameric LR complex, where the 39-LDFI-42 motif interacts as a site I with the CRH2 of a third and fourth LR (Peelman et al. 2006a; Fig. 7B). This 2:4 hexameric leptin–LR model is however not supported by the 2:2 complexes found in solution (Mancour et al. 2012, Moharan et al. 2014).

A third alternative model therefore considers the interaction of the LR IGD with a larger binding site III, containing both the S120-T121 and the 39-LDFI-41 areas (Fig. 7C). Residue R128 may also be part of this site, as an R128Q mutation affects LR signaling, but not binding to CRH2, resembling the effects of the other binding site III mutations (Verploegen et al. 1997). It should be noted, however, that R128 is largely buried in the structure of leptin and that the mutation may affect the leptin structure. Modeling of the interaction of the larger binding site III requires angles between the LR CRH2 and IGD and/or between leptin and CRH2 that deviate substantially from those found in any of the related cytokines.

**Low-resolution structures of LR extracellular assemblies**

It is clear that models for the activated LR complex evoke a number of questions, and that a better understanding of the LR activation will require experimental structure determination. The extensive glycosylation of the LR and the possible flexibility of such a multidomain receptor have undoubtedly imposed technical challenges in obtaining structural insights into the extracellular assemblies of the LR by X-ray crystallography. Thus, recent efforts to obtain structural insights into extracellular assemblies of the LR have employed alternative approaches, albeit to low resolution, via EM and SAXS.

A first breakthrough in this direction was accomplished by Mancour et al. (2012), who employed negative-stain EM to generate 2D images followed by 3D reconstructions of a leptin–LR quaternary complex. In these models, leptin induces LR ECD dimerization thereby forming a quaternary 2:2 complex, which likely represents the core signaling configuration. The EM structure of the LR ECD alone shows significant flexibility in the hinge region between the N- and C-terminal CRH2 subdomains (Fig. 8A). Interaction with leptin rigidifies the membrane-proximal CRH2 subdomain thereby stabilizing the FN III domains into a ‘leg-like’ single conformation. Based on 3D reconstruction and modeling, site III of CRH2-bound leptin (via site II) interacts with the IGD of a second receptor in a cooperative manner, resembling the 2:2 tetrameric complexes of G-CSF receptor and our 2:2 model shown in figure (Fig. 7A). No function for the CRH1 (and NTD) could be identified in this study. More recent work by Moharan et al. (2014) has revealed an analogous 2:2 quaternary complex based on SAXS experiments, and additionally provided structural insights into the antagonism of signaling via the LR by two well-characterized leptin antagonist variants. The leptin
antagonists employed in this study are in fact mutant variants of leptin (S120A-T121A vs 39-AAAA-42) that retain their high-affinity binding to the LR. Importantly, complexes of the LR ECD with S120A-T121A or 39-AAAA-42 mutants stall into a 1:1 stoichiometry, confirming that both sets of sites are crucial for the assembly of a 2:2 complex from two 1:1 complexes (Fig. 8C). This argues against the role of the 39-LDFI-42 mutant as a binding site I, as the inhibitory effect of the mutation of this site already affects formation of a 2:2 complex. 4:2 hexameric complexes or dimers or oligomers of the LR ECD were never observed in solution, in contrast with the data that suggest higher order leptin–LR complexes and the occurrence of leptin-independent LR interactions on the cellular membrane (Zabeau et al. 2004). At this point, we cannot exclude the possibility that additional receptor–receptor interactions (e.g., between FN III domains) occur at the cell-surface 2:4 or 4:4 leptin–LR complexes.

Biomedical context of the mechanism of LR activation

The structures and models for LR domains were used to map mutations in the LR that are linked with severe obesity (Kimber et al. 2008). For several mutations, the models provide a rational explanation for the resulting obesity. Leptin plays a promoting role in the development of autoimmune and neoplastic disorders. Insights into the LR activation mechanism contributed to the development of leptin antagonists that find an application in the study of leptin’s role in these disorders. In the LR complex, the interactions between the LR CRH2 and leptin’s site II, IGD with site III or the homotypic FN III–FN III clustering have the potential to be targeted for therapeutic intervention. These interactions can be blocked by i) leptin antagonistic mutants; ii) leptin peptide antagonists; iii) leptin and LR-specific (monoclonal) antibodies or nanobodies; and iv) soluble LR variants (reviewed in Zabeau et al. 2014).

Potent leptin antagonists bind to the receptor comparable with WT leptin, but fail to induce downstream signaling. R128Q and S120A-T121A mutations and mutations of the 39-LDFI-42 motif all inhibit the LR activation without affecting binding to the LR (Verploegen et al. 1997, Niv-Spector et al. 2005b, Peelman et al. 2006b). These leptin variants block the receptor in a 1:1 complex and fail to form 2:2 complexes (Moharana et al. 2014). After prolonging the circulation half-life of the mutants, these leptin mutants become potent in vivo leptin antagonists. The potency of these antagonists can be increased by an additional D23L mutation in the antagonist leptin (Elínov et al. 2009).

Helices A and C form binding site II for high-affinity LR CRH2 binding. Peptides corresponding to Aa 3–34 (helix A) and 70–95 (helix B) of human leptin have antagonistic properties in vitro and in vivo (Gonzalez & Leavis 2003). Similarly, a modified peptide corresponding to the binding site III was shown to have antagonistic effects (Otovos et al. 2008).

Leptin and LR-specific antibodies or nanobodies (the variable domain of single-chain antibodies found in camelids) can prevent productive binding of leptin to its receptor (Fazel et al. 2006, Zabeau et al. 2012), or block FN III–FN III clustering necessary for receptor activation (Zabeau et al. 2012). Soluble LR variants trap free leptin in circulation and thereby compete with the membrane-bound receptor (De Rosa et al. 2006).

Leptin antagonists induce a pronounced weight gain, but have a positive effect in several disease models (Zabeau et al. 2014). Hitherto, the undesired effects of leptin antagonists on weight gain and metabolism preclude the development into candidate drugs for most applications. Efforts to design selective peripheral antagonists that do not affect leptin’s action at the CNS may help to develop leptin antagonism into a feasible approach for therapy (Beccari et al. 2013).

Conclusions

Although the crystal structure of the leptin–LR complex has not been determined, many valuable new insights into the mechanism of this receptor have been obtained by a combination of several approaches and methods including domain deletion studies, extensive site-directed mutagenesis studies coupled with functional interrogation of binding and signaling properties, crystal structure determination of the CRH2 domain, homology modeling using related IL6 and G-CSF receptor systems, and finally structural analysis by low-resolution methods such as EM and SAXS. These insights now provide a platform to explain how mutations in the LR can lead to obesity, and to spearhead the development of leptin and LR antagonists with potential applications in autoimmune diseases, cancer, elevated blood pressure, and certain cardiovascular diseases.

The low-resolution structures indicate that leptin can drive a 2:2 complex with its receptor, and we already mapped binding sites for complex formation. However, several important aspects remain unclear, including higher order LR complex formation and leptin-independent LR
interactions on the membrane, the interactions of binding site III on leptin and the role of the NTDs. More profound insight into these aspects will not only help us to understand how the canonical signaling of this receptor works in detail, but may also lead to new findings: as leptin seems to be a very pleiotropic signaling molecule, the potential existence of a binding site I in leptin, the oligomerization of the LR and the NTDs may all have unappreciated roles in leptin biology. The determination of higher resolution structures for other domains of the LR or leptin–LR complexes therefore remains a major challenge for the future.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this review.

Funding
This work was funded by IUAP (P6/36), Research Foundation-Flanders (FWO-V, Project G.0521.12N), and J T is recipient of an ERC Advanced Grant 340941.

Acknowledgements
We apologize to our colleagues that space limitations did not allow us to cite all possible relevant literature.

References


Chehab FF, Lim ME & Lu R 1996 Correction of the sterility defect in homozygous obese female mice by treatment with the human recombinant leptin. Nature Genetics 13 318–320. (doi:10.1038/ng0396-318)


Vaisse C, Halaas JL, Horvath CM, Darnell JE, Stoffel M & Friedman JM 1996 Leptin activation of Stat3 in the hypothalamus of wild-type and ob/ob mice but not db/db mice. Nature Genetics 14 95–97. (doi:10.1038/ng0996-95)


Received in final form 18 July 2014
Accepted 24 July 2014
Accepted Preprint published online 25 July 2014