Published in final edited form as:

Proteins. 2008 July ; 72(1): 393–401. doi:10.1002/prot.21937.

Comparative structural analysis of the binding domain of follicle stimulating hormone receptor

Qing R. Fan^{1,#} and Wayne A. Hendrickson^{1,2,*}

¹Department of Biochemistry and Molecular Biophysics Columbia University, New York, NY 10032

²Howard Hughes Medical Institute, Columbia University, New York, NY 10032

Abstract

Proteins with leucine-rich repeats (LRRs) specialize in mediating protein-protein interactions. The hormone binding portion of the receptor for follicle stimulating hormone (FSH) is an LRR protein by sequence, and the crystal structure of this domain from human FSH receptor in a complex with FSH shows that it does indeed have an LRR structure. It differs from other LRR domains, however, in being an all- β protein composed of highly irregular repeats and having only slight overall curvature. Despite these distinctions and a superficial resemblance to β -helical proteins, the binding domain of FSH receptor clearly is an LRR protein. The structure does consist of two parts with distinctively different curvatures. Comparison with the structures of other LRR-containing proteins shows a correlation between curvature and main-chain hydrogen bonding pattern of the parallel β -sheet. The hormone-binding site is located at the concave surface of the receptor structure, a feature common to proteins with LRR motifs. Analysis of the ligand-binding site of LRR-containing proteins reveals that they generally utilize extensive interface area and a large number of charged residues to facilitate high-affinity protein-protein interactions.

Keywords

glycoprotein hormone receptors; leucine-rich repeats; β -helical proteins; β -sheet curvature; protein-protein interactions; interface area and electrostatic attractions

Introduction

Follicle stimulating hormone receptor (FSHR) is a G-protein coupled receptor (GPCR) exclusively expressed in the gonads to regulate reproduction in mammals^{1,2}. FSHR transmits the signal generated by the follicle stimulating hormone (FSH) to promote the development of ovarian follicles in females and maturation of germ cells in males^{1,2}. FSH is secreted from the brain as a heterodimeric glycoprotein. It is a member of the glycoprotein hormone family, which also includes the luteinizing hormone (LH), chorionic gonadotropin (CG) and thyroid stimulating hormone (TSH)³. The glycoprotein hormones exert their actions by directly binding to their cognate receptors on the target cell surface.

The glycoprotein hormone receptors belong to a subfamily of GPCRs called the leucine-rich repeat-containing GPCRs (LGRs)⁴. These receptors are distinct by having a large

*Correspondence to: Wayne A. Hendrickson, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032. Phone: 212-305-3456. Fax: 212-305-7379. wayne@convex.hhmi.columbia.edu.

#Current address: Department of Pharmacology, Columbia University, New York, NY 10032.

extracellular domain with multiple leucine-rich repeats (LRR) in addition to the seven-helix transmembrane domain. The extracellular domains of the LGRs are responsible for binding protein ligands such as glycoprotein hormones and relaxin⁴.

The LRR motif contains the consensus sequence LxxLxLxxNxL, where x denotes any amino acids⁵⁻⁷. The conserved leucine can be replaced by other hydrophobic residues such as isoleucine, valine and phenylalanine; the consensus asparagine can be substituted with a cysteine or threonine. LRRs vary in length and are usually between 20 and 29 residues long⁵⁻⁶. The first crystal structure of a LRR protein was that of porcine ribonuclease inhibitor (RI)⁸. It has a horseshoe-like structure with repeating units of a short β -strand and a α -helix. The β -strands form a parallel β -sheet along the concave surface of the curved structure, which has been found in all known LRR structures. The α -helices in RI form the convex surface of the structure. In addition to helices, the inter-strand segments can also adopt extended conformations. A survey of the known structures of LRR proteins indicates that the proteins with more irregular repeats also have more irregular 3D structures⁶.

The complex structure of human FSH bound to the hormone-binding domain of FSHR (FSHR_{HB}) has been determined⁹. It reveals a highly irregular receptor structure with ten imperfect LRRs. In this study we present a detailed comparison of FSHR_{HB} with other known structures of the LRR family, and analyze the general properties of the ligand-binding site in various LRR structures.

Results and Discussion

Curvature and main-chain hydrogen bonding

The structure of FSHR_{HB} is a right-handed solenoid formed by ten highly irregular LRRs (Fig. 1A, B). It has an overall curved shape, with a parallel β -sheet along its concave surface (Fig. 1A, B). The repeating structural unit consists of the β -strand on the concave inner side and a return loop segment on the convex back side, which for five of the repeats also contains at least one short β -strand including two with two β -strands separated by a bulge (Fig. 2B). FSHR_{HB} differs from most known LRR structures in the absence of helices on the convex surface (Fig. 1A, B).

The FSHR_{HB} structure is unique in having two parts with distinctively different curvatures. The difference in curvature correlates with the main-chain hydrogen-bonding pattern in the parallel β -sheet. One part of FSHR_{HB} comprises the C-terminal repeats 7-10, which adopt the signature horseshoe-type curvature of LRR fold. Here, the fourth peptide of each concave-surface strand is rotated such that its carbonyl oxygen points inward toward the solenoid core, and thus a row of would-be hydrogen bonds is lost from the sheet (Fig. 1C). The β -strands in these repeats terminate at the residue with the inward pointing carbonyl, and most are only three residues long. The other part of FSHR_{HB} consists of the N-terminal repeats 1-7. The parallel β -sheet formed by the strands in these repeats is essentially flat, with all of the carbonyl oxygen atoms lying within the plane of the β -sheet and forming hydrogen bonds with the amides from neighboring strands (Fig. 1C, red dotted lines). As a result, the β -strands for the N-terminal repeats extend beyond those in the C-terminal region, and are at least five residues long.

The parallel β -strands in the C-terminal portion of the FSHR_{HB} structure have the canonical hydrogen-bonding pattern observed for most LRR proteins. As illustrated by two typical repeats from the structure of porcine RI, the β -strands of RI all have the inward pointing carbonyl at their end residues and are uniform in length (three residues long) (Fig. 2A). This contrasts with the longer, regular strands in the flat portion of FSHR_{HB} (Fig. 2B). The structure of RI exhibits a single curvature as a result of its highly regular repeating units.

Like RI, most known LRR structures have a single curvature with all the β -strands being three-residues long. The FSHR_{HB} structure presents a first example of LRR fold that exhibits correlated variation in the length of the parallel β -strands, their main-chain hydrogen bonding pattern and the curvature of the structure.

Irregular repeats of the FSHR_{HB} structure

FSHR_{HB} consists of ten highly imperfect repeats, and is one of the most irregular LRR structures known to date. A structural alignment of the individual repeats within FSHR_{HB} shows that they have the consensus sequence motif of $x_{1-2}LxxLxIxxNx_{1-2}LxxIx_{2-4}Fx_{1-6}$, where x indicates any amino acid. The first consensus L in the sequence of FSHR_{HB} is sometimes replaced by hydrophilic and even charged residues (Fig. 3A, 2nd position). The consensus L or I at other positions represents mostly leucine or isoleucine but can also be other hydrophobic residues such as valine, phenylalanine and alanine (Fig. 3A). Residues at the consensus F position form a phenylalanine ladder that is occasionally substituted with leucine and valine (Fig. 3A, 21st position). As in all LRR structures, the side chains of these conserved residues form the hydrophobic core of FSHR_{HB}. At any given consensus position, however, the amino acid shows more variation in the sequence of FSHR_{HB} than most other proteins with LRR motifs. Sequence diversity is also found at the consensus positions of LHR_{HB} and TSHR_{HB} (Fig. 3C, D).

Variation also occurs at the consensus N of the FSHR_{HB} sequence (Fig. 3A, 10th position). Most LRR proteins have a conserved asparagine or cysteine residue at this location^{5,6}. In contrast, an asparagine residue is present in only four of the FSHR_{HB} repeats; it is replaced by threonine, serine and even alanine in others. More divergence can be seen in the sequences of LHR_{HB} and TSHR_{HB}, which also include leucine, isoleucine and aspartate residues at this position (Fig. 3C, D). Residues at the consensus N position are located at the turns immediately following the parallel β -sheet. In most LRR structures, the asparagines or cysteine side chains from successive turns stack together in a ladder. They form hydrogen bonds with the main chain of the previous repeat as well as the inward pointing carbonyl upstream of the same repeat^{5,6}. A similar hydrogen bonding network is also observed for the asparagines, threonine and serine side chains in FSHR_{HB}. The alanine at this position in repeat 4 simply packs against other hydrophobic residues in the core. Similar packing interactions could be expected for the hydrophobic residues found at this position in the LHR_{HB} and TSHR_{HB} sequences.

The LRRs of FSHR_{HB} vary in their overall lengths (Fig. 3A). The first repeat is an integral part of the N-terminal flanking region. This repeat is the shortest, consisting of only 19 amino-acid residues. The last repeat is incomplete in the FSHR_{HB} structure because it is partially disordered. Length of the central eight repeats range from 22 to 26 residues. The repeats of FSHR_{HB} are shorter than those observed in the structure of RI, which are either 28 or 29 residues long. The shorter repeats of FSHR_{HB} are consistent with the extended conformation on the convex side of FSHR_{HB}, in contrast with the helical return in RI.

The repeats of FSHR_{HB} also have different conformations (Fig. 3B). The predominant conformation is observed in repeats R2, R3, R4 and R8, which are 24 to 25 residues long (Fig. 3B, gray). The average r.m.s.d. for the backbone atoms of these repeats is 1.04 ± 0.18 Å. The only significant difference among them is a kink around position 13, which contains either one or two residues. Repeats R6 and R7 form a separate set because of their distinct conformation at the second position, which corresponds to the first consensus L (Fig. 3B, green). In contrast to the hydrophobic residues found among the rest of the repeats, residues at this position in R6 and R7 include a lysine and serine. Their side chains face outwardly and are solvent-exposed, consistent with the hydrophilic nature of these amino acids. As a result of the unique main-chain conformation observed at the second position of R6 and R7,

their parallel β -strands are extended by one residue at the beginning. R1, R5 and R9 exhibit dramatically different conformations from all other repeats (Fig. 3B, red).

Comparison with other LRR proteins

Among the known LRR structures, FSHR_{HB} most closely resembles the extracellular domain of Nogo receptor (NgR)_{10·11} (r.m.s.d. of 1.66 Å for 174 C α atoms) and the platelet-receptor glycoprotein Ib α (gplb α)_{12·13} (r.m.s.d. of 1.62 Å for 130 C α atoms) (Fig. 4A). The common features of FSHR_{HB}, NgR and Gplb α include an extended shape with low curvature and the lack of regular helices on the convex surface of the structures. FSHR_{HB} is also similar to the NgR structure in the presence of a Phe ladder on the convex side (Fig. 4A). These Phe residues are conserved in the sequence of LHR_{HB} and TSHR_{HB} (Fig. 3C, D). Nevertheless, NgR and gplb α structures are much more regular than FSHR_{HB} both in the length and overall conformation of their repeats. For example, all the LRRs of NgR are 24 residues long except for one repeat which contains 25 amino acids 10·11. Unlike FSHR_{HB}, which shows different curvature for different parts of the structure, both NgR and gplb α have a single curvature. The structures of NgR and gplb α are also more similar to one another (r.m.s.d. of 1.39 Å for 192 C α atoms) than either is to FSHR_{HB}.

LRR domains often have flanking regions at the N- and C-termini to cap the hydrophobic interior of the solenoid structure⁶. A homologous N-terminal flanking module is found in the structures of FSHR_{HB}, NgR and Gplb α (Fig. 4A, B). All three structures contain a β -hairpin secured by a disulfide bridge. An additional disulfide bond is present in FSHR_{HB} and NgR, which bridges the N-terminus to the β -hairpin. The N-terminal β -hairpin motif is termed the ' β -finger' in the structure of Gplb α ; it protrudes to one side of the concave surface and is involved in binding the von Willebrand factor (VWF) ligand¹². NgR has a very similar ' β -finger' which is only one residue shorter than the one in Gplb α _{10·11} (Fig. 4B). The β -hairpin in FSHR_{HB} is four residues shorter than the ' β -finger' of Gplb α , and does not protrude into the concave side of the structure as in Gplb α (Fig. 4A). Furthermore, the N-terminal β -hairpin in FSHR_{HB} does not participate in ligand-binding.

The structures of FSHR_{HB}, NgR and Gplb α all have different C-terminal caps. The C-terminal flanking region in FSHR_{HB} adopts different conformations in the two molecules of the asymmetric unit due to different crystal packing environments. This region is also partially disordered, suggesting that its conformation is flexible in the isolated structure of FSHR_{HB}. Furthermore, this region is not involved in hormone binding by FSHR_{HB}. In contrast, the C-terminal flanking region in the structure of Gplb α contains a ' β -switch', which is the site of most extensive contacts with VWF¹² (Fig. 4A). NgR has a cysteine-rich module at the C-terminus formed by several helices, which has also been implicated in ligand-binding^{10·11} (Fig. 4A).

Comparison with β -helical proteins

Proteins with LRRs share certain structural features with β -helical proteins, which also include a right-handed solenoid structure, rows of hydrophobic residues constituting the hydrophobic core of the structure, and an asparagine ladder with a hydrogen-bonding network between the asparagine side chains and the main chain of neighboring repeats^{5·6·14·17}. The L-domains of insulin receptor and epidermal growth factor receptors contain both LRR and β -helical protein features¹⁸. Because FSHR_{HB} is unusual among LRRs in being an all- β protein, its analogy with β -helical proteins is especially evident.

Similarities between LRR and β -helical proteins are somewhat superficial, however, and there are distinct differences both in sequence and structural characteristics (Fig. 5). First of all, the hydrophobic core of LRR structures are mostly filled with the LRR namesake

leucines, while the majority of interior residues in β -helical proteins are valines and isoleucines. These valines and isoleucine residues feature in cupped stacks within the solenoid¹⁷, but a recurrent sequence motif is not evident in β -helical proteins¹⁹ whereas it is in LRRs. Second, a signature characteristic of LRR proteins is their horseshoe-type curvature^{5,6,14}. In contrast, the solenoid tube of β -helix proteins such as pectate lyase C (PelC) is straight^{15,16}. Third, LRR proteins all have in common one parallel β -sheet which spans the concave surface of the LRR solenoid, whereas there can be either α -helices or extended loops on the convex side^{5,6,14}. In a typical β -helix protein, each rung consists of three β -strands¹⁵⁻¹⁷ that take part in three parallel β -sheets, each of which spans the entire solenoid. Two of the β -sheets form an antiparallel β -sandwich, while the third is perpendicular to the β -sandwich^{15,16}. Fourth, the concave-surface β -sheet of a LRR structure typically features an interrupting inward-pointing carbonyl (Fig. 2A), which is associated with LRR curvature. This peculiar hydrogen-bonding pattern is not observed in any of the β -sheets in a β -helix, where all carbonyl groups lie flat within the β -sheet (Fig. 2C). As a result, the β -sheets of a β -helical protein are not curved and they vary in strand length. Finally, in most LRR structures, the parallel β -sheet on the concave surface is flat^{5,6,14}, whereas in a β -helical structure at least one of the β -sheets has a left-handed twist with respect to the cross section of the solenoid¹⁴⁻¹⁶. An overall twist of the entire solenoid is observed in the LRR structures of three internalins²⁰⁻²² and YopM²³, but unlike the twisted β -sheets in β -helical structures, these LRR sheets track with the cross section of the tube along the entire path of the solenoid.

Although the structure of FSHR_{HB} has only β -strands, its overall architecture resembles that of a LRR module more closely than a β -helix. The C-terminal four repeats of the FSHR_{HB} structure have the signature curvature, short β -strands and canonical hydrogen-bonding pattern of all LRR structures. The N-terminal six repeats of FSHR_{HB} contain certain features of a β -helical protein such as the lack of curvature and inward pointing carbonyl groups at the end of β -strands. Two of the repeats (R3 and R4) in the N-terminal part of FSHR_{HB} structure even contain three β -strands, but these follow one another in parallel on the concave face of FSHR_{HB}. In contrast, none of the three β -strands of any repeat in a β -helix are parallel and they are always in sheets running the entire length of the solenoid. The N-terminal part of the FSHR_{HB} structure also differs from β -helical proteins in being completely untwisted.

Common ligand-binding mode by LRR structures

Proteins containing LRRs participate in a wide range of cellular functions, including signal transduction, enzyme inhibition, cell adhesion and RNA processing. A common feature of the LRR-containing proteins is their role in mediating protein-protein interactions⁵. Indeed, the adaptive immune response in sea lampreys is based on diverse LRR receptors rather than on immunoglobulins as in vertebrates²⁴, and LRR-based libraries have been designed with diverse binding potentials²⁵. In addition to the FSH-FSHR_{HB} structure⁹, other known structures of LRR protein-ligand complexes include the porcine RI-RNase A complex²⁶ and the related human RI-angiogenin complex²⁷, human platelet-receptor GpIb α bound to the plasma protein VWF A1 domain (VWF A1)^{12,28}, human spliceosomal U2B''-U2A' protein complex²⁹, internalin A (InIA) of *Listeria monocytogenes* bound to its human receptor E-cadherin²² and internalin B (InIB) bound to the ectodomain of human receptor tyrosine kinase Met³⁰, the ubiquitin ligase subunit Skp2 complexed with accessory protein Cks¹³¹, human TSH receptor ectodomain bound to a thyroid-stimulating antibody³², a complex between Toll-like receptor TLR4 and its lipopolysaccharide (LPS)-binding cofactor MD-2³³, and a human Slit-Robo receptor complex³⁴. A survey of these complexes shows that the ligand is bound to the concave surface of the LRR protein in each case, and that the ligand-binding site involves the solvent-exposed parallel β -sheet and C-terminally adjacent

loops (Fig. 6). Other surfaces of the LRR architecture can also be utilized for protein interactions. An example is the complex formed between GpIb α and thrombin where the convex surface of GpIb α contains one of the thrombin recognition sites^{35,36}. The protein-protein interactions occurring at the concave surface of LRR structures are the focus of this study.

LRR proteins use large areas of solvent-exposed surface to facilitate high-affinity protein-protein interactions. For example, the dissociation constant (K_d) of the FSH-FSHR_{HB} complex is in the nanomolar range²; RI binds to RNase A with an inhibition constant (K_i) of 5.9×10^{-14} M²⁶; and GpIb α binds to wild type VWF A1 with a K_d of 30 nM, which is reduced 2.5 fold by a gain-of-function mutation in VWF A112. The tight association is consistent with the extensive binding surface of LRR protein-ligand complexes, which ranges from 1800 Å² for U2B''-U2A' to 2600 Å² for the FSH-FSHR_{HB}, RI-RNase A and InlA-E-cadherin interfaces. In all but one of the known cases of LRR protein-ligand complexes, the binding area is greater than the 'standard-size' interfaces of 1600 ± 400 Å² found in non-covalent protein-protein complexes³⁷.

The receptor-hormone interface of the FSH-FSHR_{HB} complex exhibits relatively lower shape complementarity than most protein-protein complexes. The median value of the shape correlation statistic (Sc) is used to describe the geometric match at protein-protein interfaces³⁸. The Sc value for the FSH-FSHR_{HB} interface is 0.59. We also find similar shape complementarity values at the RI-RNaseA, InlA-E-cadherin, spliceosomal U2B''-U2A', GpIb α -VWF A1, Skp2-Cks1, InlB-Met, TLR4-MD-2 and Slit-robo interfaces (Fig. 6), with an average Sc value of 0.61. In contrast, reported Sc statistics for protease-inhibitor and protein oligomer interfaces are in the range of 0.70-0.76, and the mean Sc value for antibody-antigen interfaces is 0.66³⁸. The relatively low Sc statistics observed for LRR protein-ligand complexes may be related to the characteristics of LRR structures. The global fold of a protein containing LRRs is very constrained; the only variation at the concave face of the protein lies in its curvature. The side chains protruding from this concave face are responsible for providing all the shape complementarity toward the wide range of ligands that are recognized by LRR-containing proteins. As a result, the shape complementarity at the ligand-binding site of LRR-containing proteins are not as optimized as the hand-glove fit observed at protease-substrate and protein-protein inhibitor interfaces.

LRR proteins bind to their ligands with high affinities despite the low shape complementarity observed at the interfaces. In order to understand what contribute to the tight binding, we have examined the nature of their extensive interactions. We find that the LRR protein-ligand interactions have an average buried charge density of 0.97 ± 0.16 charges/nm², higher than for most protein-protein complexes (Fan and Hendrickson, unpublished data). Electrostatic interactions appear to play a predominant role for protein-protein recognition when the Sc value at the interface is low; when the Sc value is high, the contribution of charges can be high or low depending on the individual complex. Charged interactions are compatible with low shape complementarity because they have a less stringent requirement for geometric match than hydrophobic interactions.

In conclusion, the solenoid architecture of LRR modules provides a framework for proteins to bind a variety of ligands. The large surface area of the solvent-exposed parallel β -sheet and electrostatic attractions may play a role in the versatility of LRR motifs in protein-ligand recognition.

Materials and methods

Structural analysis

The structure of human FSHR_{HB} complexed with FSH was determined at 2.9 Å resolution with two molecules in the asymmetric unit as described⁹. Secondary structure elements were assigned according to the algorithm implemented by DSSP³⁹. Main-chain hydrogen bonds include amide to carbonyl interactions with O-N distances less than 3.4 Å and carbonyl to amide angles greater than 120°.

The solvent accessible surface area buried at protein-protein and protein-DNA interfaces were calculated with SURFACE40 using a probe radius of 1.4 Å. The surface complementarity (*Sc*) at each of these interfaces was determined using the SC program³⁸.

Our procedures for determination of buried charge density will be described separately (Fan and Hendrickson, in preparation). Briefly, buried charge density at a protein-protein or protein-DNA interface is defined as the number of charged residues buried at the interface divided by the interface area in nm². The charged residues buried at interfaces include those which have at least one charged protein side-chain atom buried by the interactions.

Structural superpositions

The program LSQMAN41 was used to perform pairwise comparison between FSHR_{HB} and other known LRR structures. The superposition of individual repeats of FSHR_{HB} was carried out using repeat 3 as the reference. The optimal structural alignment between repeat 3 and every other repeat was determined by a brute-force algorithm in LSQMAN which searched for the best match by sequentially aligning eight-residue fragments. Each repeat was then superimposed onto repeat 3 in O 42 using the transformation matrix which related their best aligned segments.

Structure-based sequence alignment of leucine-rich repeats

Sequence alignment of different repeats within each glycoprotein hormone receptor was performed manually. For human FSHR_{HB}, the alignment for human FSHR_{HB} was based on the amino acid residues aligned in the three-dimensional FSHR_{HB} structure. The core region of β-strands was placed at the beginning of each repeat. For human LHR_{HB} and TSHR_{HB}, the alignment of repeats was based on the overall sequence alignment of the three human glycoprotein hormone receptors, and aimed at maximizing the number of identical residues at each position and placing gaps at consistent locations within the repeat.

Illustrations

Figures were generated using RIBBONS43, Adobe Illustrator and Adobe Photoshop.

Acknowledgments

We thank R. Kovall, L. Mosyak, E. Alexov and C. Lusty for helpful advice. This work was supported in part by NIH grant GM68671. Q. R. F. was an Agouron Institute fellow of the Jane Coffin Childs Memorial Foundation.

References

1. Dias JA, Cohen BD, Lindau-Shepard B, Nechamen CA, Peterson AJ, Schmidt A. Molecular, structural, and cellular biology of follitropin and follitropin receptor. *Vitam Horm.* 2002; 64:249–322. [PubMed: 11898394]
2. Simoni M, Gromoll J, Nieschlag E. The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology. *Endocr Rev.* 1997; 18:739–773. [PubMed: 9408742]

3. Pierce JG, Parsons TF. Glycoprotein hormones: structure and function. *Annu Rev Biochem.* 1981; 50:465–495. [PubMed: 6267989]
4. Hsu SYT. New Insights into the evolution of the relaxin-LGR signaling system. *Trends Endocrinol Metab.* 2003; 14:303–309. [PubMed: 12946872]
5. Kobe B, Deisenhofer J. Proteins with leucine-rich repeats. *Curr Opin Struct Biol.* 1995; 5:409–416. [PubMed: 7583641]
6. Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol.* 2001; 11:725–732. [PubMed: 11751054]
7. Enkhbayar P, Kamiya M, Osaki M, Matsumoto T, Matsushima N. Structural principles of leucine-rich repeat (LRR) proteins. *Proteins.* 2003; 54:394–403. [PubMed: 14747988]
8. Kobe B, Deisenhofer J. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature.* 1993; 366:751–756. [PubMed: 8264799]
9. Fan QR, Hendrickson WA. Structure of human follicle-stimulating hormone in complex with its receptor. *Nature.* 2005; 433:269–277. [PubMed: 15662415]
10. He XL, Bazan JF, McDermott G, Park JB, Wang K, Tessier-Lavigne M, He Z, Garcia KC. Structure of the Nogo receptor ectodomain: a recognition module implicated in myelin inhibition. *Neuron.* 2003; 38:177–185. [PubMed: 12718853]
11. Barton WA, Liu BP, Tzvetkova D, Jeffrey PD, Fournier AE, Sah D, Cate R, Strittmatter SM, Nikolov DB. Structure and axon outgrowth inhibitor binding of the Nogo-66 receptor and related proteins. *EMBO J.* 2003; 22:3291–3302. [PubMed: 12839991]
12. Huizinga EG, Tsuji S, Romijn RA, Schiphorst ME, de Groot PG, Sixma JJ, Gros P. Structures of glycoprotein Iba and its complex with von Willebrand factor A1 domain. *Science.* 2002; 297:1176–1179. [PubMed: 12183630]
13. Uff S, Clemetson JM, Harrison T, Clemetson KJ, Emsley J. Crystal structure of the platelet glycoprotein Iba N-terminal domain reveals an unmasking mechanism for receptor activation. *J Biol Chem.* 2002; 277:35657–35663. [PubMed: 12087105]
14. Kobe B, Kajava AV. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. *Trends Biochem Sci.* 2000; 25:509–515. [PubMed: 11050437]
15. Yoder MD, Keen NT, Jurnak F. New domain motif: the structure of pectate lyase C, a secreted plant virulence factor. *Science.* 1993; 260:1503–1507. [PubMed: 8502994]
16. Yoder MD, Lietzke SE, Jurnak F. Unusual structural features in the parallel β -helix in pectate lyases. *Structure.* 1993; 1:241–251. [PubMed: 8081738]
17. Jenkins J, Pickersgill R. The architecture of parallel β -helices and related folds. *Prog Biophys Mol Biol.* 2001; 77:111–175. [PubMed: 11747907]
18. Ward CW, Garrett TP. The relationship between the L1 and L2 domains of the insulin and epidermal growth factor receptors and leucine-rich repeat modules. *BMC Bioinformatics.* 2001; 2:4. [PubMed: 11504559]
19. Bradley P, Cowen L, Menke M, King J, Berger B. Successful prediction of parallel β -helices from primary sequence reveals an association with many microbial pathogens. *Proc Natl Acad Sci U S A.* 2001; 98:14819–14824. [PubMed: 11752429]
20. Marino M, Braun L, Cossart P, Ghosh P. Structure of the InlB leucine-rich repeats, a domain that triggers host cell invasion by the bacterial pathogen *L. monocytogenes*. *Mol Cell.* 1999; 4:1063–1072. [PubMed: 10635330]
21. Schubert WD, Gobel G, Diepholz M, Darji A, Kloer D, Hain T, Chakraborty T, Wehland J, Domann E, Heinz DW. Internalins from the human pathogen *Listeria monocytogenes* combine three distinct folds into a contiguous internalin domain. *J Mol Biol.* 2001; 312:783–794. [PubMed: 11575932]
22. Schubert WD, Urbanke C, Ziehm T, Beier V, Machner MP, Domann E, Wehland J, Chakraborty T, Heinz DW. Structure of Internalin, a major invasion protein of *Listeria monocytogenes*, in complex with its human receptor E-cadherin. *Cell.* 2002; 111:825–836. [PubMed: 12526809]
23. Evdokimov AG, Anderson DE, Routzahn KM, Waugh DS. Unusual molecular architecture of the *Yersinia pestis* cytotoxin YopM: a leucine-rich repeat protein with the shortest repeating unit. *J Mol Biol.* 2001; 312:807–821. [PubMed: 11575934]

24. Pancer Z, Amemiya CT, Ehrhardt GR, Ceitlin J, Gartland GL, Cooper MD. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature*. 2004; 430:174–180. [PubMed: 15241406]
25. Stumpp MT, Forrer P, Binz HK, Plückthun A. Designing repeat proteins: modular leucine-rich repeat protein libraries based on the mammalian ribonuclease inhibitor family. *J Mol Biol*. 2003; 332:471–487. [PubMed: 12948496]
26. Kobe B, Deisenhofer J. A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature*. 1995; 374:183–186. [PubMed: 7877692]
27. Papageorgiou AC, Shapiro R, Acharya KR. Molecular recognition of human angiogenin by placental ribonuclease inhibitor--an X-ray crystallographic study at 2.0 Å resolution. *EMBO J*. 1997; 16:5162–5177. [PubMed: 9311977]
28. Dumas JJ, Kumar R, McDonagh T, Sullivan F, Stahl ML, Somers WS, Mosyak L. Crystal structure of the wild-type von Willebrand factor A1-glycoprotein Iba complex reveals conformation differences with a complex bearing von Willebrand disease mutations. *J Biol Chem*. 2004; 279:23327–23334. [PubMed: 15039442]
29. Price SR, Evans PR, Nagai K. Crystal structure of the spliceosomal U2B''-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature*. 1998; 394:645–650. [PubMed: 9716128]
30. Niemann HH, Jäger V, Butler JG, van den Heuvel J, Schmidt S, Ferrais D, Gherardi E, Heinz DW. Structure of the human receptor tyrosine kinase Met in complex with the *Listeria* invasion protein InlB. *Cell*. 2007; 130:235–246. [PubMed: 17662939]
31. Hao B, Zheng N, Schulman BA, Wu G, Miller JJ, Pagano M, Pavletch NP. Structural basis of the Cks1-dependent recognition of p27^{Kip1} by the SCF^{Skp2} ubiquitin ligase. *Mol Cell*. 2005; 20:9–19. [PubMed: 16209941]
32. Sanders J, Chirgadze DY, Sanders P, Baker S, Sullivan A, Bhardwaja A, Bolton J, Reeve M, Nakatake N, Evans M, Richards T, Powell M, Miguel RN, Blundell TL, Furmaniak J, Smith BR. Crystal structure of the TSH receptor in complex with a thyroid-stimulating autoantibody. *Thyroid*. 2007; 17:395–410. [PubMed: 17542669]
33. Kim HM, Park BS, Kim JI, Kim SE, Lee J, Oh SC, Enkhbayar P, Matsushima N, Lee H, Yoo OJ, Lee JO. Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist eritoran. *Cell*. 2007; 130:906–917. [PubMed: 17803912]
34. Morlot C, Thielens NM, Ravelli RB, Hemrika W, Romijn RA, Gros P, Cusack S, McCarthy AA. Structural insights into the Slit-Robo complex. *Proceedings of National Academy of Science USA*. 2007; 104:14923–14928.
35. Dumas JJ, Kumar R, Seehra J, Somers WS, Mosyak L. Crystal structure of the GpIba-thrombin complex essential for platelet aggregation. *Science*. 2003; 301:222–226. [PubMed: 12855811]
36. Celikel R, McClintock RA, Roberts JR, Mendolicchio GL, Ware J, Varughese KI, Ruggeri ZM. Modulation of α -thrombin function by distinct interactions with platelet glycoprotein Iba. *Science*. 2003; 301:218–221. [PubMed: 12855810]
37. Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. *J Mol Biol*. 1999; 285:2177–2198. [PubMed: 9925793]
38. Lawrence MC, Colman PM. Shape complementarity at protein/protein interfaces. *J Mol Biol*. 1993; 234:946–950. [PubMed: 8263940]
39. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*. 1983; 22:2577–2637. [PubMed: 6667333]
40. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol*. 1971; 55:379–400. [PubMed: 5551392]
41. Kleywegt GJ. Use of non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr D*. 1996; 52:842–857. [PubMed: 15299650]
42. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in the models. *Acta Crystallogr A*. 1991; 47:110–119. [PubMed: 2025413]
43. Ribbons, Carson M. *Methods Enzymol*. 1997; 277:493–505. [PubMed: 18488321]

Figure 1.

Structure of FSHR_{HB} and main-chain hydrogen-bonds in the parallel β-sheet.

(A) Stereo plot of the Ca trace of FSHR_{HB} with every tenth residue marked.

(B) Ribbon diagram of FSHR_{HB}. β-strands located at the concave face are shown in red and those on the convex face are in pink. Numbering of the leucine-rich repeats is indicated below the diagram.

(C) Main-chain hydrogen-bonds of the β-sheet at the concave face of FSHR_{HB} structure. Hydrogen bonds that are unique to the N-terminal six repeats are in red. The carbonyl oxygen atoms at the corresponding positions in repeats 7 through 9 are marked with red asterisks.

Figure 2. Comparison of two representative repeats from the structures of (A) porcine RI, (B) human FSHR_{HB} and (C) *E. chrysanthemi* PelC. For each structure, main-chain hydrogen-bonds of two neighboring β -strands are shown for the two repeats highlighted by a circle in the ribbon diagram.

Figure 3.

Comparison of individual repeats within FSHR_{HB}. Conserved Leu/Ile/Phe residues at structurally equivalent positions are highlighted in red for residues located within the parallel β -sheet, and pink for those located on the interstrand segments. Residues at the conserved Asn ladder position are colored cyan.

(A) Alignment of individual leucine-rich repeats within the FSHR_{HB} sequence.

(B) Stereo diagram showing the superposition of individual FSHR_{HB} repeats. C α traces of the repeats are shown. Repeats 1, 5 and 9 are colored red, repeats 6 and 7 are green, and repeats 2, 3, 4 and 8 are black. C α positions of the conserved Leu/Ile/Phe are marked with spheres for repeat 3.

(C, D) Alignment of individual leucine-rich repeats within the LHR_{HB} (C) and TSHR_{HB} (D) sequences.

Figure 4.

(A) Ribbon diagrams of FSHR_{HB} (red), NgR (green) and GpIb α (blue), showing the Phe ladder on the convex side of FSHR_{HB} and NgR as well as the hydrophobic residues at the corresponding position in GpIb α . The “ β -finger” and “ β -switch” modules are labeled for the GpIb α structure. R.m.s.d. values are indicated for each pairwise comparison.

(B) Superposition of the N-terminal β -hairpin region of FSHR_{HB} (red), NgR (green) and GpIb α (blue) structures. The disulfide bridges found in the structure of FSHR_{HB} are in black.

Figure 5. Comparison of (A) human FSHR_{HB} and (B) *E. chrysanthemi* PelC structures in three orthogonal views. The β -sheet preceding the conserved Asn ladder is considered the front in each case.

Figure 6.

Comparison of the FSH-FSHR_{HB} complex with other ligand-bound LRR protein structures. The ligand is shown in gray in each panel.

(A) Human FSHR_{HB} (red) complexed with FSH.

(B) Human Gp1bα (blue) bound to the VWF A1 domain.

(C) Human spliceosomal U2A' (orange)-U2B'' complex.

(D) Porcine RI (green) bound to RNase A.

(E) *Listeria monocytogenes* InlA (purple) complexed with the N-terminal domain of human E-cadherin.

(F) Human ubiquitin ligase Skp1-Skp2 (cyan) complexed with Cks1.

(G) *Listeria monocytogenes* InlB (pink) bound to the ectodomain of human Met.

(H) Mouse TLR4 (yellow) complexed with MD-2.

(I) The second LRR domain of human Slit (dark green) complexed with the first Ig domain of Robo receptor.