Helical Parameters and Correlations of Tandem Leucine Rich Repeats in Proteins

Purejiv Enkhbayar*, Hiroki Miyashita**, Robert H Kretsinger† and Norio Matsushima**

1Department of Biophysics and Bioinformatics, School of Biology and Biotechnology, National University of Mongolia, Ulaanbaatar 210646/377, Mongolia
2Institute of Tandem Repeats, Sapporo 004-0882, Japan
3Department of Biology, University of Virginia, Charlottesville 22904, USA

*These two authors contribute equally to this study

Abstract

Leucine rich repeats (LRRs) are present in over 20,000 proteins from viruses to eukaryotes. Most LRR units are 20-30 residues long and can be divided into a highly conserved segment and a variable segment. Eight classes have been recognized. Two to sixty-two units occur in tandem to form an LRR structure. The tertiary structures of these LRRs are helical, in which the β-strands of the highly conserved segments stack in parallel. This helix consists of a super helical arrangement of repeating structural units. We call it a coil of solenoids. We have used our program HELFIT to assign helical parameters to 642 LRRs of known structures of 114 proteins. We report these parameters and their correlations with eight classes of LRR, with the number of repeat units in the LRR, with oligomerization, and with ligand state of the LRR. The helical parameters of the eight LRR classes frequently overlap one another. However, the constant distance between parallel β-strands is the primary determinant of the helical parameters of the LRRs. When the repeat number, n, in LRRs is small, the LRR structures are more variable and, by inference, more flexible. In the LRRs with n ≥ 10, az (the rise per repeat unit) of the “RI-like” and “Cysteine-containing” classes is smaller than those of “SDS22-like”, and “Plant-specific” classes. This difference is ascribed mainly to the difference in the structural units. The helical parameters of the LRRs unambiguously describe both right handed and left handed helices, helical dimers, and subdomains if they exist. Moreover, the helical parameters sensitively detect structural changes induced by protein, protein interactions, glycosylation, and/or mutation.

Keywords: HELFIT; LRR; Solenoid; Parallel β-sheet; BR11; Oligomerization; Protein-protein interactions

Abbreviations: 3D: Three dimensional; AMIGO-1: Amphoterin-induced Proteins 1; APL1C: Anopheles Plasmodium-responsive Leucine-rich Repeat Protein 1; AtCO11: Coronative-insensitive Protein 1 from Arabidopsis thaliana; ATRPK2: LRR receptor-like serine/threonine-protein Kinase RPK2 from Arabidopsis thaliana; ATTR1: Transport Inhibitor Response 1 from Arabidopsis thaliana; AtTMK1: Probable Receptor Protein Kinase TMK1 from Arabidopsis thaliana; BAK1: BRASSINOSTEROID INSENSITIVE 1-associated Receptor Kinase 1; BR1: Protein BRASSINOSTEROID INSENSITIVE 1; BR1L: Serine/threonine-protein Kinase BR1-like 1; CD14: Myeloid Cell-Specific Leucine-Rich Glycoprotein; C: alpha-carbon; DCL1: Flagellar Outer Arm Dynein Light Chain 1; dRNA: Double-stranded RNA; FSHR: Follicle Stimulating Hormone Receptor; FSL2: LRR Receptor-Like Serine/Threonine-Protein Kinase FLS2; Glbβ: Glycoprotein Ib alpha; GPIX: Glycoprotein IX; HCS: Highly Conserved Segment; HEL: Anti-Hen Egg White Lysozyme; InlA: Internalin A; InlB: Internalin B; InlC: Internalin C; InlF: Internalin F; InlH: Internalin H; InlJ: Internalin J; InlK: Internalin K; IIR: Island region; LEGL7: Legionella Leucine-Rich Repeat-Containing Protein; LG: Leucine-Rich Repeat-Containing G Protein-Coupled Receptor; LINGO-1: Leucine-Rich Repeat and Immunoglobulin-Like Domain-Containing Nogo Receptor-Interacting Protein 1; LRIM1: Leucine-rich Immune Molecule 1; LRR: Leucine Rich Repeat; NGL-2: Ntrin-G2 Ligand/ Leucine-Rich Repeat-Containing Protein 4; NGL-3: Ntrin-G3 Ligand/ Leucine-Rich Repeat-Containing Protein 4b; Ngr-4: Reticulon-4 receptor/Nogo receptor/Nogo-66 receptor; NLR: NLR Family CARD Domain Containing Protein 4; NLRX1: NLR Family Member XI; P3AP2/ANP32A/LANP: Acidic Leucine-rich Nuclear Phospho-protein 32 Family Member A; PGIP: Polypegalacturonase Inhibitor; RabGTTA: Geranylgeranyl Transferase type-2 Subunit Alpha; RanGAP: Ran GTPase Activating Protein 1; RI: Ribonucleosome Inhibitor; RMSD: The Root Mean Square Deviation from the best fit helix; SERK1: Somatic Embroygenesis Receptor Kinase 1; Skp2: S-phase Kinase-Associated Protein 2; SpH: Salmonella Secretrt Protein H; TAP: mRNA export factor TAP; TLR: Toll-like Receptor; Trk-A: Tyrosine Kinase Receptor A; TSHR: Thyroid-Stimulating Hormone Receptor; U2 snRNP A': U2 Small Nuclear Ribonucleoprotein A'; VLR: Variable Lymphocyte Receptor; VS: Variable Segment; YopM: Leucine-rich Effector Protein from Yersinia pestis

Introduction

LRRs are present in over 20,000 proteins from viruses to eukaryotes, as listed in the data bases of PFAM, SMART, PROSITE, and InterPro [1-4]. All organisms for which sequence data are available have at least one LRR protein. Most LRR proteins are involved in protein, protein interactions, as observed in the plant immune response and the mammalian innate immune response [5-9]. Beyond immunity, extensive functional diversity occurs among LRR proteins, including their involvement in apoptosis, autophagy, ubiquitin related processes, nuclear mRNA transport, neuronal development, and the type III secretion system in pathogenic bacteria [10-16]. Furthermore, plant LRR proteins including LRR containing receptor-like kinases and LRR

*Corresponding authors: Norio Matsushima, Institute of Tandem Repeats, Sapporo 060-8556, Japan, Tel/Fax: +81 11 886 0087; E-mail: norio_irrako@outlook.jp

Robert H Kretsinger, Department of Biology, University of Virginia, Charlottesville 22904, USA, Tel: 434-982-5764; Fax: 434-982-0019; E-mail: rk5i@virginia.edu

Received June 03, 2014; Accepted June 20, 2014; Published June 24, 2014

Citation: Enkhbayar P, Miyashita H, Kretsinger RH, Matsushima N (2014) Helical Parameters and Correlations of Tandem Leucine Rich Repeats in Proteins. J Proteomics Bioinform 7: 139-150. doi:10.4172/jpb.1000314

Copyright: © 2014 Enkhbayar P et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited
containing receptor like proteins act as signal amplifiers in the cases of tissue damage, of establishing symbiotic relationships, and of affecting developmental processes [17-19].

Two to sixty-two LRR units occur in tandem. Each unit is typically 20–30 residues long and can be divided into a highly conserved segment (HCS) and a variable segment (VS) [6,20,21]. The HCS consists of an eleven residue stretch, LxxLxxNxLxL, or a twelve residue stretch, LxxLxLxxCxLx, in which “L” is Leu, Ile, Val, or Phe; “N” is Asn, Thr, Ser, or Cys; and “C” is Cys, Ser or Asn. Eight classes of LRRs—"RI-like", "Cysteine-containing", "SDS22-like", "IRREKO", "Bacterial", "Plant-specific", "Typical", and "TpLRR"—have been recognized [6,20–22]. They are characterized by different lengths and consensus sequences of the VS of their repeat units. Most of the known LRRs consist of only one class of units; however a few have a mixture of two classes. A super-motif of LRRs occurs in a group of LRR proteins including the subfamily of small LRR proteoglycan including biglycan, and decorin, and TLKs 7, 8, and 9 [23–25]. Moreover, the first, N-terminal LRR unit in the "Typical" or "IRREKO" LRR domains is frequently occupied by a "Bacterial" motif [21]. Crystal and/or solution structures of representatives of all the eight classes of LRR are available.

Three residues at positions 3 to 5, xLx, in the HCS form a short β-strand [5–9]. These β-strands from tandem LRR units stack parallel and then the LRRs form an arc or a helix. This helix is not simple but consists of a super helical arrangement of repeating structural units. We call it a coil of solenoids.

The concave face, consisting of the HCSs of each LRR unit, consists of a parallel β-sheet, one strand from each HCS unit. The convex face, consisting of the VS of each LRR unit, is made of a variety of secondary structures including the α-helix, 3_10-helix, polyproline II helix, and an extended conformation or a tandem arrangement of β-turns (Figure 1). The various secondary structures on the convex side are connected to the strands forming the β-sheet in its concave side by two loops (Figure 1). The "ascending loop" links the C-terminal end of the β-strand in the HCS to the N-terminus of the characteristic secondary structure in the VS [9] (Figure 1). The "descending loop" links the C-terminal end of the characteristic secondary structure in the VS to the N-terminus of the β-strand in the HCS of the following unit. Most of the known LRR structures have an N- and/or C-cap that shields the hydrophobic core of the first LRR unit at the N-terminus and/or the last unit at the C-terminus [5–9]. In extracellular proteins or extracellular regions, these caps frequently consist of Cys clusters consisting of two or four Cys residues on the N- and C-terminal sides of the LRRs.

Non-LRR, Island Regions, IRs, interrupting LRRs are widely distributed; they are referred to as "islands" or "loop outs" [26–28]. A large number of plant LRR proteins including LRR containing receptor like kinases and LRR containing receptor like proteins have IRs [24,28]. In our analyses we treat the LRR regions on either side of its IR as distinct helices, as well as analyzing the entire LRR as a single helix.

Enkhbayar et al. [29–31] developed a program, HELFIT, that determines the helical parameters-helix axis, pitch (P), helix radius (R), and number of repeat units per turn (N). HELFIT [29–31] computes these parameters in which the helix axis is represented by the unit vector (Figure 2A). These parameters also yield the rise per repeat unit (Azz=P/N) and the rotation per repeat unit (ΔΦ=360°/N). Moreover, HELFIT gives rmsd:

\[
\text{rmsd}=\left[\frac{\sum d_i^2}{n}\right]^{1/2}
\]

where \(d_i\) is the closest distances from data point to the trace of the helix. Here \(p=\frac{\text{rmsd}}{(n-1)^{1/2}}\) gives the regularity of helix independent of its length. If \(p\) is relatively large, the LRR is not well described by a helix and might be more easily visualized as another shape, e.g. an ellipse.

In most LRRs the three residues at positions 3-5 in the HCS form a short β-strand, which is almost completely conserved in all LRR structures [5–9]. These β-strands from adjacent repeat units form a parallel, pleated sheet on the inside, or concave, surface of the helix. As reference points for HELFIT we use the coordinates of the α-carbon (Cα) of the consensus leucine residue at position 4 (corresponding to...
Helical parameters

Three structures of LRRs with n=3 have been determined. In these cases we performed 3D circle fitting that directly fits a circle to a set of 3D data points, because only three points define a circle [32]. The 3D circle fitting, meaning P=0.0 and Δx=0.0 as an a priori assumption determines the radius of the circle, the number per circle, the center of the circle, and the root mean square deviation from the best fit circle. The 3D circle fitting was also utilized for the estimation of the average distance between adjacent repeats by the helical parameters. Furthermore, the center of the circle was used for the calculation of the distance between the two monomers in TLR dimers.

For grouping of eight LRR classes of known LRR domains we performed similarity searches by FASTA and by BLAST using representative LRR domains clearly belonging to individual classes as query sequences. The representative LRR domains were selected from those for which the 3D structures have been solved. “RI-like” LRRs are from RI, “Cysteine-containing” from Skp2, “SDS22-like” from InlA, “1RREKO” from InlJ, “Bacterial” from YopM, “Plant-specific” from FLS2, “Typical” from LINGO-1, and “TpLRR” from BACCAC_03700.

Secondary structure assignments based on 3D coordinates were made using the DSSP program (http://crdd.osdd.net/raghava/ccpdb/beta2_up.php) [33].

The correlation of Δz, ΔΦ, and R

\[ D = \left[ \frac{2R \sin(\Delta \Phi)}{2} + (\Delta z)^2 \right]^{1/2} \]  

The data points of 2·R·sin(ΔΦ/2) versus Δz fall on a circle with radius D, although, in the “TpLRR” class the data deviate from this circle (Figure 3). The circle fittings using in total 627 data points (except for those of the “TpLRR” proteins) gives D=5.02 ± 0.00 Å [32]. This D corresponds to the inter-strand distance that is in the range of 4.5 to 5.5 Å [146]; this distance allows the formation of hydrogen bonds between parallel strands. In a canonical, parallel β-sheet the strands are aligned at the same level parallel to the axes of the individual strands. This occurs in LRR structures; the observed D is consistent with the distance between strands in the canonical parallel β-sheet.

The D value in the “TpLRR” class is 5.52 ± 0.14 Å and thus is larger and more variable than those in other classes. The calculation of the Ca(i)-Ca(i+1) distance revealed D=6.9 Å between neighboring β-strands, as observed in EAEPRAA2165_0102, EUVBEN_01088, BACCAC_03700, and BACVOA_01565. This greater distance arises from water bridges in the hydrogen bonds between the strands.

Most of the LRRs have R=15 → 30 Å (Supplementary Table 1). However, five LRRs in human Trk-A shows very small R (=0.89 Å) [143], while LRRs 7-13 in BT_1240 has very large R (=277 Å). Even in the extreme cases the 2·R·sin(ΔΦ/2) and Δz values fall on the circle in figure 3. In the former the helix axis is nearly perpendicular to the β-strand axis. In contrast, in the latter the helix axis is nearly parallel to the β-strand axis.

The helical parameters of the eight LRR classes frequently overlap one another (Figure 2B). The Δz and 2·R·sin(ΔΦ/2) values in the “Typical” class are distributed broadly (Figure 3).

Structures of LRRs in GPIBβ, a GPIβ/GPIX chimera, and ZP_02034617.1 with n=3 are available (Supplementary Table 1).
but also the entire dimer forms a single helix. We found four examples. Mouse CD14, in which the LRR does not conform to any of the eight classes, forms a dimer in the crystallographic asymmetric unit as well as in solution [139]. Dimerization in the crystal is mediated by LRR residues in the loop between β-strands in the HCS of repeats 12 and 13. The C-terminal β-strands of the two β-sheets from the two monomers interact in an antiparallel fashion and form a large and continuous β-sheet encompassing the entire CD14 dimer. The two monomers are related by a two fold axis of rotation, perpendicular to their common helix axis. The CD14 monomer forms a left handed helix, as does the CD14 dimer (Figure 5 and supplementary Table 1).

Human RI complexed with human angiogenin or ribonuclease I forms a dimer with human RI-angiogenin/ribonuclease I [36,37]. The whole molecule of human RI consists of only LRRs; all belong to “RI-like” class. The two complexes are held together by many hydrogen bonds formed between the N-terminal β-strands of the two human RI molecules, leading to the formation of an antiparallel β-sheet. The entire RI dimer adopts a right handed helix whose helical parameters are similar to those of its individual monomers. These two hRI molecules are related by an approximate two fold axis of rotation, nearly perpendicular to the common helical axis.

Mouse RI bound to mouse ribonuclease I forms a tetramer in the crystallographic asymmetric unit. The tetramer consists of two dimers that are related by a glide plane. The monomers of both dimers adopt a right handed helix whose parameters are quite similar to those in the dimer of human RI- angiogenin/ribonuclease I (Supplementary Table 1).

Frog LGR4, whose LRRs belong to the “RI-like” class, also forms a homo-dimer in the crystal and in solution [104]. The two monomers in the LGR4 dimer are in close proximity at their C-terminal sides and are related by a two fold axis of rotation—TLRs 1-6, TLR8, RP105, LGR4, LGR5, and Drosophila Toll, “TpLRR” class—FAEPRAA2165_01021 and [144]. The circle fitting analysis yields \( D = 4.90 \pm 0.01 \, \text{Å} \) that seems to be smaller than in LRRs with \( n \geq 4 \). This smaller \( D \) might reflect interactions of side chains bridging N- and C-caps [144].

**Left handed helices**

Most LRRs form right handed helices, that is \( \Delta z \approx 0.0 \, \text{Å} \). However, left handed helices are adopted by LEGL7, CD14, Lmo2365_1307, FSHR, TLR1, and TLR2, in which the repeat number \( n \geq 7; \Delta z = -0.78 \rightarrow -3.17 \, \text{Å} \) (Figure 4 and supplementary Table 1).

**Helical dimers in CD14, RI, and LGR4**

In the homo-dimers of some LRRs not only the individual monomer
Most of the p (error) values are larger than 0.20 (Supplementary Table 1). The angle between the two helix axes (Ω) best characterizes the discontinuity (Supplementary Table 2). The LRRs with two subdomains may be grouped into four categories. The overall shape of the entire LRR in TLRs 1-6, TLR8, RP105, and EUBVEN_01088, FAEPRAA2165_01021, and BT_1240 form a homo-dimer in the crystal, in which the two molecules are in close proximity at the C-terminal sides. The LRR unit 7 is 30 residues long and disrupts the continuity of the parallel β-sheet; Ω=64 °→ 70° (Supplementary Table 2).

The LRR domains in human LGR4 and LGR5 with n=19, and frog LGR4 with n=17 are kinked between the central eleventh and twelfth repeats [104-107]; Ω=19.1 °→ 40.1°. The units of LRR11 and of LRR12 do not strictly obey the LxxLxLxxNxL rule. Instead, the Asn residues are replaced by A309 and T332, respectively. As a result, the asparagine ladder (in which the Asn side chains from different turns are stacked and form hydrogen bond connecting turns) breaks in this region, thereby generating two longer β-strands [104-107].

**Homo-oligomers**

Proteins that contain LRRs form homo-dimers, -trimers, -tetramers, -pentamers, -hexamers, and -octamers in crystals. The LRRs of these homo-oligomers are related by non-crystallographic symmetry and hence experience slightly different packing environments. There are two patterns. The first case is that the helical parameters of individual monomers constituting the homo-oligomers are similar to each other. In this case, their repeat numbers are large, n ≥ 11. Examples include, homo-octamers-NLRC4 (n=16), Skp1 (n=11), ATTR1 (n=18), and ATCOI1 (n=18); homo-hexamers-SphH1 (n=10) and biglycan (n=12); homo-pentamer-EUBVEN_01088 14; homo-tetramers-Rangap1 (n=11), LINGO-1 (n=14), Drosophila Toll (n=8), and CARMIL (n=16), and homo-dimers-NGL-3 (n=11) and InLA.
(n=16). In the second pattern the helical parameters differ significantly among the individual monomers. In this case, their repeat numbers are small; n=5 or 7. They include homo-tetramers-TAP, n=5, VLR2913, n=5, bovine coupling factor B, n=5, and human pp32, n=5; the homohexamer-human pp32, n=5; and the homo-octamer-InIK, n=7.

**Eight classes of LRR**

When the repeat number, n, is ≤8, the helical parameters for individual monomers of the homo-oligomers are sometimes variable. The helix parameters were not determined in these cases.

The consensus sequence of the "RI-like" class is LxxLxLxxNx(L/C)xxxoxxLxxoLxxxx. The repeat length is 28-29. Most of their VSs adopt an α-helical conformation (β-α structural units) (Figure 1). "RI-like" LRR proteins include RI (n=17), NLRCA (n=16), CARML, (n=16), and NLRX1 (n=8) [34-40]. The helix parameters range over: \( \text{P} = 6.3 \rightarrow 35.6 \text{ Å}, \Delta z = 0.21 \rightarrow 1.76 \text{ Å}, \Delta \Phi = 11.4 \rightarrow 18.5 \text{°}, N = 19.5 \rightarrow 31.6 \text{ units/repeat}, \text{and} R = 14.7 \rightarrow 24.8 \text{ Å} \) (Supplementary Table 1).

The LRR domains of RanGAP from human and fungi, n=11, and tropomodulin from chicken and Caenorhabditis elegans, n=5, show little similarity to the RI LRRs [41-45]. However, their repeat units are in β-α conformations that are the same as those of the "RI-like" class. Correspondingly, their helical parameters are similar to those of the four "RI-like" LRR proteins.

The LRR domain in LEGL7 is similar to that in CARMIL (E-value=6.2 · 10^-4 in FASTA). We tentatively assign the LEGL7 LRR to the "RI-like" class. The LRR domain consists of eleven units. The consensus is VxxLxLxxNxLxxxNLxxLxxPxx with 29 residues. DSSP analysis indicates that the "ascending loops" in the eight repeats adopt a \( \text{3}_1 \text{-} \text{helix of one turn at the underlined residues and the VS adopts an \( \alpha \)-helix with 6 \rightarrow 10 \) residues. The LRR domain adopts a left-handed helix; \( \text{P}=26.9 \text{ Å}, \Delta z = -1.53 \text{ Å}, \Delta \Phi = 20.5\text{°}, N = 17.6 \text{ units/repeat, and} R = 13.7 \text{ Å} \) (Supplementary Table 1).

"Cysteine-containing" LRR proteins include F-box/LRR-repeat protein 20 from *Rattus norvegicus*, grr1 from *Saccharomyces cerevisiae* and grmA from *Emericella nidulans*. The consensus sequence is LxxLxLxxCxxTDxxoxxl(a/g)x(C/L)x.x. Tertiary structures are available for Skp2 with n=8 or 11 [46,47]. The units of Skp2 are relatively variable; they are in β-α conformations (Figure 1). The helix parameters range over: \( \text{P}=28.1 \rightarrow 40.8 \text{ Å}, \Delta z = 1.17 \rightarrow 1.75 \text{ Å}, \Delta \Phi = 15.0 \rightarrow 16.2\text{°}, N = 22.3 \rightarrow 23.4 \text{ units/repeat, and} R = 16.6 \rightarrow 18.9 \text{ Å} \) (Supplementary Table 1).

The "Cysteine-containing" LRR domain in Skp2 is similar to those in ATIR1 and ATCOI1 [48-50]. For both, n=18. The HCS of the LRR domains in ATIR1 and in ATCOI1 are highly variable because the CxxI sequence is not conserved in most of these LRRs. In contrast, the VSs are conserved. We, therefore, assume that both the LRR domains belong to the "Cysteine-containing" class. The variable HCSs form relatively long β-strands with three to six residues and the VSs form α-helices with twelve residues. The \( p \) values are large: \( p=0.33 \rightarrow 0.42 \) Å and thus the right handed helix is distorted. The helix parameters range over: \( \text{P}=19.3 \rightarrow 20.4 \text{ Å}, \Delta z = 1.02 \rightarrow 1.05 \text{ Å}, \Delta \Phi = 18.5 \rightarrow 19.3\text{°}, N = 18.7 \rightarrow 19.1 \text{ units/repeat, and} R = 14.9 \rightarrow 15.3 \text{ Å} \).

The consensus sequence of the "SDS22-like" class is LxxLxLxxNxLxxxLxxLxx. The repeat length is 21-23. The VSs strongly prefer a \( \text{3}_1 \)-helix at the underlined residues. The individual units are in β-3_2_6_ conformation (Figure 1). "SDS22-like" LRR proteins include five distinct internalins (except for InlJ) with n=6 → 16 [51-68]. The "SDS22-like" LRR domains adopt a right-handed helix: \( \text{P}=6.27 \rightarrow 7.7 \text{ Å}, \Delta z = 0.23 \rightarrow 2.91 \text{ Å}, \Delta \Phi = 11.8 \rightarrow 15.1\text{°}, N = 23.9 \rightarrow 30.7 \text{ units/repeat, and} R = 17.4 \rightarrow 22.8 \text{ Å} \) (Supplementary Table 1).

The LRR domain in InIK shows high similarity to that in Lmo2365_1307. These two LRR domains show significant similarity to those in InlA and internalin C2 (with E-values <10^4 in FASTA). We tentatively assume that the LRR domain in InIK and Lmo2365_1307 belong to the "SDS22-like" LRR class. Three variable HCSs at the N-terminal side adopt long β-strands with five or eleven residues. These putative "SDS22-like" domains adopt a left-handed helix: \( \text{P}=62.9 \rightarrow -97.4 \text{ Å}, \Delta z = -1.90 \rightarrow -3.32 \text{ Å}, \Delta \Phi = 10.9 \rightarrow 14.8\text{°}, N = 25.8 \rightarrow 33.1 \text{ units/repeat, and} R = 12.9 \rightarrow 24.7 \text{ Å} \).

The consensus sequence of the "IRREKO" class is LxxLxLxxNxLxxLxxLxxN/L(x)(x)/xx or LxxLxxCxxNxLxxLxxN/L(x)/xx [21]. The only available 3D structure is that of InlJ [69]. InlJ with n=15 belongs to this class; although, the first and the last units are "SDS22-like". The VS of TxlDL(T/S)x(N/L)Tx adopt an extended conformation and \( \alpha \)-turns (Figure 1). Its helical parameters are \( P=73.3 \text{ Å}, \Delta z = 2.80 \text{ Å}, \Delta \Phi = 13.7\text{°}, N = 26.2 \text{ units/repeat, and} R = 17.7 \text{ Å} \) (Supplementary Table 1).

The consensus sequence of the "Bacterial" class is LxxLxLxxNxLxxLxxLxxP(x)xxLpxx. The repeat length is 20-22. The structural units are in β-polypeptide II helix conformation (Figure 1). "Bacterial" LRR proteins include YopM (n=16), SpH2 (n=13), SpH1 (n=10), and ipaH3 (n=9) [147-150]. The helix parameters range over: \( \text{P}=46.7 \rightarrow 103 \text{ Å}, \Delta z = 1.64 \rightarrow 2.61 \text{ Å}, \Delta \Phi = 8.71 \rightarrow 12.7\text{°}, N = 28.4 \rightarrow 36.0 \text{ units/repeat, and} R = 20.6 \rightarrow 24.6 \text{ Å} \) (Supplementary Table 1).

"Plant-specific" LRR proteins include protein FLS2, BRL1, BRL1, BAK1, SERK1, ATTMK1, ATRPK2, and PGIP [70-77]. BRL1 and BRL1 are homologs. BAK1 and SERK1 are also homologs. The consensus of the "plant-specific" class is xxLxxxxLxxNxxGxxLxLxxLxx with 24 residues [22,151]. The AtTMK LRR domain with n=15 is kinked at the central eleventh and twelfth repeats [75]. The consensus sequence, LxGxLp, at positions 11 to 16 forms a second parallel β-strand in the "ascending loop". Thus, the structural unit is β-β-3_2_6_ (Figure 1). The helix parameters with n=9 → 29 range over: \( \text{P}=57.2 \rightarrow 87.0 \text{ Å}, \Delta z = 2.39 \rightarrow 3.37 \text{ Å}, \Delta \Phi = 13.7 \rightarrow 15.0\text{°}, N = 24.0 \rightarrow 26.2 \text{ units/repeat, and} R = 15.2 \rightarrow 17.5 \text{ Å} \) (Supplementary Table 1).

"Typical" LRRs are the most abundant LRR class [6,20]. The consensus sequence is LxxLxLxxNxLxxLxxLxxLxxLxx in which uppercase indicates more than 50% occurrence of a given residue in a certain position; lowercase indicates 30-50% occurrence; "L" is Leu, "N" is Asn, "p" is Pro, "o" indicates a non-polar residue, "F" are Phe, and "x"
Citation: Enkhbayar P, Miyashita H, Kreitsinger RH, Matsushima N (2014) Helical Parameters and Correlations of Tandem Leucine Rich Repeats in Proteins. J Proteomics Bioinform 7: 139-150. doi:10.4172/jpb.1000314

is a non-conserved residue. "Typical" LRR proteins include 53 proteins (Supplementary Table 1) [71, 78-137].

Most of their VSs adopt a tandem arrangement of three β-turns (which we call consecutive -turns) on the convex faces (Figure 1). The consecutive -turns form a flat-tish amphiphilic structure with main chain hydrogen bonds in a linear arrangement. Adjacent β-turn repeat motifs form a regular parallel packing arrangement with the core hydrophobic residues pointing alternately in and out of the plane of the structure interlocking with the corresponding hydrophobic residues from the adjacent repeat. The consecutive -turns have already been reported in the underlined eight residues of the VS consensus of xLxPxGLLxxGLx seen in three LRR units in GPIbα with n=9 [125]. Thus, the structural units may be represented by consecutive -turns.

Phenylalanine at position 19 is internally buried and is surrounded by four Leu's, which form the hydrophobic core of the LRR structure. These Phε's of consecutive units form the spine of the LRR [89]. This Phε spine involves aromatic, aromatic interactions; the benzyl groups form stacks of Ph side chains. The first, N-terminal LRR unit in most of the "Typical" LRR domains is "Bacterial" LRR. The VS adopts the polyproline II helix instead of consecutive β-turns.

The helix parameters with n=7-27 range broadly: P=62.0 → 93.2 Å, Δz=1.53 → 3.24 Å, ΔΦ=7.41 → 150.0°, N=23.9 → 51.2 units/repeat, and R=15.5 → 38.9 Å (Supplementary Table 1). The LRR domains may be grouped into a right handed helix, a left-handed helix, and a near flat.

Human Slit2 contains four tandem arrays (D1 to D4) of LRRs that consist of six to eight units. The structures of D2, n=7, D3, n=7, and D4, n=6, have been determined in the free state [91-93]. The helical parameters differ significantly from each other; D2>D4>D3 in D4, consist of six to eight units. The structures of D2, flat.

be grouped into a right handed helix, a left-handed helix, and a near descending loops consist of a single β-turn (Figure 1). Most of their VSs use nearly its entire concave surface to bind HEL. This binding causes large changes of the helical parameters; a decrease of P, Δz, and ΔΦ, and an increase in R (Supplementary Table 3). The structures of D3 in human Slit2 have been determined in the free state and in the complex with the Ig domain from Robo1 [91]. As seen in the VLRB.2D-HEL complex, the Slit D3 domain uses its concave surface to bind Ig. The formation of the complex causes large changes of the helical parameters. Conversely, an increase of P, Δz, and ΔΦ, and a decrease in R is observed (Supplementary Table 3). The same structural change occurs in the InLA-cadherin complex; although, it is minor (Supplementary Tables 1 and 3) [56].

The crystal structures of free LRIM1 and of free APL1 have been determined and have small changes in structural change induced by protein, protein interactions, glycosylation and/or mutation

The LRRs have been determined in the free state and in complex with REP (Supplementary Table 3). In the LRRs the concave surface, the convex surface, N-terminal domain including N-cap, C-terminal domain including C-cap, or their combinations are involved in protein, protein interactions (Supplementary Table 3).

The LRRs of VLR and Slit2 belong to the "Typical" class, as noted. VLRB.2D, n=5, bound to a protein antigen, HEL, as well as non-bound VLR has been determined [83]. VLR uses nearly its entire concave surface to bind HEL. This binding causes large changes of the helical parameters; a decrease of P, Δz, and ΔΦ, and an increase in R (Supplementary Table 3). The structures of D3 in human Slit2 have been determined in the free state and in the complex with the Ig domain from Robo1 [91]. As seen in the VLRB.2D-HEL complex, the Slit D3 domain uses its concave surface to bind Ig. The formation of the complex causes large changes of the helical parameters. Conversely, an increase of P, Δz, and ΔΦ, and a decrease in R is observed (Supplementary Table 3). The same structural change occurs in the InLA-cadherin complex; although, it is minor (Supplementary Tables 1 and 3) [56].

The crystal structures of free LRIM1 and of free APL1 and of their hetero-dimer complex have been determined [134]. The LRIM1/ APL1 complex has a single intermolecular disulfide bond. LRIM1 and APL1 contain 11 and 15 repeat units, respectively. The HELFIT analysis detects structural change in both LRR domains due to the complex formation. The LRIM1 LRR domain in the complex shows a significant increase in P and Δz and decrease in R (Supplementary Table 3). In contrast, the helical parameters in the APL1 LRR domain are reversed and have small changes in Δz and R.

The structures of frog LGR4 have been determined in the free state and in complex with R-spondin [104]. The formation of the complex induces small structural changes and increases in P, Δz, and R (Supplementary Table 3).

The crystal structures of the complexes with RabGGTα and its complex bound to escort protein 1 have been determined [53,54]. The helical handedness of the LRR domain with n=6 reverses upon binding to REP (Supplementary Table 1).

The crystal structures of the complexes of Skp1-Skp2 and Skp1-Skp2-Cks1 are available [46,47]. The binding of Cks1 causes slight structural changes; an increase of P, Δz, and ΔΦ, and a decrease in R. Moreover, the crystal structures of the complexes with Ran-GPPNPHE- RanBP1-RanGAP and Ran-GDP-AIFx-RanBP1-RanGAP have been determined [43]. The helical parameters of these two complexes differ slightly from one another.

Many structures of GPIbα have been reported in the free state and
in complexes with one or two α-thrombins, von Willebrand Factor A1, and PEP inhibitor [124-131]. We do not find a clear pattern of structural changes. However, it appears that the binding of two α-thrombins causes a structural change.

The crystal structures of NgR-4, n=10, in both the non-glycosylated state and N-glycosylated at Asn-82 and Asn-179, are available [88,89]. N-glycosylation induces a decrease of P, Δz, and ΔΦ and an increase of N and of R (Supplementary Table 3). The G2B4 mutant of bovine coupling factor B, n=5, has a smaller R than does the wild type (Supplementary Table 1) [152].

Spatial arrangement of TLR dimers

The structures of the unliganded and ligand activated human TLR8 dimer were determined [119]. Helix parameters of the two monomers in the unliganded dimer are generally similar to those in the liganded dimer. However, upon ligand binding the spatial arrangement of the two monomers in the dimer changes. The distance between the two monomers at their C-terminal regions changes from 53 Å to 30 Å [119]. HELFIT calculates the distance and the angle between the two helical axes of the two monomers in the dimers, L and Ψ. Correspondingly, the ligand binding drastically changes these two parameters; ΔL=- 2.6 Å and ΔΨ=17° (Figure 7).

The structures of TLR3 complexed with dsRNA and with six different Fabs (Fab15 light and heavy chains, Fab12 light and heavy chains, and Fab1068 light and heavy chains) have been determined, as well as the structure of free TLR3 [112,113]. The formations of the complexes induce decreases in the helix radii. The dsRNA TLR3 dimer has L=74.5 Å and Ψ=148°.

Furthermore, the hetero-dimers of human TLR1 and TLR2, and of mouse TLR2 and TRL6 have been determined [108,109]. The values of L and Ψ are 44.0 Å and 168° in the TLR1-TLR2 dimer and 53.0 Å and 169° in the TLR2-TLR6 dimer, respectively.

Discussion

The constant inter-strand distance is the primary determinant of the helical parameters of the LRRs

The HELFIT analysis demonstrates that the inter-strand distance (D) is constant in the seven classes except for the "TpLRR" class; D=5.02 ± 0.00 Å (Figure 2B). In contrast, D in the "TpLRR" class is significantly larger; D=5.52 ± 0.14 Å. This can be attributed to non-uniformity of the inter-strand distance. The helix radius R in some LRRs belonging to the "TpLRR" class is extremely small or very large (Supplementary Table 1). Thus, the overall shape of these LRRs of the "TpLRR" class is best described as a prism; while those of the other classes are helices or semi-ellipses.

The "ascending loops" influence the helical parameters of the LRRs

The LRRs in RI and LEGL7 belong to the "RI-like" class. The RI LRR adopts a right handed helix or a horse shoe shape. In contrast, the LEGL7 LRR adopts a left handed helix (Figure 4). The LRRs form β-strands on the concave face (in the HCS) and α-helix of 9 → 15 residues on the convex face (in the VS). Both of the "descending loops" also consist of one β-turn. On the other hand, the ascending loops of the LEGL7 LRR include a 310-helix of one turn; while the RI LRR does not adopt this helix in the loops.

Some LRRs in NLRC4, InlK, Lmo2365_1307, AttIR1, AtCO11, TLRs 1-6, and TLR8 do not obey the LxxLxLxNxLxL consensus, especially at the underlined residues. In addition, some LRRs in these proteins frequently form longer β-strands consisting of four to seven residues in the HCS. This affects not only the ascending loops but also the descending loops. Indeed, the right handed helices of NLRC4, AttIR1, AtCO11, TLRs 1-6 and TLR8 are distorted. Also the LRRs in InlK and Lmo2365_1307, which belong to the "SDS2-like" class, adopt left handed helices; while those in the normal "SDS2-like" class form right handed helices.

The HCS of the "TpLRR" class with ten residues is shorter than eleven or twelve residues in the "Typical" class. On the other hand, the VS of the "TpLRR" class is quite similar to that of the "Typical" class. Both VSs consist mainly of consecutive -turns (Figure 1). However, the helical parameters indicate that the "Typical" class adopts a helix, while the "TpLRR" class resembles a prism. These observations can be attributed to a structural difference of the ascending loops and the inter-strand distance.

The helix parameters vary among the "Typical" LRRs. Seven of the ten repeats in NgR-4 and eleven of the fourteen repeats LINGO-1 conform to the "Typical" consensus. However, their structures differ from one another. The NgR-4 LRR domain adopts a right handed helix with Δz=2.09 Å; while the LINGO-1 LRR arc is nearly flat; |Δz|=0.06 Å. The LINGO-1 LRRs show a structural periodicity that consists of four tandem repeats of three consecutive LRR units. The first and second VSs in the three LRR units adopt a β-strand plus consecutive β-turns, while in the last VS the corresponding β-strand breaks. In contrast, this periodicity is not seen in the corresponding residues in NgR-4. The different helical parameters may be due to a difference in structure of the ascending loops and the conformations on the convex face.

In conclusion, the above observations suggest that the helical parameters more strongly depend on the structures of the "ascending loops" than of the "descending loops". Moreover, the helical parameters of the LRRs are influenced by helical elements on the convex face and the uniformity of parallel strand stacking on the concave face.

LRR domains having small repeat number are highly flexible

The values R and Δz versus n for all 642 LRRs show a lot of variation, if n is small (Figure 3). Here we describe some examples. The first is the family of VLR. The lysozyme bound VLR.B2D forms a tetramer in the crystal [83]. The two monomers adopt a right handed helix, while the other two monomers adopt a left handed helix. Sea lamprey VLR.B, n=5, bound to the BclA protein forms a trimer [84]. One monomer forms a right handed and the other monomer forms a left handed helix. The third monomer is not visible in the electron density map of the crystal structure. Tetrameric VLR 2913 has a range of helical parameters in its four monomers (Supplementary Table 1).

The crystal structures of InlB, n=8, in wild type, mutant, and complexes have been determined [60-64,85]. The InlB molecules form a monomer, dimer, and hexamer in the crystals. The helical parameters show a range of variation. InlB with an additional LRR unit inserted, n=9, forms a trimer, All the monomers have similar helical parameters. Right handed and left handed helices exist in the dimer of SERK1, n=5, bound to BRII [71] and in the tetramer of bovine Coupling Factor B, n=5 [152]. InlK, n=7, forms a homo-octameric in the crystal [68]. The helix parameters of these eight monomers are highly variable (Supplementary Table 1). Even with n=17 in RI, the helix parameters are significantly different among four vertebrate species: both P and Δz decrease in order of pig, bovine, human, and mouse [34-37].
These observations strongly suggest that LRR domains with \( n \leq 7 \) are more variable and are more flexible in solution. NMR data of LC1 with \( n=6 \) support this inference; within the \( \beta \)-sheets the consensus leucine residues at position 4 have relatively low order parameters \([52]\).

### Structural features of individual LRR classes

As noted, the helical parameters of the eight LRR classes frequently overlap one another. Moreover, when the repeat number in LRRs is small, the LRR structures are more variable. Taken together, the helical parameters of the LRRs with \( n \geq 10 \) were compared to find the structural features (Figure 3B and supplementary Figure 1). The "TpLRR" class deviates significantly from other classes. The "Typical" class has highly variable values. The rise per subunit, \( \Delta z \), of the "RI-like" and "Cysteine-containing" classes is clearly smaller than those of the "SDS22-like", and "Plant-specific" classes. This is mainly ascribed to the difference in the structural units.

### Structures of non-LRR, IRs

The structures of non-LRR, IRs may be considered in two groups. In the first group, IRs allow the continuity of the parallel \( \beta \)-sheet and the LRR units to form a single domain of a regular LRR structure. Examples include the IRs of BR11, AIRPK2, BT_0210, and TLR8.

The BR11 LRR domain has 25 units interrupted by a 70 residue IR between unit 21 and unit 22 \([22]\). B. thetaiaomicron BT_0210 contains 18 units interrupted by a 76-residue IR between units 2 and 3. In the two LRR domains the helical handedness reverses (Supplementary Table 1). The IR forms a small domain those folds back into the interior of the helix, where it has extensive polar and hydrophobic interactions. The IR between repeat units 15 and 16 of TLR8 is 39 residues long \([24]\). The additional residues loop out from the expected helical path of a regular LRR before rejoining it some residues later \([119]\). Consequently, this IR is in the interior of the helix, in contrast to the IRs in BR11 and BT_0210.

In the second case IRs disrupt the continuity of the parallel \( \beta \)-sheet and form two distinct subdomains. This case is observed in AtTMK1 \([75]\) and Drosophila Toll \([95]\). The AtTMK1 LRR domain contains 15 repeats interrupted by a 41 residue IR between LRR11 and LRR12 \([28]\), while the Toll LRR domain with 24 repeats has a 68 residue IR between LRR10 and LRR11\([95]\) .

### Conclusions

We have assigned helical parameters to the 642 LRRs of known structure using the program HELFIT. These parameters and their correlations with class of LRR, with the number of repeat units in the LRR, and with oligomerization and ligand state of the LRR are described. The helical parameters of the eight LRR classes frequently overlap one another. However, the constant distance between parallel \( \beta \)-strands is the primary determinant of the helical parameters of the LRRs. When the repeat number, \( n \), is small, the LRR structures are more variable and, by inference, more flexible. In the LRRs with \( n \geq 10 \), rise per repeat unit, \( \Delta z \), of the "RI-like" and "Cysteine-containing" classes are smaller than those of "SDS22-like", and "Plant-specific" classes. This is ascribed to the difference in the structural units. The helical parameters of the LRRs unambiguously define both right handed and left handed helices, and helical dimers and subdomains if they exist. Moreover, the helical parameters sensitively detect structural changes induced by protein, protein interactions, glycosylation, and/or mutation.

### Acknowledgements

We are grateful to Tuvdendorj Nomin-Erdene for her help of the HELFIT analysis. We also thank Chai JiJie of Tsinghua University for his providing of the PDB file of AIRPK2.

### References

1. Punta M, Coggill PC, Eberhardt RY, Misty J, Tate J, et al. (2012) The Pfam protein families database. Nucleic Acids Res 40: D290-301.
2. Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res 40: D302-305.
3. Sigrist CJ, de Castro E, Cerutti L, Cucchi BA, Hulo N, et al. (2013) New and continuing developments at PROSITE. Nucleic Acids Res 41: D344-347.
4. Burge S, Kelly E, Lonsdale D, Mutowu-Muellenet P, McAnulla C, et al. (2012) Manual GO annotation of predictive protein signatures: the InterPro approach to GO curation. Database (Oxford) 2012: bar068.
5. Kobe B, Deisenhofer J (1994) The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci 19: 415-421.
6. Kobe B, Kajava AV (2001) The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol 11: 725-733.
7. Matsushima N, Enkhbayar P, Kamiya M, Osaki M, and Kretsinger R (2005) Leucine-Rich Repeats (LRRs): Structure, Function, Evolution and Interaction with Ligands. Drug Design Rev 2: 305-322.
8. Matsushima N, Tachi N, Kuroki Y, Enkhbayar P, Osaki M, et al. (2005) Structural analysis of leucine-rich-repeat variants in proteins associated with human diseases. Cell Mol Life Sci 62: 2771-2791.
9. Bella J, Hindle KL, McEwan PA, Lovell SC (2008) The leucine-rich repeat structure. Cell Mol Life Sci 65: 2307-2333.
10. Kofoed EM, Vance RE (2012) NAIPs: building an innate immune barrier against bacterial pathogens. NAIPs function as sensors that initiate innate immunity by detection of bacterial proteins in the host cell cytosol. Bioessays 34: 589-598.
11. Ng A, Xavier RJ (2011) Leucine-rich repeat (LRR) proteins: integrators of pattern recognition and signaling in immunity. Autophagy 7: 1082-1084.
12. de Wit J, Hong W, Luo L, Ghosh A (2011) Role of leucine-rich-repeat proteins in the development and function of neural circuits. Annu Rev Cell Dev Biol 27: 697-729.
13. McPhee JB, Mena P, Zhang Y, Bliska JB (2012) Interleukin-10 induction is an important virulence function of the Yersinia pseudotuberculosis type III effector YopM. Infect Immun 80: 2519-2527.
14. Furlan G, Klinkenberg J, Trujillo M (2012) Regulation of plant immune receptors by ubiquitination. Front Plant Sci 3: 238.
15. Ho DN, Coburn GA, Kang Y, Cullen BR, Georgiadis MM (2002) The crystal structure and mutational analysis of a novel RNA-binding domain found in the human Tap nuclear mRNA export factor. Proc Natl Acad Sci U S A 99: 1888-1893.
16. Jiang J, Zhang C, Wang X (2013) Ligand perception, activation, and early signaling of plant steroid receptor brassinosteroid insensitive 1. J Integr Plant Biol 55: 1198-1211.
17. Tör M, Lotze MT, Holton N (2009) Receptor-mediated signalling in plants: molecular patterns and programmes. J Exp Bot 60: 3645-3654.
18. Steinwand BJ, Kieber JJ (2010) The role of receptor-like kinases in regulating cell wall function. Plant Physiol 153: 479-484.
19. Bernoux M, Ellis JG, Dodds PN (2011) New insights in plant immunity signaling activation. Curr Opin Plant Biol 14: 512-518.
20. Kajava AV (1998) Structural diversity of leucine-rich repeat proteins. J Mol Biol 277: 519-527.
21. Matsushima N, Miyashita H, Mikami T, Kuroki Y (2010) A nested leucine rich repeat (LRR) domain: the precursor of LRRs is a ten or eleven residue motif. BMC Microbiol 10: 235.
22. Kajava AV, Anisimova M, Peeters N (2008) Origin and evolution of GALA-LRR, a new member of the CC-LRR subfamily: from plants to bacteria? PLoS One 3: e1694.
23. Matsushima N, Ohyanagi T, Tanaka T, Kretsinger RH (2000) Super-motifs and evolution of tandem leucine-rich repeats within the small proteoglycans-biglycan, decorin, lumican, fibromodulin, PRELP, keratocan, osteoadherin, epiphycan, and osteoglycin. Proteins 38: 210-225.
24. Matsushita N, Tanaka T, Enkhbayar P, Mikami T, Taga M, et al. (2007) Comparative sequence analysis of leucine-rich repeat (LRRs) within vertebrate toll-like receptors. BMC Genomics 8: 124.

25. Miyashita H, Kuroki Y, Matsushita N (2014) Novel leucine rich repeat domains in proteins from unicellular eukaryotes and bacteria. Protein Pept Lett 21: 292-305.

26. Fritz-Laylin LK, Krishnamurthy N, Tör M, Sjölander KV, Jones JD (2005) Phylogenetic analysis of the receptor-like proteins of rice and Arabidopsis. Plant Physiol 138: 611-623.

27. Matsushita N, Mikami T, Tanaka T, Miyashita H, Yamada K, et al. (2009) Analyses of non-leucine-rich repeat (non-LRR) regions intervening between LRRα in proteins. Biochim Biophys Acta 1790: 1217-1237.

28. Matsushita N, Miyashita H (2012) Leucine-rich Repeat (LRR) domains containing intervening motifs in plants. Biopolymers 2: 288-311.

29. Enkhbayar P, Hikiki K, Osaki M, Kretzinger RH, Matsushita N (2006) 3(10)-helices and P helices paralelling a 3(10)-helix. Proteins 64: 691-699.

30. Enkhbayar P, Damdinsuren S, Osaki M, Matsushita N (2008) HELIXFIT: Helix fitting by a total least squares method. Comput Biol Chem 32: 307-310.

31. Enkhbayar P, Matsushita N (2012) HELIXFIT: Algorithm and applications. AIP Conf. Proc 1479: 1942-1947.

32. Enkhbayar P, Kamiya M, Osaki M, Matsutomo T, Matsushita N (2004) Structural principles of leucine-rich repeat (LRR) proteins. Proteins 54: 394-403.

33. Kabach W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22: 2577-2637.

34. Kobe B, Deisenhofer J (1995) A structural basis of the interactions between leucine-rich repeats and protein ligands. Nature 374: 183-186.

35. Kobe B, Deisenhofer J (1996) Mechanism of ribonuclease inhibition by ribonuclease inhibitor protein based on the crystal structure of its complex with ribonuclease-A. J Mol Biol 264: 1029-1043.

36. Johnson RJ, McCoy JG, Bingman CA, Phillips GN Jr, Raines RT (2007) Inhibition of human pancreatic ribonuclease by the human ribonuclease inhibitor protein. J Mol Biol 368: 434-449.

37. Papageorgiou AC, Shapiro R, Acharya KR (1997) Molecular recognition of the RNA-binding element of the NLRX1 innate immune modulator. Immunity 7: 781-791.

38. Hillig RC, Renault L, Vetter IR, Drell T (2004) Structural genomics analyses of non-leucine-rich repeat (non-LRR) regions intervening between LRRα in proteins. J Mol Biol 342: 2716-2725.

39. Wu H, Maciejewski MW, Marintchev A, Benashski SE, Mullen GP, et al. (2000) Solution structure of a dynein motor domain associated light chain. Nat Struct Biol 7: 575-579.

40. Wu H, Blackledge M, Maciejewski MW, Mullen GP, King SM (2003) Relaxation-based structure refinement and backbone molecular dynamics of the dynein motor domain-associated light chain. Biochemistry 42: 57-71.

41. Zhang H, Seabra MC, Deisenhofer J (2000) Crystal structure of Rab geranylgeranyltransferease at 2.0 A resolution. Structure 8: 241-251.

42. Pylepenko O, Rak A, Reents R, Niculceas A, Sidoriwich V, et al. (2003) Structure of Rab escort protein-1 in complex with Rab geranylgeranytransferase. Mol Cell 11: 483-494.

43. Price SR, Evans PN, Nakagai K (1998) Crystal structure of the spilosecional U2BF-*U2A* protein complex bound to a fragment of U2 small nuclear RNA. Nature 394: 645-650.

44. Schubert WD, Urbaric C, Ziehm T, Beier V, Machner MP, et al. (2002) Structure of internalin B protein complex from Listeria monocytogenes in complex with its human receptor E-cadherin. Cell 111: 825-836.

45. Wollet T, Heinz DW, Schubert WD (2007) Thermodynamically reengineering the llisterial invasion complex InlAE-cadherin. Proc Natl Acad Sci U S A 104: 13960-13965.

46. Marino M, Braun L, Cossart P, Ghosh P (1999) Structure of the InlB leucine-rich repeat, a domain that triggers host cell invasion by the bacterial pathogen L. monocytogenes. Mol Cell 4: 1063-1072.

47. Schubert WD, Göbel G, Diephotlz M, Darji A, Kloer D, et al. (2001) Internalins from the human pathogen Listeria monocytogenes combine three distinct folds into a contiguous internalin domain. J Mol Biol 312: 753-784.

48. Ferraris DM, Gherardi E, Di V, Heinz DW, Niemann HH (2010) Ligand-mediated dimerization of the Met receptor tyrosine kinase by the bacterial invasion protein InlB. J Mol Biol 395: 522-532.

49. Marino M, Banerjee M, Jonquière R, Cossart P, Ghosh P (2002) GW domains of the Listeria monocytogenes invasion protein InlB are SH3-like and mediate binding to host ligands. EMBO J 21: 5623-5634.

50. Marino M, Banerjee M, Copp J, Dromsi S, Chapman T, et al. (2004) Characterization of the calcium-binding sites of Listeria monocytogenes InlB. Biochem Biophys Res Commun 316: 379-386.

51. Niemann HH, Jäger V, Butler PJ, van den Heuvel J, Schmidt S, et al. (2007) Structure of the human receptor tyrosine kinase in complex with the Listeria monocytogenes invasion protein InlB. Cell 130: 235-246.

52. Niemann HH, Gherardi E, Bleymüller WM, Heinz DW (2012) Engineered variants of InlB with an additional leucine-rich-repeat domain discriminate between physiologically relevant and packing contacts in crystal structures of the InlB:MET complex. Protein Sci 21: 1528-1539.

53. Breitsprecher D, Gherardi E, Bleymüller WM, Niemann HH (2014) Crystal structure of an engineered YopM-InlB hybrid protein. BMC Struct Biol 14: 12.

54. Ooi A, Hussain S, Seyedarabi A, Pickersgill RW (2006) Structure of internalin A from Yersinia pestis: oligomeric and functional characterization. EMBO J 25: 5623-5634.

55. Breitsprecher D, Gherardi E, Bleymüller WM, Niemann HH (2014) Crystal structure of an engineered YopM-InlB hybrid protein. BMC Struct Biol 14: 12.

56. Ooi A, Hussain S, Seyedarabi A, Pickersgill RW (2006) Structure of internalin A from Yersinia pestis: oligomeric and functional characterization. EMBO J 25: 5623-5634.

57. Breitsprecher D, Gherardi E, Bleymüller WM, Niemann HH (2014) Crystal structure of an engineered YopM-InlB hybrid protein. BMC Struct Biol 14: 12.

58. Ooi A, Hussain S, Seyedarabi A, Pickersgill RW (2006) Structure of internalin A from Yersinia pestis: oligomeric and functional characterization. EMBO J 25: 5623-5634.

59. Breitsprecher D, Gherardi E, Bleymüller WM, Niemann HH (2014) Crystal structure of an engineered YopM-InlB hybrid protein. BMC Struct Biol 14: 12.

60. Ooi A, Hussain S, Seyedarabi A, Pickersgill RW (2006) Structure of internalin A from Yersinia pestis: oligomeric and functional characterization. EMBO J 25: 5623-5634.

61. Breitsprecher D, Gherardi E, Bleymüller WM, Niemann HH (2014) Crystal structure of an engineered YopM-InlB hybrid protein. BMC Struct Biol 14: 12.
69. Bublitz M, Holland C, Sabet C, Reichelt J, Cossart P, et al. (2008) Crystal structure and standardized geometric analysis of HiL, a listerial virulence factor and leucine-rich repeat protein with a novel cysteine ladder. J Mol Biol 378: 87-96.

70. Sun Y, Li L, Macho AP, Han Z, Hu Z, et al. (2013) Structural basis for fg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex. Science 342: 624-628.

71. Santiago J, Henzler C, Hothorn M (2013) Molecular mechanism for plant steroid receptor activation by somatic embryogenesis co-receptor kinases. Science 341: 889-892.

72. Hothorn M, Belkhaled Y, Dreux M, Dabi T, Noel JP, et al. (2011) Structural basis of steroid hormone perception by the receptor kinase BR1. Nature 474: 467-471.

73. She J, Han Z, Kim TW, Wang J, Cheng W, et al. (2011) Structural insight into brassinosteroid perception by BR1. Nature 474: 472-476.

74. She J, Han Z, Zhou B, Chai J (2013) Structural basis for differential recognition of brassinolide by its receptors. Protein Cell 4: 475-482.

75. Liu P, Hu Z, Zhou B, Liu S, Chai J (2013) Crystal structure of an LRR protein with two selenoïds. Cell Res 23: 303-305.

76. Song W, Han Z, Sun Y, Chai J (2014) Crystal structure of a plant leucine rich repeat protein with two island domains. Sci China Life Sci 57: 137-144.

77. Di Matteo A, Federici L, Mattei B, Salvi G, Johnson KA, et al. (2003) The crystal structure of polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. Proc Natl Acad Sci U S A 100: 10124-10128.

78. Kim HM, Oh SC, Lim KJ, Kamatsumas J, Heo JY, et al. (2007) Structural diversity of the hagfish variable lymphocyte receptors. J Biol Chem 282: 6726-6732.

79. Han BW, Herrin BR, Cooper MD, Wilson IA (2008) Antigen recognition by variable lymphocyte receptors. Science 321: 1834-1837.

80. Deng L, Velikovsky CA, Xu G, Iyer LM, Tasumi S, et al. (2010) A structural basis for antigen recognition by the T cell-like lymphocytes of sea lamprey. Proc Natl Acad Sci U S A 107: 13408-13413.

81. Luo M, Velikovsky CA, Yang X, Siddiqui MA, Hong X, et al. (2013) Recognition of the Thomsen-Friedenreich pan cancroma carbohydrate antigen by a lamprey variable lymphocyte receptor. J Biol Chem 288: 23597-23606.

82. Kanda R, Sutoh Y, Kasamatsu J, Kasahara M, et al. (2014) Crystal structure of the lamprey variable lymphocyte receptor C reveals an unusual feature in its N-terminal capping module. PLoS One 9: e85875.

83. Velikovsky CA, Deng L, Tasumi S, Iyer LM, Kerzic MC, et al. (2009) Structure of a lamprey variable lymphocyte receptor in complex with a protein antigen. Nat Struct Mol Biol 16: 725-730.

84. Kirchdoerfer RN, Herrin BR, Han BW, Turnbough CL Jr, Cooper MD, et al. (2003) The crystal structure of leucine-rich repeat protein with two island domains. Sci China Life Sci 57: 137-144.

85. Jiang X, Liu H, Chen X, Chen PH, Fischer D, et al. (2012) Structure of fлизide-stimulating hormone in complex with the entire ectodomain of its receptor. Proc Natl Acad Sci U S A 110: 20461-20466.

86. Kim HM, Oh SC, Lim KJ, Kamatsumas J, Heo JY, et al. (2007) Structural diversity of the hagfish variable lymphocyte receptors. J Biol Chem 282: 6726-6732.

87. Sanders, J, Chirgadze DY, Sanders, P, Baker, S, Sullivan, A, et al. (2007) Crystal structure of the TSH receptor (TSHR) bound to a blocking-type TSHR autoantibody. J Mol Endocrinol 46: 81-99.

88. Sanders, J, Kundra-Panula, J, Rauvala, H, Goldman, A (2011) Crystal structure and role of glycans and dimerization in folding of neuronal leucine-rich repeat protein AMIGO-1. J Mol Biol 413: 1001-1015.

89. Luo K, Xu Y, Rajashehankar KR, Robev D, Nikolov DB (2013) Crystal structures of Lgr4 and its complex with R-spondin1. Structure 21: 1683-1689.

90. She J, Han Z, Kim TW, Wang J, Cheng W, et al. (2011) Structural basis for differential recognition of Lgr4 and its complex with R-spondin. Genes Dev 27: 1345-1350.

91. Sanders, J, Chirgadze, D, Sanders, P, Baker, S, Sullivan, A, et al. (2007) Crystal structure of the TSH receptor in complex with a thyroid-stimulating autoantibody. Thyroid 17: 395-410.

92. Morlot C, Hemrika W, Romijn RA, Gros P, Cusack S, et al. (2007) Production of Slit2 Lissabonocytes in mammalian cells for structural studies and the structure of human Slit2 domain 3. Acta Crystallogr D Biol Crystallogr 63: 961-968.

93. Seiradake E, von Philipsborn AC, Henry M, Fritz M, Lortat-Jacob H, et al. (2014) Structural basis and functional relevance of the Slit2 homodimerization domain. EMBO Rep 10: 736-741.

94. Howitt JA, Clout NJ, Hohenerste E (2004) Binding site for Robo receptors revealed by dissection of the leucine-rich repeat region of Slit. EMBO J 23: 4406-4412.

95. Panteri X, Stelter M, Ursei C, Fandrich U, Lilie H, et al. (2014) Structure of the Toll-Splaztle complex, a molecular hub in Drosophila development and innate immunity. Proc Natl Acad Sci U S A 111: 6281-6286.

96. Gangloff M, Arnott CJ, Lewis M, Gay NJ (2013) Functional insights from the crystal structure of the N-terminal domain of the prototypical toll receptor. Structure 21: 143-153.

97. Gangloff M, Moreno A, Gay NJ (2013) Liesegang-patterns of Toll crystals grown in gel. J Appl Crystallogr 46: 337-345.

98. Lewis M, Arnott CJ, Beehston E, McCoy A, Ashcroft AE, et al. (2013) Cytokine Spalze binds to the Drosophila immunoreceptor Toll with a neurotrophin-like specificity and couples receptor activation. Proc Natl Acad Sci U S A 110: 20461-20466.

99. Jiang X, Liu H, Chen X, Chen PH, Fischer D, et al. (2012) Structure of fлизide-stimulating hormone in complex with the entire ectodomain of its receptor. Proc Natl Acad Sci U S A 110: 12491-12496.

100. Fan QR, Hendrickson WA (2005) Structure of human fлизide-stimulating hormone in complex with its receptor. Nature 433: 269-277.
115. Park BS, Song DH, Kim HM, Choi BS, Lee H, et al. (2009) The structural basis of lipopolysaccharide recognition by the Toll-TRD-2 complex. Nature 458: 1191-1195.

116. Ohno U, Yamakawa N, Akashi-Takamura S, Miyake K, Shimizu T (2012) Structural analyses of human Toll-like receptor 4 polymorphisms D299G and T399I. J Biol Chem 287: 40611-40617.

117. Zhou K, Kanai R, Lee P, Wang HW, Modis Y (2012) Toll-like receptor 5 forms asymmetric dimers in the absence of flagellin. J Struct Biol 177: 402-409.

118. Yoon SI, Kurnasov O, Natarajan V, Hon M, Gudkov AV, et al. (2012) Structural basis of TLR5-flagellin recognition and signaling. Science 335: 859-864.

119. Tanji H, Ohoto U, Shibata T, Miyake K, Shimizu T (2013) Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. Science 339: 1426-1429.

120. Kokatla HP, Sil D, Tanji H, Ohoto U, Malladi SS, et al. (2014) Structure-based design of novel human Toll-like receptor 8 agonists. ChemMedChem 9: 719-733.

121. Yoon SI, Hong M, Wilson IA (2011) An unusual dimeric structure and assembly for TLR4 regulator RIP105-2-1. Nat Struct Mol Biol 18: 1028-1035.

122. Ohno U, Miyake K, Shimizu T (2011) Crystal structures of mouse and human RIP105-2-1 complexes reveal unique dimer organization of the Toll-like receptor family. J Mol Biol 413: 815-825.

123. Zhao Y, Malinauskas T, Harlos K, Jones EY2 (2014) Structural insights into the inhibition of Wnt signaling by cancer antigen 514/Wnt-activitated inhibitory factor 1. Structure 22: 612-620.

124. Celikel R, McClintock RA, Roberts JR, Mendicoll-Mobile GL, Ware J, et al. (2003) Modulation of alpha-thrombin function by distinct interactions with platelet glycoprotein Ibalpha. Science 301: 218-221.

125. Uff S, Clemenson JM, Harrison T, Clemenson KU, Emsley J (2002) Crystal structure of the platelet glycoprotein Ibalpha (alpha) N-terminal domain reveals an unmasking mechanism for receptor activation. J Biol Chem 277: 35657-35663.

126. Varughese KI, Ruggeri ZM, Celikel R (2004) Platinum-induced space-group transformation in crystals of the platelet glycoprotein Ib alpha N-terminal domain. Acta Crystallogr D Biol Crystallogr 60: 405-411.

127. Huizinga EG, Tsuji S, Romijn MA, Schiphorst ME, de Groot PG, et al. (2002) Structure of the α3 chain of von Willebrand factor 1. Structure 22: 612-620.

128. Zarpellon A, Celikel R, Roberts JR, McClintock RA, Mendicoll-Mobile GL, et al. (2011) Binding of alpha- thrombin to surface-anchored platelet glycoprotein Ib alpha sulfotyrosines through a two-site mechanism involving exosite I. Proc Natl Acad Sci U S A 108: 8626-8633.

129. Dumas JJ, Kumar R, Seeba J, Somers WS, Mosyak L (2003) Crystal structure of the GpIbalpha-thrombin complex essential for platelet aggregation. Science 301: 222-226.

130. Dumas JJ, Kumar R, McDonagh T, Sullivan F, Stahl ML, et al. (2004) Crystal structure of the wild-type von Willebrand factor A1-glycoprotein Ibalpha complex reveals conformation differences with a complex bearing von Willebrand disease mutations. J Biol Chem 279: 23327-23334.

131. Fukuda K, Doggett T, Lauzenzi IJ, Liddington RC, Diacono TG (2005) The snake venom protein botrocetin acts as a biological brace to promote dysfunctional platelet aggregation. Nat Struct Mol Biol 12: 152-159.

132. McEwan PA, Andrews RK, Emsley J (2009) Glycoprotein Ibalpha inhibitor complex structure reveals combined steric and allosteric mechanism of von Willebrand factor antagonism. Blood 114: 4883-4885.

133. Blennner MA, Dong X, Springer TA (2014) Structural basis of regulation of von Willebrand factor binding to glycoprotein Ib. J Biol Chem 289: 5655-5579.

134. Baxter RH, Steinert S, Chelliah Y, Voloshynsky G, Levashina EA, et al. (2010) A heterodimeric complex of the LRR proteins LR RM1 and APL1C regulates complement-like immunity in Anopheles gamierei. Proc Natl Acad Sci U S A 107: 16817-16822.

135. Basquin J, Roudko VV, Rode M, Basquin C, Seraphin B, et al. (2012) Architecture of the nuclease module of the yeast CrC4-not complex: the Not1-Caf1-C4-1 interaction. Mol Cell 48: 207-218.

136. Scott PG, McEwan PA, Dodd CM, Bergmann EM, Bishop PN, et al. (2004) Crystal structure of the dimeric protein core of decorin, the archetypal small leucine-rich repeat proteoglycan. Proc Natl Acad Sci U S A 101: 15633-15638.

137. Scott PG, Dodd CM, Bergmann EM, Sheehan JK, Bishop PN (2006) Crystal structure of the biglycan dimer and insights that BeautySOx is essential for folding and stability of class I small leucine-rich repeat proteoglycans. J Biol Chem 281: 13324-13332.

138. Kelley SL, Lucc T, Nair SK, Tapping RI (2013) The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket. J Immunol 190: 1304-1311.

139. Kim Ji, Lee CJ, Jin MS, Lee CH, Paik SG, et al. (2005) Crystal structure of CD14 and its implications for lipopolysaccharide signaling. J Biol Chem 280: 1347-1351.

140. de Chiara C, Menon RP, Pastore A (2008) Structural bases for recognition of Amp32/LAPN proteins. FEBS J 275: 2548-2560.

141. Huyton T, Wolberger C (2007) The crystal structure of the tumor suppressor protein p32 (Amp32): structural insights into Amp32 family of proteins. Protein Sci 16: 1308-1315.

142. Tochino N, Umehara T, Munemasa Y, Suzuki T, Sato S, et al. (2010) Solution structure of histone chaperone ANP32B: interaction with core histones H3-H4 through its acidic concave domain. J Mol Biol 401: 97-114.

143. Wehrman T, He XL, Raab B, DuPattla A, Blau H, et al. (2007) Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. Nature 53: 25-38.

144. McEwan PA, Yang W, Carr KH, Mo X, Zheng X, et al. (2011) Quaternary organization of GPIb-IX complex and insights into Bernard-Soulier syndrome revealed by the structures of GPIbβ and a GPIbβ/GPIX chimera. Blood 118: 5292-5301.

145. Trastoy B, Lomino JV, Pierce BG, Carter LG, Günther S, et al. (2014) Crystal structure of Streptococcus pyogenes EnoS, an immunomodulatory endoglycosidase specific for human IgG antibodies. Proc Natl Acad Sci U S A 111: 6714-6719.

146. Nagahara Y, Lee, Blundell TL (1999) Analysis and prediction of inter-strand packing distances between beta-sheets of globular proteins. Protein Eng 12: 1055-1062.

147. Evdokimov AG, Anderson DE, Rountzah K, Waugh DS (2001) Unusual molecular architecture of the Yersinia pestis cytotoxin YopM: a leucine-rich repeat protein with the shortest repeating unit. J Mol Biol 312: 807-821.

148. Quezada CM, Hicks SW, Galán JE, Stebbins CE (2009) A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. Proc Natl Acad Sci U S A 106: 4864-4869.

149. Keselj AF, Tang X, McCormick C, Zeqiraj E, Rohde JR, et al. (2014) Structure of an SspH1-PKN1 complex reveals the basis for host substrate recognition and mechanism of activation for a bacterial E3 ubiquitin ligase. Mol Cell Biol 34: 362-373.

150. Zhu Y, Li H, Hu L, Wang J, Zhou Y, et al. (2008) Structure of a Shigella effector reveals a new class of ubiquitin ligases. Nat Struct Mol Biol 15: 1302-1308.

151. Miyashita H, Kretzing R, Kuroki Y, Matsumasa N (2013) Horizontal gene transfer of plant-specific leucine-rich repeats between plants and bacteria. Nat Sci 5: 580-598.

152. Lee JK, Belogrudov GI, Stroud RM (2008) Crystal structure of bovine mitochondrial factor B at 0.96-A resolution. Proc Natl Acad Sci U S A 105: 13378-13394.