Arch Immunol Ther Exp, 2004, 52, 104–112
PL ISSN 0004-069X

WWW.AITE–ONLINE.ORG
Review

Received: 2003.11.19
Accepted: 2004.01.20
Published: 2004.04.15

Leucine-rich repeats in host-pathogen interactions

Łukasz Kędzierski, Jacqui Montgomery, Joan Curtis and Emanuela Handman

The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia

Source of support: the Australian National Health and Medical Research Council and the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases.

Summary

Leucine-rich repeats (LRRs) are versatile binding motifs found in a variety of proteins and are involved in protein-protein interactions. The LRR domain is composed of repeats forming a characteristic solenoid horse-shoe structure, which provides a scaffold for numerous insertions involved in binding to pathogen-associated molecular patterns and surface receptors. LRRs have been shown to be involved in the host defense systems of both plants (resistance genes) and mammals (Toll-like receptors and nucleotide-binding oligomerisation domain proteins), where they sense specific pathogen-associated molecules and activate the innate immune system. Paradoxically, LRRs have also been shown to be part of microbial virulence factors involved in the interaction with host cells and establishment of infection. The potential of LRRs to bind a vast array of structurally unrelated ligands and their well-documented involvement in microbial pathogenesis make them a potential target for vaccines and new drugs. The recent identification of LRRs in the obligate intracellular protozoan parasite *Leishmania* and their participation in the macrophage-parasite interaction have added new insight into the role of LRRs in the host cell invasion.

Key words: **leucine-rich repeat • host-pathogen interaction • Toll-like receptors • *Leishmania***

Full-text PDF: http://www.aite-online/pdf/vol_52/no_2/5239.pdf

Author's address: Dr. Emanuela Handman, Division of Infection and Immunity, The Walter and Eliza Hall Institute of Medical Research, 1G Royal Pde, Parkville 3050, Victoria, Australia, tel.: +61 3 93452476, fax: +61 3 93470852, e-mail: handman@wehi.edu.au

INTRODUCTION

The first array of leucine-rich repeats (LRRs) was identified in the sequence of a leucine-rich α_2-glycoprotein of human serum[73], and since that time the list has expanded to at least 200 proteins[35]. The LRR domains were predicted to form distinct modules that could be incorporated into a variety of proteins, and in fact LRR motifs are very common throughout the phylogenetic kingdom. The motif consists of irregularly spaced but conserved leucine residues, particularly the LxxLxL motif, which forms a β-strand. Additional residues are unique to each particular motif, but these tend to differ among the different proteins. The more closely related the proteins structurally and functionally, the more similar the unique amino acids. In general, the β-sheet structure is highly conserved among different proteins, while the remaining regions display more variation[7, 38]. The motif's β-strand and α-helix units are connected in loops and arranged parallel to a common axis, forming a non-globular, horseshoe structure[7, 35, 38, 40]. The β-sheets cover the concave face of the horseshoe structure, while the helices flank its outer circumference. The entire structure is compact and stabilized by disulfide linkages.

The crystal structure of the porcine ribonuclease inhibitor was the first to reveal the horseshoe shape of the motif and its characteristic arrangement of β--sheets and α-helices[36]. Subsequently, structures of several other LRR-containing proteins and protein complexes featuring motifs of different length were determined[14, 39, 45, 49, 61]. All the structures showed striking similarities despite some differences in overall shape, ranging from a slightly bowed tube to the now classic horseshoe shape[40].

Kajava[35] subdivided the LRR domains into 6 subfamilies based on their length and consensus sequence. These include typical, ribonuclease inhibitor-like, bacterial, plant, cysteine-containing, and SD22-like subfamilies. Proteins in different families tend to share features such as cellular localization, origin, and flanking sequences. Importantly, there seems to be some functional overlap between motifs displaying sequence homology.

All known LRR-containing proteins have been implicated in protein-protein interactions[7, 38]. They participate in a wide range of processes, such as DNA repair, cell adhesion, signal transduction, development, transcription, and RNA processing. There are no obvious features shared between the ligands and receptors of the LRR domains that have been described to date. This indicates that LRRs are a versatile binding motif that can accommodate ligands with minimal sequence identity and make unique and strong interactions with them. Therefore, it is hardly surprising that LRR motifs have been exploited by both hosts and microbes in the "war of the DNAs"[11]. The former can utilize them to recognize pathogen molecules or pathogen-specific molecular patterns, whereas the latter can use LRRs to attach to the host cell and/or gain entry into it. Here we review the involvement of the LRR motifs in host-pathogen interactions, with emphasis on recent studies on the functional and structural characterization of leishmanial proteins containing LRR domains, which are implicated in *Leishmania*-macrophage interaction and establishment of infection.

LRRs IN HOST-PATHOGEN INTERACTIONS: THE HOST'S PERSPECTIVE

LRRs in plant defenses

Plant-pathogen interactions involve the interaction between the product of the pathogen avirulence gene and the host resistance (*R*) gene and result in the hypersensitive response[7, 12]. The absence of one of the genes or lack of a functional protein encoded by one of the genes results in disease. There is a wide range of pathogen-resistance genes, which encode proteins grouped in only 5 different classes. The striking feature of *R* genes is the presence of LRR motifs.

The majority of plant disease-resistance genes identified to date belong to the cytoplasmic nucleotide--binding site (NBS)-LRR class[12, 48]. Apart from the LRR domain, these gene products also contain a NBS, which forms part of a larger domain with homology to interleukin (IL)-1 receptors and may also have coil-coiled domains. Other products of the *R* genes are membrane associated and feature extracellular LRRs. The specificity of the plant disease--resistance genes is confined to the LRR domains, which are either directly or indirectly involved in ligand-binding. NBS-LRR R proteins can also form protein complexes which function as a receptor for many fungal and bacterial ligands[12], but in this case LRRs are not directly involved in the actual ligand recognition. Studies conducted on the plant defense mechanisms and *R* genes are rather sparse, but recently published reports indicate that NBS-LRR genes and their products are ubiquitous in plants such as barley[48], wheat[71] or *Arabidopsis*, where 149 NBS-LRR genes were identified encoding proteins able to detect a vast range of diverse pathogen-derived ligands[54].

Some insight into the structure and function of the plant-specific LRR protein family was recently pro-

vided with the elucidation of the crystal structure of polygalacturonase-inhibiting protein (PGIP), which is involved in plant defense against many phytopathogenic fungi[14]. This molecule has a unique architecture, which consists of two β-sheets most likely to be conserved in other plant LRR-containing proteins. The PGIP also possesses a cluster of negatively charged amino acids on the concave surface of the molecule's structure, which are involved in binding of polygalacturonases released by fungi.

Mammalian LRRs

Recent findings have demonstrated that there are similarities between the pathogen-associated molecular pattern (PAMP) receptors in animals and recognition complexes involved in plant defenses, both of which comprise LRRs and probably share a common evolutionary origin[59]. PAMPs are recognized by pattern recognition receptors (PRRs), which distinguish between self structures and microbial structures and transduce this recognition signal to activate host innate immune responses. The recognition of PAMPs is carried out by a variety of molecules belonging to the Toll-like receptor (TLR) family[52, 68, 74], nucleotide-binding oligomerisation domain (NOD) protein subfamily[31, 33], and NALP (for NACHT (for NAIP – neuronal apoptosis inhibitory protein, CIITA – MHC class II transcription activator, HET-E – incompatibility locus protein from *Podospora anserine* and TP1 – telomerase-associated protein), LRR and pyrin domain (PYD) domains) subfamily[9, 75]. All of these families of molecules form part of the innate immune system, the most ancient and universal system of defense. In many organisms, such as plants and insects, it is the sole means of defense against microbial attack since, unlike vertebrates, these organisms lack an adaptive immune system. The innate immune system consists of many soluble and cell-associated receptors as well as effector cells involved in host-pathogen interaction[1].

The TLRs are an ancient group of molecules first identified in *Drosophila*, where they are involved in antifungal[42] and antibacterial[23] responses and form a part of the innate immune system. The TLRs form a group of pattern-recognition receptors essential for the induction of the rapid response characteristic of innate immunity. They are expressed by a variety of cells, such as monocytes/macrophages, different subsets of dendritic cells and mast cells, as well as other cells contributing to inflammatory responses[74]. Their most striking characteristic is the presence of an extracellular LRR domain and an intracellular Toll/IL-1 receptor (TIR) domain. The LRRs act as ligand-recognition domains and they also play a role in signal transduction[52]. The ability of TLRs to sense a large variety of molecules, such as bacterial lipoprotein and lipopolysaccharide, peptidoglycan, CpG oligodeoxynucleotides, flagellin, as well as additional fungal and yeast PAMPs[68, 77], makes them a critical arm of the host defense system. Recently, TLRs have been shown to also recognize glycolipid molecules in protozoan parasites such as *Trypanosoma cruzi*[8] and *Leishmania*[13, 58].

While the TLRs act as membrane PRRs, there are also cytosolic proteins executing a similar function. These have been grouped into the NOD family of proteins. To date, only two members of the NOD family have been identified, namely NOD1 and NOD2. Both are structurally related to the NBS--LRR class of plant R proteins. They are composed of a caspase-recruitment domain (CARD), the NOD domain, and a stretch of LRRs at the carboxy terminus of the protein[31]. The LRRs present in NODs have been shown to be involved in the PAMP-recognition process, and act as cytosolic pattern-recognition receptors recognizing the same conserved structures as TLRs[28, 32, 60]. An intriguing role for these molecules in the etiology of Crohn's disease, a chronic inflammatory bowel disease, is currently the focus of much interest[28, 60].

Another family of cytosolic proteins structurally related to plant R proteins and NODs is the recently discovered NALPS subfamily. These are NBS-LRR proteins which contain a PYD domain instead of the CARD domain found in NODs[9]. There are 14 NALP protein-encoding genes in the human genome, all of which are more or less identical structurally, and feature an LRR-rich C-terminal region[75]. This family of proteins is involved in bacteria-induced inflammation and the formation of the caspase-NALP complex known as the inflammasome. The mechanisms that NALPs employ to form the inflammasome remain to be elucidated. However, it has been postulated that NALPs are activated through the detection of microbial patterns by their LRRs [9].

LRRs IN HOST-PATHOGEN INTERACTIONS: THE MICROBES' PERSPECTIVE

As described above, the LRR domains present in plant and mammalian proteins are major players in the host defense against pathogens. Their role in the recognition of bacterial PAMPs has been well documented (reviewed in[74]), and there is mounting evidence that they are also involved in innate immunity to many parasites[8, 51]. Paradoxically, many pathogens, such as bacteria and protozoan parasites, appear to utilize LRRs for the exactly opposite purpose, that is

to invade the host cell and/or suppress the host defenses.

Bacterial virulence factors-containing LRR domains

Internalins of *Listeria monocytogenes* (*Li. monocytogenes*) are the most extensively studied bacterial proteins containing LRR domains. There are at least nine proteins in this family; all contain 22-residue LRRs (reviewed in[49]) and have been implicated in the invasion of mammalian host cells. The exact function of all of them has not been elucidated to date, but internalin A (InlA) and InlB have been studied extensively. InlA mediates entry into intestinal epithelial cells, while InlB promotes invasion of hepatocytes, fibroblasts and endothelial cells[46]. These ligands utilize the host receptors E-cadherin and gC1q-R/p32, respectively[4, 67]. It has been shown that InlB is necessary and sufficient to promote entry into mammalian cells[6, 62]. Moreover, the LRR domain is the region of InlB responsible for entry and initiation of signal transduction[5, 50], and antibodies directed to the LRR region block cell invasion by *Li. monocytogenes*[53].

Other bacterial virulence factors possessing an LRR domain include *Yersinia pestis* (*Y. pestis*) YopM protein[44], *Salmonella typhimurium* SspH1 and SspH2 secreted proteins[55] and SlrP protein[76], and *Shigella flexneri* IpaH protein[21]. YopM is the primary virulence factor of *Y. pestis* and displays thrombin-binding activity, although that seems to be a function of a small pool of extracellular YopM[25, 43]. Proteins expressed by the enteric bacteria *Salmonella* and *Shigella* have been shown to be involved in the inhibition of nuclear factor (NF)-κB-dependent gene expression and as such contribute to bacterial persistence and colonization of host tissues[20].

LRR-CONTAINING PROTEINS OF PROTOZOAN PARASITES

Recent studies have described LRR-containing proteins in protozoan parasites, but these have not been as well characterized as those of pathogenic bacteria. However, it is likely that protozoan proteins with LRR motifs have a role in virulence in these parasites, as they do in bacteria.

The waterborne protozoan parasite *Giardia lamblia* possesses three cyst wall proteins, each containing LRR domains[47, 57, 72]. These proteins participate in the formation of a protein complex which is targeted to the encystation-specific secretory granules before incorporation into the cyst wall. The LRRs mediate the association of these proteins to the cell wall and the formation of its fibrillar structure[24]. An LRR-containing protein has also been identified in *Trypanosoma brucei* (*T. brucei*). In these organisms, antigenic variation of the surface coat is controlled by specific expression sites where expression site-associated genes (ESAG) have also been localized[63, 65, 70]. The ESAG8 protein is almost exclusively composed of LRRs, but also contains a putative DNA-binding zinc finger[65]. The cytoplasmic pool of ESAG8 has been shown to interact with a *Pumilio* family protein, PUF1 (Pumilio/FBF (Fem-3 mRNA binding factor))[27], which is essential for cell viability and is an RNA-regulatory protein. Recently, a novel protein containing 12 LRRs has been identified in *T. vaginalis*[26]. The protein, named TvBspA-like-625, is thought to be involved in interactions between *T. vaginalis* and the vaginal epithelial cells as well as the host extracellular matrix proteins, both of which are required for the parasite to establish a vaginal infection.

LEISHMANIAL PROTEINS-CONTAINING LRRS

LRR-containing proteins such as proteophosphoglycan (PPG)[30] and parasite surface antigen (PSA)-2[34] have recently been identified in the extracellular glycocalyx of *Leishmania*. *Leishmania* are protozoan parasites that shuttle between blood-feeding sandflies, where they grow as flagellated promastigotes, and mammalian hosts, where they exist as obligate intracellular amastigotes in phagocytes. The parasites cause a disease known as leishmaniasis, with clinical manifestation ranging from self-healing skin lesions to severe systemic disease and even death. A critical point in the host-parasite interaction involves the attachment to and invasion of host macrophages, initially by the promastigotes and subsequently by amastigotes. Both promastigotes and amastigotes use receptor-mediated phagocytosis for invasion[17, 19], and a number of host receptors and leishmanial molecules are involved in this complex interaction[19, 66].

On the parasite side, membrane glycoconjugates, such as the 63 kDa protease leishmanolysin (gp63), lipophosphoglycan (LPG) and PPG, have been implicated in the attachment and invasion of host macrophages[17, 18]. LPG and PPG are expressed extracellularly and are ideally placed to interact with receptors on host macrophages. It has been previously shown that secreted PPG can bind to macrophages and modulate their biology[64]. It was assumed that this interaction was mediated by glycans shared by PPG and LPG. However, when the gene encoding a membrane-bound PPG was cloned[30], sequence analysis revealed the presence of an LRR motif with remarkable similarity to the LRRs previously described in

members of the PSA-2 family of glycoproteins[34, 56]. There are some striking similarities between PPG and PSA-2 and the listerial internalin family. For example, both PPG and PSA-2 are encoded by gene families and are expressed as secreted and membrane-anchored forms. Given the proven role of internalins in host cell invasion, it is likely that PPG and PSA-2 are also involved in similar processes. Recent studies in our laboratory have demonstrated that PSA-2 binds to macrophages via the complement receptor 3 (CR3) and that the monoclonal antibodies directed against the LRR region of PSA-2 inhibit macrophage invasion (Kędzierski et al., manuscript submitted). An analogous type of interaction between LRR and CR3 has been described for the platelet glycoprotein Ibα, which binds to leukocyte CR3[69]. Binding has been mapped to the glycoprotein Ibα LRR domain and the I (inserted) domain present on the CD11b subunit of CR3. The I domain has been shown to be a major ligand-binding domain of CR3[15].

Analysis of the predicted PSA-2 and PPG protein sequences revealed that the similarity extends beyond the LRR and includes a di-cysteine motif located N-terminal to the LRRs. Cysteine clusters often flank LRR domains, particularly at the C-terminal end of the LRR motif, although N-terminal motifs have been identified as well[35, 37]. A cysteine-rich motif, designated as NF domain, for N-terminal cysteine-containing flanking domain, has been described by Kajava[35]. Our recent analysis of the LRR domains of the PPG and PSA-2 from *Leishmania major, L. amazonensis, L. chagasi, L. infantum* and *L. tropica* identified a novel motif containing two cysteines which was located N-terminal to the LRR domain. This motif is not restricted to *Leishmania,* but is also found in the same location in other LRR-containing plant and parasite proteins of *Arabidopsis thaliana*, *Aegilops tauschii* and *Ozyra sativa,* and the cyst wall proteins of *Giardia* (Fig. 1A). Interestingly, in PSA-2 polypeptides from different *Leishmania* species the di-cysteine motif is located at a similar distance of 43-44 amino acids N-terminal to the LRR domain. In PPG, the LRRs are present in a main block as well as scattered single motifs. In PPG the di-cysteine motif is located upstream of the single LRR domain between amino acids 153-176. This places the motif at an equivalent distance from the first LRR repeat as in the PSA-2 sequences. Although conserved, the distance between the di-cysteine motif and the LRR is much shorter in plant and *Giardia* proteins and varies between 26 to 28 residues from the first LRR motif (Fig. 1A).

The consensus sequence of the di-cysteine motif (CxWxxxxC) contains highly conserved cysteine residues as well as a single tryptophan residue (Fig. 1B). The cysteine residues most likely form a disulfide bridge, while the single tryptophan residue may be important for the stability of the entire structure, a function that has been previously demonstrated for tryptophan residues[16]. Such conservation in both the sequence and location of the motif in several diverse proteins suggests a common functional and/or structural role. There is evidence that LRR domains can fold as independent protein modules and therefore can be easily linked to other domains without steric hindrance[38]. Flanking cysteine clusters in several LRR-containing proteins, including the TLRs[22] and small proteoglycans[41], have been shown to form intrachain disulfide linkages that are thought to define and separate the LRRs from other domains in the protein[3]. It is likely that the di-cysteine motif plays a structural role and is involved in exposure of the LRR, thus facilitating contact between the LRR domain and its interacting partner molecules.

In the case of *Leishmania*, the additional motif may influence folding of the LRR domain or may promote dimerization, contributing directly to parasite attachment and host cell invasion. A similar observation was made with the LRRs of InlB of *Li. monocytogenes*, which was more effective at promoting bacterial internalization when linked to an adjacent "inter-repeat" sequence[6]. It has been also demonstrated that the N-terminal cysteine residues of glycoprotein Ibα form a β-hairpin with disulfide bonds linking the β-strands of LRR, thus stabilizing the LRR structure[29].

The LRR-containing proteins of *Leishmania* possess all the attributes found in LRR-containing bacterial virulence factors. Their striking similarity to the *Li. monocytogenes* internalin family strongly suggests that LRRs in PPG and PSA-2 are involved in the macrophage invasion process and contribute to the overall pathogenesis of *Leishmania*.

CONCLUSIONS

Discovery of a LRR domain allowed the identification of a large number of proteins containing such domains. Since LRRs are implicated in protein-protein interaction, it is not surprising that they are involved in host-pathogen interactions and are exploited by both host and pathogens. It has been postulated that LRRs form a versatile scaffold that can accommodate a variety of functional insertions, thus allowing recognition of a vast array of ligands[2]. Their presence in a diverse range of host microbial-recognition proteins (TLRs, NODs) implies evolutionary conservation of the innate immune defenses.

A

B

```
L.major PPG                          94 LSKTWVCRNF--------CWDGVRCSAE-GVEVRLDSRQVVGRLPEVPTAVDAEQVMVTHFAVEGEGDGIKGGL 160 AJ243459
L.major PSA-2                        60 LGEKWAGNDF--------CWEAVLCNAP-DVYVSGISPTYAGTLPEMPVNVDYRHVVIKQLDFSEMGPG----L 122 L43630
L.major TR1                          59 LGEKWAGNDF--------CWEAVLCNAP-DVYVSGISPTYAGTLPEMPENVDYRHVVIRRLDFSEMGPG----L 121 X57009
L.major clone 2.1                    60 LGEKWAGNDF--------CWEAVLCNAP-DVYVSGISPTYAGTLPEMPVNVDYRHVVIKQLDFSEMGPG----L 122 L43630
L.major clone 2.5                    60 LGEKWAGNDF--------CWDFIRCTQR-AVYVRVISTTYAGTLPEMPVNVDYRHVMIRELYIWNMP------L 120 X57134
L.major clone 4.6                    60 LEKKWAGNDF--------CWDLIRCNKM-GVYVTGISPTYAGTLPEMPVNVDYRHVMIRDLWLWNMS------L 120 X57135
L.major clone 6.4                    60 LGEKWTGNDF--------CWEAVLCDAP-DVYVSGISPTYAGTLPEMPVNVDYTTVMIKQLDFSKMGPG----L 122 X56810
L.major seq01                        60 LEKKWTGNDF--------CWDHVACVSV-NVFVSLDVSTYAGTLPEMPVGVDYRHVMIRDLGFWNMP------L 120 AL390114
L.major seq03                        60 LGEKWAGNDF--------CWDFIVCNVI-GVNVRGISPTYAGTLPEMPVNVDYRHVVIKQLDFSEMGPG----L 122 AL390114
L.major seq07                        60 LGEVWKGDDF--------CWDHLLCLPL-SVTIRVFVFTYAGTLPEMPVNVDYGHVMIGNLGLWNMPQVVGT-L 125 AL390114
L.infantum PSA-2                     60 LGKNWTGDDF--------CWDHVTCAPL-SVAVGLDVSTYAGTLPEIPVGVDYRYVMIRDLGFWNMP------L 120 Y10816
L.tropica PSA-2                      60 LGEKWTGDDF--------CWDHVTCVSL-SVFVRLDVSTYAGTLPEMPVGVDYSHVMIRDLGFWNMP------L 120 AF164027
L.amazonensis PSA-2                  60 LGDTWTGSDF--------CWDHILCYSS-GVGVWMHNVDYTGTLPEMPASVDYKDVMILALDFGAMGQG----L 122 M38368
L.chagasi PSA-2                      63 L---WKGDDF--------CWKYVYCYPS-AVGVLMDSVDYAGSLPEMPAGVDYTNVVITQLDFSAKGQR----L 122 AF006588
G.lamblia CWP1                       43 L-----SGDS-------ICWWGVTCEAS-NNYVIALDLSDMG-LTGTIPENIGC----------------L 85  U09330
G.lamblia CWP2                       43 L----TPDVS-------YCWWGITCDS-NNNVIGIDLSDMG--LTGALPADIGCFP--------------L 87  U28965
G.lamblia CWP3                       40 L----QSDVY--------CWWGVSCDDN-DNVVSLNVQHMG--LNGQLSDLTN----------------L 80  AY061927
A.thaliana LRR-like                  47 L-RQWNSDNIN------YCWWGVTCDNTGLFRVIALNLTGLG-LTGSISPWFGRFDN------------L 98  AL161552
A.thaliana protein kinase-like       47 L-TSWNLSTTF-------CWWGVTCDVS-LRHVTSLDLSGLN-LSGTLSSDVAH----------------L 93  AL132965
O.sativa hypothetical protein        49 L-ASWNSSTSF-------CWWGVTCDRRTPARVAALTLPSGN-LAGGLPPVIGN----------------L 96  AP002522
Ae.tauschii LRR-like                 54 L-SSWTINSSSNGSTHGFCWWGVRCSRTHPGHVMALRLQGIG-LSGTISPFLGN----------------L 108 AF497474
```

Figure 1. Leucine-rich repeats (LRRs) in protozoa and plants. **A** – schematic representation of the relative positions of LRR motifs and the associated conserved, C-terminal-flanking di-cysteine sequence in *Leishmania* species, *Gardia lamblia* (*G. lamblia*) cyst wall proteins, and plant species *Arabidopsis thaliana* (*A. thaliana*), *Aegilops tauschii* (*Ae. tauschii*) and *Ozyra sativa* (*O. sativa*). **B** – sequence alignment of 21 sequences from protozoan and plant species containing the CxWxxxxC di-cysteine motif. The alignment was generated using Clustal W with minimal manual editing. Clones 2.1, 2.5, 4.6, and 6.4 and seq 01, 03 and 07 belong to the *Leishmania major* (*L. major*) PSA-2 family of genes and represent separate genes. New sequences of seq 01, 03 and 07 were identified by NCBI BLAST searches of the *L. major* genome on a chromosome fragment designated LMFLCHR12 (*L. major* chromosome 12/24 clone Chr.12/24 strain Friedlin), which also contains the previously identified clones 2.1, 2.5, 4.6, and 6.4. GenBank accession numbers are given on the right of each sequence.

The recent discovery of a scavenger receptor-like protein in *Plasmodium falciparum* (*P. falciparum*)[10] which comprises several domains associated with PRRs raises the intriguing possibility of convergent evolution of defense mechanisms in protozoan parasites. The presence of LRR domains in these molecules remains to be determined, but several contigs containing LRRs have been already identified in the recently completed *P. falciparum* genome [51].

ACKNOWLEDGMENT

Łukasz Kędzierski is a recipient of an NHMRC Peter Doherty Postdoctoral Fellowship.

REFERENCES

1. Basset C., Holton J., O'Mahony R. and Roitt I. (2003): Innate immunity and pathogen-host interaction. Vaccine, 21 (suppl. 2), S12–23.
2. Bell J. K., Mullen G. E., Leifer C. A., Mazzoni A., Davies D. R. and Segal D. M. (2003): Leucine-rich repeats and pathogen recognition in Toll-like receptors. Trends Immunol., 24, 528–533.
3. Bhowmick N., Huang J., Puett D., Isaacs N. W. and Lapthorn A. J. (1996): Determination of residues important in hormone binding to the extracellular domain of the luteinizing hormone/chorionic gonadotropin receptor by site-directed mutagenesis and modeling. Mol. Endocrinol., 10, 1147–1159.
4. Braun L., Ghebrehiwet B. and Cossart P. (2000): gC1q-R/p32, a C1q-binding protein, is a receptor for the InlB invasion protein of *Listeria monocytogenes*. EMBO J., 19, 1458–1466.
5. Braun L., Nato F., Payrastre B., Mazie J. C. and Cossart P. (1999): The 213-amino-acid leucine-rich repeat region of the *Listeria monocytogenes* InlB protein is sufficient for entry into mammalian cells, stimulation of PI 3-kinase and membrane ruffling. Mol. Microbiol., 34, 10–23.
6. Braun L., Ohayon H. and Cossart P. (1998): The InIB protein of *Listeria monocytogenes* is sufficient to promote entry into mammalian cells. Mol. Microbiol., 27, 1077–1087.
7. Buchanan S. G. and Gay N. J. (1996): Structural and functional diversity in the leucine-rich repeat family of proteins. Prog. Biophys. Mol. Biol., 65, 1–44.
8. Campos M. A., Almeida I. C., Takeuchi O., Akira S., Valente E. P., Procopio D. O., Travassos L. R., Smith J. A., Golenbock D. T. and Gazzinelli R. T. (2001): Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. J. Immunol., 167, 416–423.
9. Chamaillard M., Girardin S. E., Viala J. and Philpott D. J. (2003): NODs, NALPs and NAIP: intracellular regulators of bacterial--induced inflammation. Cell. Microbiol., 5, 581–592.
10. Claudianos C., Dessens J. T., Trueman H. E., Arai M., Mendoza J., Butcher G. A., Crompton T. and Sinden R. E. (2002): A malaria scavenger receptor-like protein essential for parasite development. Mol. Microbiol., 45, 1473–1484.
11. Cohn M. (2002): The immune system: a weapon of mass destruction invented by evolution to even the odds during the war of the DNAs. Immunol. Rev., 185, 24–38.
12. Dangl J. L. and Jones J. D. (2001): Plant pathogens and integrated defence responses to infection. Nature, 411, 826–833.
13. De Veer M. J., Curtis J. M., Baldwin T. M., Didonato J. A., Sexton A., Mcconville M. J., Handman E. and Schofield L. (2003): MyD88 is essential for clearance of *Leishmania major*: possible role for lipophosphoglycan and Toll-like receptor 2 signaling. Eur. J. Immunol., 33, 2822–2831.
14. Di Matteo A., Federici L., Mattei B., Salvi G., Johnson K. A., Savino C., de Lorenzo G., Tsernoglou D. and Cervone F. (2003): The crystal structure of polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. Proc. Natl. Acad. Sci. USA, 100, 10124–10128.
15. Diamond M. S., Garcia-Aguilar J., Bickford J. K., Corbi A. L. and Springer T. A. (1993): The I domain is a major recognition site on the leukocyte integrin Mac-1 (CD11b/CD18) for four distinct adhesion ligands. J. Cell Biol., 120, 1031–1043.
16. Greene L. H., Chrysina E. D., Irons L. I., Papageorgiou A. C., Acharya K. R. and Brew K. (2001): Role of conserved residues in structure and stability: tryptophans of human serum retinol-binding protein, a model for the lipocalin superfamily. Protein Sci., 10, 2301–2316.
17. Handman E. (1999): Cell biology of *Leishmania*. Adv. Parasitol., **44,** 1–39.
18. Handman E. (1992): Host-parasite interactions in leishmaniasis. Adv. Mol. Cell Biol., 5, 133–155.
19. Handman E. and Bullen D. V. (2002): Interaction of *Leishmania* with the host macrophage. Trends Parasitol., 18, 332–334.
20. Haraga A. and Miller S. I. (2003): A *Salmonella enterica* serovar typhimurium translocated leucine-rich repeat effector protein inhibits NF-κB-dependent gene expression. Infect. Immun., 71, 4052–4058.
21. Hartman A. B., Venkatesan M., Oaks E. V. and Buysse J. M. (1990): Sequence and molecular characterization of a multicopy invasion plasmid antigen gene, *ipaH*, of *Shigella flexneri*. J. Bacteriol., 172, 1905–1915.
22. Hashimoto C., Gerttula S. and Anderson K. V. (1991): Plasma membrane localization of the Toll protein in the syncytial *Drosophila* embryo: importance of transmembrane signaling for dorsal-ventral pattern formation. Development, 111, 1021–1028.
23. Hedengren M., Asling B., Dushay M. S., Ando I., Ekengren S., Wihlborg M. and Hultmark D. (1999): Relish, a central factor in the control of humoral but not cellular immunity in *Drosophila*. Mol. Cell, 4, 827–837.
24. Hehl A. B., Marti M. and Kohler P. (2000): Stage-specific expression and targeting of cyst wall protein-green fluorescent protein chimeras in *Giardia*. Mol. Biol. Cell., 11, 1789–1800.
25. Hines J., Skrzypek E., Kajava A. V. and Straley S. C. (2001): Structure-function analysis of *Yersinia pestis* YopM's interaction with α-thrombin to rule on its significance in systemic plague and to model YopM's mechanism of binding host proteins. Microb. Pathog., 30, 193–209.
26. Hirt R. P., Harriman N., Kajava A. V. and Embley T. M. (2002): A novel potential surface protein in *Trichomonas vaginalis* contains a leucine-rich repeat shared by micro-organisms from all three domains of life. Mol. Biochem. Parasitol., 125, 195–199.
27. Hoek M., Zanders T. and Cross G. A. (2002): *Trypanosoma brucei* expression-site-associated-gene-8 protein interacts with a *Pumilio* family protein. Mol. Biochem. Parasitol., 120, 269–283.
28. Hugot J. P., Chamaillard M., Zouali H., Lesage S., Cezard J. P., Belaiche J., Almer S., Tysk C., O'Morain C. A., Gassull M., Binder V., Finkel Y., Cortot A., Modigliani R., Laurent-Puig P., Gower-Rousseau C., Macry J., Colombel J. F., Sahbatou M. and Thomas G. (2001): Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature, 411, 599–603.
29. Huizinga E. G., Tsuji S., Romijn R. A., Schiphorst M. E., de Groot P. G., Sixma J. J. and Gros P. (2002): Structures of glycoprotein Ibα and its complex with von Willebrand factor A1 domain. Science, 297, 1176–1179.
30. Ilg T., Montgomery J., Stierhof Y. D. and Handman E. (1999): Molecular cloning and characterization of a novel repeat-containing *Leishmania major* gene, *ppg1*, that encodes a membrane-associated form of proteophosphoglycan with a putative glycosylphosphatidylinositol anchor. J. Biol. Chem., 274, 31410–31420.
31. Inohara N. and Nunez G. (2003): NODs: intracellular proteins

involved in inflammation and apoptosis. Nat. Rev. Immunol., 3, 371–382.

32. Inohara N., Ogura Y., Chen F. F., Muto A. and Nunez G. (2001): Human NOD1 confers responsiveness to bacterial lipopolysaccharides. J. Biol. Chem., 276, 2551–2554.

33. Inohara N., Ogura Y. and Nunez G. (2002): NODs: a family of cytosolic proteins that regulate the host response to pathogens. Curr. Opin. Microbiol., 5, 76–80.

34. Jimenez-Ruiz A., Boceta C., Bonay P., Requena J. M. and Alonso C. (1998): Cloning, sequencing, and expression of the PSA genes from *Leishmania infantum*. Eur. J. Biochem., 251, 389–397.

35. Kajava A. V. (1998): Structural diversity of leucine-rich repeat proteins. J. Mol. Biol., 277, 519–527.

36. Kobe B. and Deisenhofer J. (1993): Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. Nature, 366, 751–756.

37. Kobe B. and Deisenhofer J. (1994): The leucine-rich repeat: a versatile binding motif. Trends Biochem. Sci., 19, 415-421.

38. Kobe B. and Deisenhofer J. (1995): Proteins with leucine-rich repeats. Curr. Opin. Struct. Biol., 5, 409–416.

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

40. Kobe B. and Kajava A. V. (2001): The leucine-rich repeat as a protein recognition motif. Curr. Opin. Struct. Biol., 11, 725–732.

41. Kresse H., Hausser H. and Schonherr E. (1993): Small proteoglycans. Experientia, 49, 403–416.

42. Lemaitre B., Nicolas E., Michaut L., Reichhart J. M. and Hoffmann J. A. (1996): The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell, 86, 973–983.

43. Leung K. Y., Reisner B. S. and Straley S. C. (1990): YopM inhibits platelet aggregation and is necessary for virulence of *Yersinia pestis* in mice. Infect. Immun., 58, 3262–3271.

44. Leung K. Y. and Straley S. C. (1989): The *yopM* gene of *Yersinia pestis* encodes a released protein having homology with the human platelet surface protein GPIbα. J. Bacteriol., 171, 4623–4632.

45. Liker E., Fernandez E., Izaurralde E. and Conti E. (2000): The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain. EMBO J., 19, 5587–5598.

46. Lingnau A., Domann E., Hudel M., Bock M., Nichterlein T., Wehland J. and Chakraborty T. (1995): Expression of the *Listeria monocytogenes* EGD inlA and inlB genes, whose products mediate bacterial entry into tissue culture cell lines, by PrfA-dependent and -independent mechanisms. Infect. Immun., 63, 3896–3903.

47. Lujan H. D., Mowatt M. R., Conrad J. T., Bowers B. and Nash T. E. (1995): Identification of a novel *Giardia lamblia* cyst wall protein with leucine-rich repeats. Implications for secretory granule formation and protein assembly into the cyst wall. J. Biol. Chem., 270, 29307–29313.

48. Madsen L. H., Collins N. C., Rakwalska M., Backes G., Sandal N., Krusell L., Jensen J., Waterman E. H., Jahoor A., Ayliffe M., Pryor A. J., Langridge P., Schulze-Lefert P. and Stougaard J. (2003): Barley disease resistance gene analogs of the NBS-LRR class: identification and mapping. Mol. Genet. Genomics, 269, 150–161.

49. Marino M., Braun L., P. Cossart P. and Ghosh P. (2000): A framework for interpreting the leucine-rich repeats of the *Listeria* internalins. Proc. Natl. Acad. Sci. USA, 97, 8784–8788.

50. Marino M., Braun L., P. Cossart P. and Ghosh P. (1999): Structure of the lnlB leucine-rich repeats, a domain that triggers host cell invasion by the bacterial pathogen *L. monocytogenes*. Mol. Cell, 4, 1063–1072.

51. McGuinness D. H., Dehal P. K. and Pleass R. J. (2003): Pattern recognition molecules and innate immunity to parasites. Trends Parasitol., 19, 312–319.

52. Medzhitov R. (2001): Toll-like receptors and innate immunity. Nat. Rev. Immunol., 1, 135–145.

53. Mengaud J., Lecuit M., Lebrun M., Nato F., Mazie J. C. and Cossart P. (1996): Antibodies to the leucine-rich repeat region of internalin block entry of *Listeria monocytogenes* into cells expressing E-cadherin. Infect. Immun., 64, 5430–5433.

54. Meyers B. C., Kozik A., Griego A., Kuang H. and Michelmore R. W. (2003): Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell, 15, 809–834.

55. Miao E. A., Scherer C. A., Tsolis R. M., Kingsley R. A., Adams L. G., Baumler A. J. and Miller S. I. (1999): *Salmonella typhimurium* leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. Mol. Microbiol., 34, 850–864.

56. Montgomery J., Ilg T., Thompson J. K., Kobe B. and Handman E. (2000): Identification and predicted structure of a leucine-rich repeat motif shared by *Leishmania major* proteophosphoglycan and parasite surface antigen 2. Mol. Biochem. Parasitol., 107, 289–295.

57. Mowatt M. R., Lujan H. D., Cotten D. B., Bowers B., Yee J., Nash T. E. and Stibbs H. H. (1995): Developmentally regulated expression of a *Giardia lamblia* cyst wall protein gene. Mol. Microbiol., 15, 955–963.

58. Muraille E., de Trez C., Brait M., de Baetselier P., Leo O. and Carlier Y. (2003): Genetically resistant mice lacking MyD88--adapter protein display a high susceptibility to *Leishmania major* infection associated with a polarized Th2 response. J. Immunol., 170, 4237–4241.

59. Nurnberger T. and Brunner F. (2002): Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. Curr. Opin. Plant Biol., 5, 318–324.

60. Ogura Y., Bonen D. K., Inohara N., Nicolae D. L., Chen F. F., Ramos R., Britton H., Moran T., Karaliuskas R., Duerr R. H., Achkar J. P., Brant S. R., Bayless T. M., Kirschner B. S., Hanauer S. B., Nunez G. and Cho J. H. (2001): A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature, 411, 603–606.

61. Papageorgiou A. C., Shapiro R. and Acharya K. R. (1997): Molecular recognition of human angiogenin by placental ribonuclease inhibitor – an X-ray crystallographic study at 2.0 A resolution. EMBO J., 16, 5162–5177.

62. Parida S. K., Domann E., Rohde M., Muller S., Darji A., Hain T., Wehland J. and Chakraborty T. (1998): Internalin B is essential for adhesion and mediates the invasion of *Listeria monocytogenes* into human endothelial cells. Mol. Microbiol., 28, 81–93.

63. Pays E., Tebabi P., Pays A., Coquelet H., Revelard P., Salmon D. and Steinert M. (1989): The genes and transcripts of an antigen gene expression site from *T. brucei*. Cell, 57, 835–845.

64. Piani A., Ilg T., Elefanty A. G., Curtis J. and Handman E. (1999): *Leishmania major* proteophosphoglycan is expressed by amastigotes and has an immunomodulatory effect on macrophage function. Microbes Infect., 1, 589–599.

65. Revelard P., Lips S. and Pays E. (1990): A gene from the VSG expression site of *Trypanosoma brucei* encodes a protein with both leucine-rich repeats and a putative zinc finger. Nucleic Acids Res., 18, 7299–7303.

66. Rittig M. G. and Bogdan C. (2000): *Leishmania*-host-cell interaction: complexities and alternative views. Parasitol. Today, 16, 292–297.

67. Schubert W. D., Urbanke C., Ziehm T., Beier V., Machner M. P., Domann E., Wehland J., Chakraborty T. and Heinz D. W. (2002): Structure of internalin, a major invasion protein of *Listeria monocytogenes*, in complex with its human receptor E-cadherin. Cell, 111, 825–836.

68. Sieling P. A. and Modlin R. L. (2002): Toll-like receptors: mammalian "taste receptors" for a smorgasbord of microbial invaders. Curr. Opin. Microbiol., 5, 70–75.

69. Simon D. I., Chen Z., Xu H., Li C. Q., Dong J., Mcintire L. V., Ballantyne C. M., Zhang L., Furman M. I., Berndt M. C. and Lopez J. A. (2000): Platelet glycoprotein ib α is a counterreceptor for the leukocyte integrin Mac-1 (CD11b/CD18). J. Exp. Med., 192, 193–204.

70. Smiley B. L., Stadnyk A. W., Myler P. J. and Stuart K. (1990): The trypanosome leucine repeat gene in the variant surface glycopro-

tein expression site encodes a putative metal-binding domain and a region resembling protein-binding domains of yeast, *Drosophila*, and mammalian proteins. Mol. Cell Biol., 10, 6436–6444.

71. Spielmeyer W. and Lagudah S. (2003): Homologous set of NBS--LRR genes located at leaf and stripe rust resistance loci on short arms of chromosome 1 of wheat. Funct. Integr. Genomics, 3, 86–90.

72. Sun C. H., McCaffery J. M., Reiner D. S. and Gillin F. D. (2003): Mining the *Giardia lamblia* genome for new cyst wall proteins. J. Biol. Chem., 278, 21701–21708.

73. Takahashi N., Takahashi Y. and Putnam F. W. (1985): Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich α 2-glycoprotein of human serum. Proc. Natl. Acad. Sci. USA, 82, 1906–1910.

74. Takeda K., Kaisho T. and Akira S. (2003): Toll-like receptors. Annu. Rev. Immunol., 21, 335–376.

75. Tschopp J., Martinon F. and Burns K. (2003): NALPs: a novel protein family involved in inflammation. Nat. Rev. Mol. Cell Biol., 4, 95–104.

76. Tsolis R. M., Townsend S. M., Miao E. A., Miller S. I., Ficht T. A., Adams L. G. and Baumler A. J. (1999): Identification of a putative *Salmonella enterica* serotype typhimurium host range factor with homology to IpaH and YopM by signature-tagged mutagenesis. Infect. Immun., 67, 6385–6393.

77. Underhill D. M. and Ozinsky A. (2002): Toll-like receptors: key mediators of microbe detection. Curr. Opin. Immunol., 14, 103–110