Pattern-Recognition Receptors in Human Disease


Building an immune system from nine domains

E.M. Pålsson-McDermott and L.A.J. O’Neill
School of Biochemistry and Immunology, Trinity College, Dublin 2, Ireland

Abstract

Four families of PRRs (pattern-recognition receptors) have been identified as important components of innate immunity, participating in the sensory system for host defence against the invasion of infectious agents. The TLRs (Toll-like receptors) recognize a variety of conserved microbial PAMPs (pathogen-associated molecular patterns) derived from bacteria, viruses, protozoa and fungi. They work in synergy with the cytosolic NLRs [NOD (nucleotide binding and oligomerization domain)-like receptors] (which sense bacteria), RLRs [RIG-I (retinoic acid-inducible gene 1)-like receptors] (which sense viruses) and CLRs (C-type lectin receptors) (which sense fungi). All of these receptor families signal an increase in the expression of a range of immune and inflammatory genes. The structural architecture of these receptors is conserved, involving seven distinct domains: the LRR (leucine-rich repeat) domain, the TIR (Toll/interleukin-1 receptor) domain, the NBS (nucleotide-binding site), the CARD (caspase recruitment domain), the PYD (pyrin domain), the helicase domain and the CTLD (C-type lectin domain). Two other domains, the Ig domain and the ITAM (immunoreceptor tyrosine-based activation motif) domain also participate and are also found in antibodies and TCRs (T-cell receptors), key proteins in adaptive immunity. This total of nine domains can therefore be used to construct immune systems which are common to many, if not all, species, allowing us to speculate on the minimum requirement for a complex immune system in structural terms. These insights are important for our overall understanding of the regulation of immunity in health and disease.

Key words: C-type lectin receptor (CLR), leucine-rich repeat (LRR), NOD-like receptor (NLK), RIG-I-like receptor (RLR), Toll/interleukin-1 receptor domain (TIR), Toll-like receptor (TLR).

Abbreviations used: Apaf-1, apoptotic protease-activating factor 1; CARD, caspase recruitment domain; ASC, apoptosis-associated speck-like protein containing a CARD; CLR, C-type lectin receptor; CTD, C-type lectin domain; DD, death domain; dsRNA, double-stranded RNA; ECD, ectodomain; EMDV, encephalomyocarditis virus; IFN, interferon; IRAK, IL-1R-associated kinase; IRF, interferon regulatory factor; ITAM, immunoreceptor tyrosine-based activation motif; LPS, lipopolysaccharide; MDA, melanoma differentiation-associated gene; MEF2, myeloid differentiation factor 88; Mal, MyD88 adaptor-like protein; NLK, NAK/TLR/p53 domain-containing protein; NBS, nucleotide-binding site; NF-κB, nuclear factor κB; NOD, nucleotide-binding and oligomerization domain; NLK, NOD-like receptor; PAMP, pathogen-associated molecular pattern; PRR, pattern-recognition receptor; PYD, pyrin domain; RIG-I, retinoic acid-inducible gene 1; RIP, receptor-interacting protein; RLK, RIG-I-like receptor; TRIF, TIR-like receptor; TMV, tobacco mosaic virus; TRAF, TNF (tumour necrosis factor) receptor-associated factor; VLR, variable lymphocyte receptor.

1To whom correspondence should be addressed (email palsson@tcd.ie).

Introduction

The recent and ongoing identification of germline-encoded innate immune PRRs (pattern-recognition receptors), as a first line of host defence, has revolutionized our understanding of innate immunity [1,2]. These receptors are central to the innate immune system, as well as helping to shape the subsequent adaptive immune response of vertebrates and the ‘effector-triggered’ immunity of plants. Invertebrates rely solely on PRRs as a universal and ancient defence mechanism for clearing invading pathogens. These PRRs all share the distinguishing feature of being able to initiate downstream signalling cascades which ultimately lead to altered gene expression. Other innate pathogen sensors, however, such as the serum complement components and mannose-binding lectin, initiate one of three complement activation pathways, classical, alternative and lectin, resulting in opsonization, followed by phagocytosis.
Figure 1 | Change in gene expression in response to invasive microbes is conveyed by four types of PRRs: TLRs, NLRs, RLRs and CLRs

The TLRs and CLRs activate the transcription factor NF-κB and also MAPKs, which are required for the expression of many immune and inflammatory genes, most notably cytokines and chemokines. Certain TLRs and the RLRs activate another family of transcription factors, the IRFs, which are required for expression of antiviral genes. Certain NLRs (e.g. Nalp3) activate caspase 1, which processes the pro-forms of IL-1β and IL-18. Signalling is activated via a receptor-specific subset of adaptors, with MyD88, Mal, TRIF and TRAM mediating the signalling. MyD88 is also used by IL-1R and IL-18R. RLRs signal via the adaptor IPS-1, whereas CLRs signal via CARD9. Cardif, CARD adaptor inducing IFNβ, MAVS, mitochondrial antiviral signalling protein; VISA, virus-induced signalling adaptor.

The PRRs recognize PAMPs (pathogen-associated molecular patterns) and can be subdivided into four distinct families: (i) TLRs (Toll-like receptors), (ii) CLRs (C-type lectin receptors), (iii) NLRs [NOD (nucleotide binding and oligomerization domain)-like receptors], and (iv) RLRs [RIG-I (retinoic acid-inducible gene 1)-like receptors] (Figure 1). Through studying the structural homologies of these pathogen sensors across the species, a picture of an evolutionarily conserved sensing system is beginning to emerge. In the present paper, we discuss these families of PRRs in terms of a set of structural domains. By including the Ig domain and the ITAM (immunoreceptor tyrosine-based activation motif), we can define a total of nine domains, used throughout Nature in various combinations to build immune systems with highly conserved features. These nine domains suggest that conserved immune systems evolved from one or a few ancient pathogen-sensing proteins, which have since evolved into a highly regulated, pathogen-specific and exceptionally efficient signalling network aimed at producing a unique response to a specific pathogen. This unexpected complexity and specificity continues the process whereby innate immunity is no longer seen as a non-specific unsophisticated aspect of host defence, but instead can be viewed as the core system of pathogen recognition, essential for the activation of appropriate responses for the elimination of infectious pathogens.

TLRs

As a starting point, we will first define the four receptor families. The trigger for the breakthrough in this field can be dated from the discovery of the TLRs. These type 1 integral membrane glycoproteins are key components of the host’s first line of defence, recognizing a variety of microbial products and initiating a complex immune response all designed to eliminate the invading pathogen [3]. The extracellular domain of TLRs is made up of multiple LRRs (leucine-rich repeats) which provide a platform for pathogen binding, giving each distinct TLR a specific pattern-recognition site. Intracellularly, the TLRs hold a signalling domain known as the TIR [Toll/IL (interleukin)-1 receptor] domain, which is the defining motif of the now well-established TLR/IL-1R (IL-1 receptor) superfamily [4,5]. The founding member, Toll, is a protein involved in establishing dorsoventral polarity in Drosophila melanogaster, which was subsequently found to also play a role in antimicrobial peptide production in the adult fruitfly [6]. Homologies between downstream signalling components of the Toll pathway and the type I IL-1R (IL-1RI) provided a template for understanding the pathway leading to the activation of NF-κB (nuclear factor κB) by IL-1, a pathway which, to a large extent, is shared by the majority of the 13 TLRs identified to date (13 TLRs in mice and ten in humans [7,8]). Lipid-based structures form the ligands for TLR2 and TLR4. The best characterized of these ligands are bacterial or mycobacterial lipopeptides, or glycerophosphatidylinositol anchors from parasites, both of which are recognized by TLR2 (in combination with TLR1 or TLR6), and bacterial LPS (lipopolysaccharide), which is recognized by TLR4. TLR3, TLR7, TLR8 and TLR9 all recognize viral and/or bacterial nucleic acids. The most studied examples are the recognition of dsRNA (double-stranded RNA) by TLR3, and the recognition of unmethylated CpG motifs in DNA by TLR9. Lastly, TLR5 recognizes bacterial flagellin and TLR11 (in mouse only) recognizes the flagellin-like molecule profilin [9].

Much effort has been focused on establishing the role of four adaptors involved in TLR signalling (Figure 1) [10]. Ligand binding, as in the fruitfly homologue Toll, results in two TLR molecules coming together, which results in a
TIR–TIR interaction to form a dimer [11–14]. MyD88 (myeloid differentiation factor 88), which was the first TLR adaptor described [15], also has a TIR domain which allows it to dock through a TIR–TIR interaction to the intracellular domain of the TLR dimer. MyD88 is used by all of the TLRs, except for TLR3, as well as by receptors for two important pro-inflammatory cytokines, IL-1R and IL-18R. Broadly speaking, downstream events involve the death domain of MyD88 interacting with the death domain of IRAK (IL-1R-associated kinase)-4, allowing it to become active and phosphorylate IRAK-1 which in turn activates TRAF [TNF (tumour necrosis factor) receptor-associated factor]-6. Subsequent ubiquitination of TRAF-6 and the protein kinase TAK-1 [TGF (transforming growth factor)-β-activated kinase 1] allows the IKK (inhibitor of NF-κB kinase) complex to become active, leading to the activation of NF-κB, which is one of the main players in regulating transcription of genes that are involved in immunity and inflammation [16]. Furthermore, this pathway gives rise to the activation of two MAPKs (mitogen-activated protein kinases), p38 and JNK (c-Jun N-terminal kinase), also resulting in an increase in transcriptional events and in the stabilizing of AU repeat-containing mRNAs.

The discovery of three more adaptors followed, starting with Mal (MyD88 adaptor-like protein), also known as TIRAP (TIR domain-containing adaptor protein) [17], which provided the first clue to the potential downstream specificity involved in TLR signalling [18]. Using Mal-deficient mice, Mal was shown to be specific to TLR2 and TLR4 signalling, although single and double Mal- and MyD88-knockout mice retained the ability to produce IFN (interferon)-β through the activation of IRF (interferon regulatory factor) 3, suggesting a pathway that is independent of Mal and MyD88. The reason for this has since been found through the discovery of two further adaptors, TRIF (TIR-related adaptor protein inducing IFNβ), which is required by TLR3 and TLR4 [19] and TRAM (TRIF-related adaptor molecule), which is specific to 'TLR4 signalling [20, 21]. These two adaptors make up the second delayed ‘MyD88-independent pathway’, leading to the activation of IRF3, through the IKK-like TBK-1 [TANK (TRAF-family member-associated NF-κB activator)-binding kinase-1] and induction of cytokines such as IFN-β. Clearly therefore, although the TIR domain is the common signalling domain, differences occur in the utilization of the different TIR-domain-containing adaptors which might provide specific tailor-made responses to an invading pathogen. The only clear example of this to date, however, is the ability of TRIF to induce cytokines such as IFNβ, which are more associated with antiviral responses, or T-cell activation via induction of the co-stimulatory molecules CD80 and CD86 in dendritic cells. The MyD88 pathway is more associated with inflammation.

**NLRs**

The NLRs, which have multiple other names, notably the CATERPILLER [CARD (caspase recruitment domain), transcription enhancer, R (purine)-binding, pyrin, lots of leucine repeats] gene family [22], are exclusively found intracellularly and are part of a family made up from at least 23 members whose defining motifs include a ligand-binding C-terminal LRR, a central NBS (nucleotide-binding site), which is thought to regulate self-oligomerization, and an N-terminal protein–protein interaction or effector domain, composed of a CARD or a PYD (pyrin domain) [23].

Broadly speaking, recognition of live bacteria in the cytosol by NLRs synergizes with extracellular recognition with TLRs, helping to clear the pathogen by up-regulating host cell synthesis of pro-inflammatory factors such as chemokines [including IL-8 and MIP (macrophage inflammatory protein)-2] and cytokines (IL-1β, IL-6 and IL-18) [23]. The NLR family members Nod1 (CARD4) and Nod2 (CARD15) detect distinct substructures from bacterial peptidoglycans, although the actual mechanism of ligand binding has not yet been clearly defined. Nod1 and Nod2 both activate NF-κB by forming oligomers upon ligand binding, which enables recruitment of RIP (receptor-interacting protein) 2 [also known as RICK [RIP-like interacting CLARP (caspase-like apoptosis-regulatory protein) kinase] or CARDIAK [CARD-containing ICE (IL-1β-converting enzyme)-associated kinase]] through a CARD–CARD interaction, and this complex in turn activates the IKK complex, resulting in NF-κB activation.

In contrast with Nod1 and Nod2, the NALPs (NACH/ LRR/PYD-containing proteins), which include NaIp2, NaIp3 and Ipaf, are important components of caspase-1-containing inflammasomes [24]. The inflammasomes are high-molecular-mass multiprotein complexes serving as scaffolds for caspase 1 recruitment and activation. Caspase 1 is a particularly important enzyme for the inflammatory response since it cleaves pro–IL-1β and pro–IL-18 into their mature forms. Central to all inflammasome assembly is ASC (apoptosis-associated speck-like protein containing a CARD), since mice deficient in ASC fail to produce IL-1β in response to stimulation [25]. ASC contains both a PYD and a CARD, allowing it to act as a direct bridge between the upstream sensors NALP and the downstream effector caspase 1 through homotypic protein–protein interactions. Three distinct types of inflammasomes have been described [25–27]. First, the NaIp3 inflammasome, comprising NaIp3, ASC, cardinal and caspase 1, is required for the activation of caspase 1 in response to LPS followed by ATP, the bacterial lipopeptide mimetic Pam3CSK4, R848 or bacterial CpG DNA. Activation of caspase 1 in response to flagellin occurs in a TLR5-independent manner through the assembly of the second type of inflammasome, the Ipaf inflammasome, which is made up by Ipaf, ASC and caspase 1, and the final inflammasome described to date is thought to be induced by mechanical lysis and consists of NaIp2, NaIp1 (CARD7), ASC, caspase 5 and Caspase 1.

**RLRs**

The RLRs are similar to NLRs in that they are cytosolic. They differ, however, in that they are exclusively viral sensors that complement the TLR viral sensor system of dendritic cells made up by TLR3, TLR7, TLR8 and TLR9. Similarly...
to the NLRs, the RLRs contain a CARD. They also contain a helicase domain. The best characterized RLRs are RIG-I and MDA (melanoma differentiation-associated gene) 5 [28,29]. They detect intracellular viral dsRNA, with RIG-I detecting uncapped 5′-triphosphate dsRNA or ssRNA (single-stranded RNA) and MDA5 detecting poly(I:C) and dsRNA from EMCV (encephalomyocarditis virus). Both receptors mediate immunity by inducing the production of type I interferons. RIG-I has been implicated in sensing dsRNA from Newcastle disease virus, vesicular stomatitis virus and Japanese encephalitis virus, whereas MDA5 detects dsRNA from EMCV and causes an increase in the expression of type I interferons in response to poly(I:C).

RIG-I and MDA5 both signal through a CARD–CARD interaction with a common adaptor, IPS-1 (IFN-β promoter stimulator 1), also known as MAVS (mitochondrial antiviral signalling protein), VISA (virus-induced signalling adaptor) and Cardif (CARD adaptor inducing IFN-β) (Figure 1), which in turn can signal through TRAF3, TBK1, IKKi and IRF3, giving rise to an increased expression of type I interferons, similarly to that induced by TLR3 [30,31].

Another member of the RLR family is LGP2 which may have a negative regulatory role in RLR signalling by sequestering some of the RNA ligand, but failing to signal due to a lack of a CARD [32].

The similarities in response elicited by antiviral TLRs and the RLR system suggest that the two sensor systems have distinct cellular responses to eliminate a viral infection, with the TLRs covering any viral infection occurring in dendritic cells [33], and the RLRs being responsible for a more widely occurring antiviral reaction [34].

**CLRs**

The final group of PRRs are the CLRs, which refers to a large family of proteins found almost exclusively in metazoans [35]. The term C-type lectin is somewhat of a misnomer, as the term refers to calcium carbohydrate-binding proteins and the family also includes proteins that do not bind calcium or have any carbohydrate specificity, but refers to a group of proteins with one or more defining C-type lectin domains. This structural motif was initially described as the protein fold of the carbohydrate-recognition domain of mannose-binding lectin and has since been found in over 1000 proteins, of which those CLRs found in vertebrates alone can be divided into 17 different groups [35]. The majority of CLRs mediate endocytosis and/or phagocytosis, play a role in antigen presentation and keep endogenous glycoprotein levels constant, although a small subset of CLRs gives rise to altered gene expression in response to various PAMPs, and these belong to the CLR V subfamily in vertebrates. Dectin-1 is a member of this subgroup and has a defining intracellular ITAM. Dectin-1 recognizes the fungal component zymosan through a single extracellular CTLD (C-type lectin-like domain), once again providing a backup or synergizing response to the fungal-sensing TLR system [36]. This becomes clear from recent studies using dectin-1-null mice, which are rendered more susceptible than wild-type mice to certain fungal infections such as Candida albicans and Pneumocystis carinii. Upon ligand recognition, the ITAM of dectin-1 recruits the tyrosine kinase Syk, ultimately resulting in the activation of NF-κB through the adaptor CARD9, MAL1 (mucosa-associated-lymphoid-tissue lymphoma-translocation gene 1) and Bcl-10; another transcription factor NFAT (nuclear factor of activated T-cells) and p42/p44 MAPK, all culminating in an increased expression of pro-inflammatory cytokines (Figure 1) [36]. A Syk kinase-independent signal also occurs, mediating phagocytosis independently of TLR2 [37].

The CLR adaptor CARD9 consists of an N-terminal CARD and a C-terminal coiled-coil domain and is structurally closely homologous with the well-characterized CARD-containing adaptor CARMA1 (CARD/membrane-associated guanylate kinase 1) [38]. CARD9 is expressed in myeloid cells, including macrophages and dendritic cells, and has also been implicated in Nod2 signalling in response to peptidoglycan and TLR-mediated activation of MAPK (but not NF-κB) through RIP2 [39,40]. A proposed model is emerging, with TLR, NLR and CLR signalling all converging on CARD9, eliciting distinct cytokine responses, although undoubtedly future studies will reveal more detail with regard to how this occurs [38].

**The nine domains required to build an immune system**

Having defined the four receptor families, we will now describe in detail the nine structural domains used to build each receptor.

**The TIR domain**

As mentioned above, the TIR domain (Table 1) is a key signalling region of all the TLRs and IL-1R, but is also shared by the intracellular signalling adaptors MyD88, Mal, TRIF and TRAM [41]. It also occurs in a fifth cytosolic adaptor SARM (sterile α-motifs and β-catenin/armadillo repeat motif). The TIR domain is found in many species, including mammals, plants and insects, suggesting that it arose in the common unicellular ancestor to plants and animals.

Signalling through the intracellular TIR domains of TLRs occurs through receptor dimerization upon ligand binding. This allows the two TIR domains to come into close proximity to each other and form a symmetrical association [11,14]. Subsequent structural reorganization creates a signalling platform which allows members of the adaptor subfamily to become recruited. The approx. 160 amino acids that define the TIR domain comprise five β-sheets (βA–βE) surrounding five α-helices (αA–αE) connected by flexible loops, and the structural importance of this particular folding pattern is evident through evolutionary conservation. The loops are named according to the structural elements that they connect (for example, the DD loop connects strand βD and helix αD). Within the sequence of the TIR domain lie three highly conserved regions denoted box 1, box 2, and box 3, and a structural model of the TIR domain shows that many of the conserved residues lie buried within the core of
Table 1 | The nine domains, which are evolutionarily conserved throughout the plant and animal kingdoms, that play key roles in sensing PAMPs and triggering a host defence response

See the text for further details.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Receptor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ab/TCR</td>
</tr>
<tr>
<td>Ig</td>
<td></td>
</tr>
<tr>
<td>LRR</td>
<td></td>
</tr>
<tr>
<td>CTLD</td>
<td></td>
</tr>
<tr>
<td>ITAM</td>
<td></td>
</tr>
<tr>
<td>TIR</td>
<td></td>
</tr>
<tr>
<td>NBS</td>
<td></td>
</tr>
<tr>
<td>Death Domain superfamily</td>
<td></td>
</tr>
<tr>
<td>CARD</td>
<td></td>
</tr>
<tr>
<td>Pyrin</td>
<td></td>
</tr>
<tr>
<td>Helicase</td>
<td></td>
</tr>
</tbody>
</table>

The nine domains, which are evolutionarily conserved throughout the plant and animal kingdoms, that play key roles in sensing PAMPs and triggering a host defence response. Hence the variable amino acids are exposed on the surface of the structure which is where the loop structures reside. Of structural importance here are the BB loop and DD loop of the TIR domain, both of which are critical for TIR–TIR interaction and specificity, possibly through a DD loop to BB loop homotypic interaction [4].

Ig domain

The TIR domain can be found in association with two other domains in PRRs, the Ig domain and the LRR domain. The Ig domain is made up of approx. 110 amino acids and, within each domain, an intrachain disulfide bond forms a characteristic loop of approx. 60 amino acids, called an Ig fold. Members of the IL-1R family have extracellular Ig domains [43]. IL-1RI has three such domains, which is the case with all other IL-1R-like receptors. One notable exception is a protein termed SIGIRR (single Ig domain IL-1R-related protein) which, as the name suggests, has only one Ig domain [44]. The Ig domains in IL-1RI, combined with the TIR domain, obviously play a central part in innate immunity, but, interestingly, from an evolutionary perspective, the Ig domain also dominates the adaptive immune response, forming the basic structure of antibodies and the extracellular portion of the TCR (T-cell receptor) and its accompanying CD3γ, ε and δ chains (Table 1). What marks out the Ig domain in this regard is the variable region which can generate tremendous diversity, forming the basis for high-affinity recognition of antigenic peptides and acquired immunity. It is therefore intriguing that the same domain is used in mammals in association with the ancient TIR domain, in the form of IL-1RI and related receptors.

LRR

The second domain found in association with the TIR domain is the LRR domain (Table 1). The LRR–TIR combination is found in animal and plant TLRs. Our understanding of ligand recognition and binding through this region has improved since Bell et al. [11] published the 2.4 Å (1 Å = 0.1 nm) crystal structure of the ECD (ectodomain) of TLR3. The overall structure is described as a horseshoe-shaped solenoid composed of 23 LRRs, with one side of the ECD exhibiting potential ligand-interacting properties [11].

A more recent intriguing finding has been that, in certain species, the LRR, as opposed to the Ig domain, is used to generate diversity. Through the discovery of the VLRs (variable lymphocyte receptors) in jawless fish, it has also been demonstrated how an adaptive immune response can be generated, not by the traditional Ig antibody gene rearrangement, but from a single germline VLR gene by genomic rearrangement of flanking LRR cassettes [45,46]. These VLRs are subsequently shed during an infection and the increased titre of VLR proteins specific to an immunized antigen can be measured in the plasma, as can the increased number of VLR-positive large lymphocytes in the blood. This suggests that the hagfish and the lamprey (the two sole survivors of the jawless fish of the cyclostome taxon) are the earliest creatures with an adaptive immune system, proving that the adaptive immune response may not be unique to jawed vertebrates.

©The Authors Journal compilation ©2007 Biochemical Society
Similarly to the Ig domain, the LRR domain can also associate with the TIR domain to generate a signalling PRR. This combination is also found in non-vertebrate organisms, re-emphasizing the evolutionary importance of this structure. In a group of plant proteins, a third domain is found with the LRR and TIR domains, termed the NBS [or NBD (nucleotide-binding domain)] [47]. These proteins have been compared with the NLRs of mammals as they detect intracellular pathogen-associated proteins, usually of viral origin. Some of these plant proteins lack a TIR domain, and the family overall is termed the NBS–LRR family. The NBS–LRR proteins form part of the effector-triggered immunity of plants, corresponding to mammalian adaptive immunity, and genes encoding proteins which confer specificity are referred to as resistance (R) genes, of which the majority encode NBS–LRR proteins [48]. The microbe in turn has to express a corresponding dominant avirulence (Avr) gene in order to trigger a resistance response, and for the plant to avoid becoming diseased. A classical model for this 'gene-for-gene' response is represented by the interplay between TMV (tobacco mosaic virus) and tobacco plants bearing the so-called N gene [49]. The product of this gene confers resistance to TMV on the tobacco plant, as becomes evident in plants lacking the N gene. The N protein has a TIR–NBS–LRR structure, providing further evidence of the importance and general use of the evolutionarily conserved TIR domain.

The CARDs and PYDs

LRRs also form the key recognition domain in the NLRs, providing another commonality with the TLRs. However, in the case of the NLRs, the LRRs are found in association with the NBS, as occurs in the plant TLRs, and also a CARD or a PYD (Table 1). CARDs and PYDs are both part of a large group of proteins that go by the name of the 'death domain superfamily' [50]. These proteins, which also include the DD (death domain) subfamily and the DED (death effector domain) subfamily mediate homotypic interactions within each domain subfamily, conveying signals which induce assembly and activation of apoptotic and inflammatory protein complexes. The defining feature of the DD superfamily is a six-helical bundle (H1–H6) structural fold which was first described in an NMR study where the CARD of RIADD (RIP-associated protein with DD) and the PYD of NALP1 were included. Although this fold is evolutionarily conserved, both the CARD and the PYD subfamilies exhibit distinct structural and sequence characteristics. The CARD and PYD both play critical roles in gene induction in a pro-inflammatory response, as is seen from the RLRs which contain a CARD, and the NLRs which contain a CARD, but in some instances also contain a PYD. The DD superfamily is also of key importance to downstream components in TLR-induced signal transduction, since the DD domain occurs in both MyD88 and the IRAK proteins.

From an evolutionary perspective, the DD superfamily is found in numerous multicellular organisms, with far fewer members being present in lower organisms, which matches the increased complexity of inflammatory and apoptotic responses occurring in higher organisms.

The crystal structure of a complex with Apaf-1 (apoptotic protease-activating factor 1) CARD and caspase 9 CARD has revealed in detail how this interaction occurs via three positively charged helices of caspase 9 and the convex surface made up of two negatively charged helices on Apaf-1 CARD [50,51]. Whether this interaction also holds true for other CARD–CARD interactions remains to be seen. In addition, it is also presently unclear whether CARD-containing proteins self-associate through a stable CARD–CARD interaction or whether they mostly form homotypic associations with other CARD proteins.

Although at present the PYD subfamily of DD proteins is the least well characterized out of the four mentioned, the recent NMR structures have removed any doubt as to whether it really is a true DD family member or not. PYD displays the typical six-helical bundle fold, but with one key difference, since helix H3 is completely replaced by a flexible loop. This is particularly interesting as H3 seems to play a critical role in protein–protein interactions in the entire DD superfamily, suggesting that PYD–PYD interactions may be significantly different from other interactions within the DD superfamily [52]. How PYD–PYD interactions occur is presently unknown.

NBS

As part of the structure of all NLRs, plant TLRs and the inflammasome components ASC and pyrin, the NBS (of which there are three subfamilies termed NACHT, NB–ARC (nucleotide-binding adaptor shared by Apaf-1, R proteins and CED-4) (in plant TLRs) or NOD) plays an important role in self-oligomerization (Table 1) [53]. Preserved motifs such as these usually reflect the importance functionally of the region, and this holds true for the NBS as mutations in this domain give rise to several diseases. For example, mutations within the NBS of NOD2, flanking both sides of the Mg^{2+}-binding site give rise to Blau syndrome with inflammatory characteristics such as uveitis, arthritis and dermatitis. Furthermore, mutations of the NLR member Nalp3 (or CIASI) are associated with multiple auto-inflammatory conditions, e.g. FCAS (familial cold autoinflammatory syndrome), commonly known as FCU (familial cold urticaria), NOMID (neonatal–onset multisystem inflammatory disease) and MWS (Muckle–Wells syndrome) [54,55], which manifest themselves in recurrent episodes of inflammatory attacks associated with arthritis, fever and rash. A total of 27 different disease-associated mutations have been identified and, quite astonishingly, all of these lie within the central NBS exon of Nalp3 [55], re-emphasizing the critical role of a functioning NBS for controlled inflammatory responses.

CTLD

The major non-opsonic fungal β-glucan CLR, dectin-1, has two defining motifs: an extracellular ligand-recognition motif termed the CTLD and the intracellular effector domain, ITAM. The crystal structure of the β-glucan-recognition
domain of dectin-1 was published recently [56]. The CTLD is connected via a stalk to the transmembrane domain which is followed by the intracellular ITAM. The structure suggests a binding mechanism which involves a variety of interactions, including both hydrophobic packing and hydrogen bonds. This detailed study of dectin-1 ligand recognition may suggest how ligand binding by other CTLD-containing receptors occurs.

ITAM
The ITAM as described above is the intracellular effector domain of certain CLR receptors, such as dectin-1, but the ITAM is also an essential part of signalling through the TCR complex (Table 1) [57]. Ten copies of the ITAM appear on the intracellular face of a fully assembled TCR complex, six of which occur on the TCRζ homodimer subunits, and a further four ITAMs are contributed by the accompanying CD3 heterodimers (γε and δε). In addition to TCRs and CLRs, the ITAM also occurs in the B-cell receptor, NK (natural killer) receptors and particular Fc receptors. Therefore, similarly to the Ig domain, the ITAM domain occurs in receptors that are important for both innate and adaptive immunity.

RNA helicase domain
The final domain described here is critical to the innate antiviral immune system and occurs in the RLRs, namely the RNA helicase domain. This domain was first described in DEXD/H-box helicases and is of key importance to RLR function [31]. Although the CARD of RIG-I mediates the signal, the RNA helicase domain is critical for the function of CARD, and it appears to be the helicase that recognizes and interacts with viral intracellular dsRNA. This activates the innate ATPase activity of the helicase which helps to regulate signal transduction in an ATPase-dependent manner, although precise details are still lacking.

Conclusions
The last 5 years have seen remarkable progress in our understanding of innate immunity. Four major receptor families have been found that recognize pathogens, and nine protein domains are known to be used to construct these receptors. These domains, used in various combinations, are core to the majority of all known immune systems of multicellular organisms ranging from plants to mammals. These domains probably arose in a common ancient pathogen-recognition system. As a consequence of their effectiveness, the domains were conserved over time and their uses have since expanded, giving rise to diversity within each receptor family, and also, in the case of the LRR and Ig domains, forming the basis for adaptive immunity in certain organisms. These continuing insights will allow us to provide a detailed molecular description of the events that occur during innate immunity.

References

Received 6 September 2007
doi:10.1042/BST0351437