THE SOLE GATEWAY TO ENDOTOXIN RESPONSE: HOW LPS WAS IDENTIFIED AS TLR4, AND ITS ROLE IN INNATE IMMUNITY

BRUCE BEUTLER AND ALEXANDER POLTORAK

The Scripps Research Institute, La Jolla, California

This paper is available online at http://dmd.aspetjournals.org

ABSTRACT:

TLR4 has emerged as a specific conduit for the bacterial lipopolysaccharide (LPS) response. The fact that such a protein exists, and furthermore, the fact that it is one member of a family of proteins expressed by mononuclear cells, yields considerable insight into the mechanism by which phagocytes “see” the microbial universe. It cannot yet be assumed that all of the TIRs have specificity comparable to that of TIR4, but it is probable that they do, given the molecular constraints to which all proteins are subject. Indeed, it is remarkable that TIR4 is able to sense so diverse an array of LPS molecules as it does. The total number of TIR proteins is not yet known. Although approximately 30 leucine-rich proteins bearing Toll-like cytoplasmic domains might be anticipated based on a survey of the genes in Drosophila, far fewer Toll-like genes have been found in mammals to date, although approximately 2 million expressed sequence tag sequences are now archived, and much of the genome has been covered. Some of the Toll-like proteins are, in fact, cytokine receptors. Ten leucine-rich TIRs have been reported so far. Even a small number of receptors might be sufficient to confer recognition of most pathogens, be they fungi, bacteria, or protozoa. Some such receptors may also play developmental roles. The mutational deletion of TIR genes alone and in combination with one another may help to establish the functions of each member of this newly emergent family of proteins.

In 1908, Elie Metchnikoff and Paul Ehrlich shared the Nobel Prize in Medicine. Metchnikoff was honored for his discovery of phagocytes, which as he correctly surmised, were an essential component of host defense against infection. In effect, he identified the cellular basis of what we now call “innate immunity”. Ehrlich was honored for his discovery of “antitoxins”, known today as antibodies. In so doing, he laid the foundations of the science of humoral immunity. As the Nobel Committee clearly recognized at the time, phagocytes and antibodies operate in concert with one another, each providing an important measure of protection against microbial pathogens (Mörner, 1908).

It is worth noting that Metchnikoff did not initially discover phagocytes in humans or in mice, nor in any other vertebrate, but in the larvae of starfish impaled by tangerine thorns, and in freshwater Daphnia, where he observed ingestion of fungal spores and anthrax bacilli. By contrast, antibodies have long been known to be produced by vertebrates only. As such, it is clear that innate immunity antedated humoral immunity. The latter system of defense was built to supplement the former.

Humoral immunity is one of the crowning glories of evolution. It arose in a remarkably short period of time. To do so, it needed to satisfy two requirements of any immune system. First, it was necessary for specific immunity to deal with a diverse range of pathogens. Second, it was necessary for the system to discriminate between self and non-self, hence avoiding the scenario of horror autotoxicus, envisioned by Ehrlich himself. The first requirement was met through reliance upon an anticipatory strategy in which avid receptors (T-cell receptors and immunoglobulins) are fashioned for virtually any molecule that might ever be encountered. This process requires a mechanism of genomic rearrangement that is, as far as we know, unique in all of nature. The second requirement was met by the process of clonal selection; also something quite unique, and as yet, only partly understood.

As might be assumed, the innate immune system overcame the dual problems of pathogen diversity and self-recognition long before vertebrates evolved. Yet only now have we begun to understand how this was accomplished. Lacking any known means by which to generate receptor diversity, cells of the innate immune system came to rely upon plasma membrane proteins that engage phylogenically conserved determinants on pathogens. These “pattern recognition receptors” (Janeway, 1992; Medzhitov and Janeway, 1998) comprise the afferent limb of innate immunity. One such receptor—and its cognate ligand—has now been identified. It points the way to the identification of all others. It is the receptor for bacterial lipopolysaccharide (LPS1), or “endotoxin”.

Endotoxin and Gram-Negative Infection

As recently recounted (Rietschel and Westphal, 1999), the word “endotoxin” was first coined by Pfeiffer (1858–1945) to describe the abundant component of Gram-negative organisms that evokes fever, shock, and other disturbances in mammals. Ultimately, at least the major endotoxic principle was defined as lipopolysaccharide (LPS), a variable amphiphilic structural component of the outer leaflet of the outer membrane of nearly all Gram-negative bacteria. LPS is not synthesized by Gram-positive organisms. Although other molecular components of Gram-negative and Gram-positive bacteria (e.g., lipoteichoic acid, lipopeptides, and other amphiphilic molecules) also exhibit pyrogenicity, LPS is probably the most biologically significant component that might ever be encountered. This process requires a mechanism of genomic rearrangement that is, as far as we know, unique in all of nature. The second requirement was met by the process of clonal selection; also something quite unique, and as yet, only partly understood.

As might be assumed, the innate immune system overcame the dual problems of pathogen diversity and self-recognition long before vertebrates evolved. Yet only now have we begun to understand how this was accomplished. Lacking any known means by which to generate receptor diversity, cells of the innate immune system came to rely upon plasma membrane proteins that engage phylogenically conserved determinants on pathogens. These “pattern recognition receptors” (Janeway, 1992; Medzhitov and Janeway, 1998) comprise the afferent limb of innate immunity. One such receptor—and its cognate ligand—has now been identified. It points the way to the identification of all others. It is the receptor for bacterial lipopolysaccharide (LPS1), or “endotoxin”.

1 Abbreviations used are: LPS, lipopolysaccharide; IL, interleukin; NF-κB, nuclear factor-κB.
of these. LPS is distinguished by a lipid A moiety, in which primary and secondary acyl chains are linked to a disaccharide-phosphate backbone, a ketodeoxyoctulosonic acid moiety, and a polysaccharide moiety of highly variable structure. The lipid A portion of LPS is toxic; the other constituents are not.

The degree of lipid A toxicity is not fully definable on a chemical basis, but at least partly, it is related to the number and density of secondary acyl chains. There are marked interspecies differences in LPS toxicity. Moreover, sensitivity to LPS within a species may be modulated by a number of factors, including glucocorticoids (which dampen sensitivity), and interferon-γ (which augments sensitivity).

Infection by facultative intracellular bacteria such as *Bacillus Calmette-Guerin* can boost the lethal effect of LPS by several orders of magnitude (Shands et al., 1971; Peavy et al., 1979; Berendt et al., 1980).

It is believed that LPS initiates many of the devastating effects observed in Gram-negative sepsis. Hypotension, coagulopathy, and injury to several organ systems are particularly severe in Gram-negative infection. Although it was once believed that LPS caused these effects through a direct action on the endothelium and parenchymal organs concerned, it is now quite clear that it acts indirectly, causing cells of hematopoietic origin (Michalek et al., 1980) (and particularly macrophages (Freudenberg et al., 1986; Galanos and Freudenberg, 1993)) to secrete pro-inflammatory cytokines that in turn elicit these effects. TNF is the most important of these. Blockade of TNF activity yields substantial attenuation of LPS toxicity (Beutler et al., 1985), and direct administration of TNF will mimic LPS toxicity in normal animals (Tracey et al., 1986).

Tens of thousands of people succumb to Gram-negative infection in the US annually, and it is likely that hundreds of thousands of deaths occur worldwide. Hence, there is great interest in understanding the mechanism of endotoxicity. At the same time, genetic studies suggest that timely recognition of LPS is important to permit containment and eradication of a Gram-negative infection soon after inoculation (see below). Moreover, logical considerations suggest that LPS comprises a superb target for innate immune recognition. Present in abundance on nearly all Gram-negative organisms, it is indispensable for their growth and survival. It has no vertebrate structural analog. These considerations suggested long ago that the “receptor” for LPS might serve an important sensing function in the innate immune system.

**Endotoxin-Resistant Mice: Evidence for a Solitary LPS Response Pathway**

The LPS receptor remained shrouded in mystery until very recently. In fact, LPS was once thought to exert its effects by intercalation into the plasma membrane or by interaction with a variety of plasma membrane proteins. However, over 30 years ago, the identification of LPS-resistant mice (Heppner and Weiss, 1965) made such speculations appear improbable, if not altogether untenable. Rather, it appeared that a single pathway for LPS responses must exist, since mutations of a single gene could entirely ablate LPS responses (Watson et al., 1977, 1978).

Two separate defects in LPS signal transduction developed as the result of independent mutations that arose spontaneously and became fixed in the C3H/HeJ and C57BL/10ScCr substrains (Coutinho et al., 1977). The locus involved was named *Lps*, and in mice of the C3H/HeJ strain, the responsible allele was termed *Lps<sup>d</sup>* to denote a defective response to lipopolysaccharide. Closely related animals (e.g., C3H/HeN and C57BL/10ScSn) retained normal LPS sensitivity. They served as controls in investigations of the effects of LPS resistance, and permitted an educated guess as to the date at which each mutation occurred.

C3H/HeJ mice were found to have a codominant defect in LPS signal transduction, whereas animals of the C57BL/10ScCr strain were noted to have a recessive defect. Allelism was proven by the fact that F<sub>1</sub> animals derived from the cross of these two strains were profoundly unresponsive to LPS (Coutinho and Meo, 1978). The C3H/HeJ mutation was mapped to mouse chromosome 4 by Watson and coworkers (1978), who placed the *Lps* locus between Mup-1 and Ps loci.

Mice of the C3H/HeJ strain were noted to be abnormally susceptible to Gram-negative infection, suggesting that the mutation could prevent timely recognition and clearance-injected bacteria (O’Brien et al., 1980; Rosenstreich et al., 1982). The animals were also found to be sensitized to LPS by the adoptive transfer of spleen cells (Rosenstreich et al., 1977) [and later macrophages (Freudenberg et al., 1986)] from C3H/HeN animals; hence, it became clear that the macrophage mediates the toxic effects of LPS. Beyond this, cell from C3H/HeJ mice were used in a vast number of situations wherein the confounding effects of LPS were undesirable; e.g., in the isolation of interleukin-1 (Lachman et al., 1977; Mizel and Mizel, 1981). Yet, despite the enormous interest engendered by the *Lps* locus, all efforts to isolate its protein product were frustrated, as were attempts to clone the *Lps* cDNA.

Early speculations that *Lps* might encode the endotoxin receptor were turned aside by the demonstration that CD14, the glycosylphosphoinositol-anchored protein product of an unlinked gene, acts as the principal LPS-binding protein on the plasma membrane (Wright et al., 1990). Since CD14 has no clear means of signaling across the membrane, it appeared likely that *Lps* might encode a co-receptor capable of serving such a role, but no such molecule could be found.

**Fine Mapping of the Lps Locus**

In 1994, several groups began to map *Lps* to resolution far exceeding that achieved by earlier workers. Relying upon the growing density of microsatellite markers in the mouse genome, thousands of meioses were studied by Beutler and coworkers (University of Texas, Southwestern Medical Center), Schwartz and coworkers (University of Iowa), and Malo and coworkers (Montreal General Hospital). All three groups defined overlapping genetic intervals that, in retrospect, contained the *Lps* locus. However, the efforts to isolate and clone *Lps* remained independent from beginning to end.

Whether for reasons of bad luck or merely as the result of conservative phenotypic analysis, the critical region defined by the Dallas group (Poltorak et al., 1998b) was far larger than those defined by the Montreal (Qureshi et al., 1996) or Iowa City (Peiffer-Schneider et al., 1997) teams. It was spanned by an exceptionally dense contig that covered 3.2 Mb with a seamless array of YAC and BAC clones. A series of distance measurements, performed by fluorescence in situ hybridization on interphase nuclei using independent BACs as probes, established the order and distance of separation between clones. These measurements were supported by other methods, including the use of pulsed-field gel electrophoresis to determine BAC size, statistical estimates of physical distance based on BAC size, and incontrovertible genetic data derived from phenotypic analysis of animals with crossovers between novel microsatellites found within the interval. The critical region was ultimately reduced in size to 2.6 Mb of DNA, delimited by the novel markers “B” and “83.3” (proximal and distal, respectively). The entire contig was explored by a combination of exon trapping, hybridization/selection, and shotgun sequencing.
**TLR4—A Member of an Ancient Family of Defensive Proteins—Is the Only Gene in the Lps Critical Region**

Numerous pseudogenes were identified within the B \(\rightarrow\) 83.3 interval. Moreover, several genes were identified centromeric to B, before the exclusion of the proximal contig on genetic grounds. These included genes encoding a zinc finger protein, an isologue of Tera, the majority of the *Pappa* gene, the arylacetamide deacetylase gene, and an isologue of the astrotactin gene. Within the critical region itself, the gene-encoding mouse Toll-like receptor 4 (*Tlr4*) stood as the sole candidate (Poltorak et al., 1998a). The identification of novel mutations within the *Tlr4* locus of C3H/HeJ mice and C57BL/10ScCr mice (not present in the control substrains) immediately confirmed that *Tlr4* and *Lps* were identical. In C3H/HeJ mice, a point mutation modifies a conserved residue of the Tlr4 cytoplasmic domain (P712H). In C57BL/10ScCr mice a null allele was observed (Poltorak et al., 1998a). With the presentation of these data, the search for *Lps* was formally concluded on September 29, 1998.

The identity of *Tlr4* and *Lps* came as a startling revelation. In addition to its developmental function, the prototypic Toll gene of *Drosophila* was already known to subserve an important role in defense against fungal infection. The Toll pathway had a well-established counterpart in the IL-1 signaling cascade (Gay and Keith, 1991; Heguy et al., 1992; Muzio et al., 1997; Kopp et al., 1999), and *Tlr4* itself was known to share at least some components of this cascade (Muzio et al., 1997; Medzhitov et al., 1998), engaging MyD88. An immunodeficient human patient with coreistance to LPS and IL-1 had been described (Kuhns et al., 1997), suggesting that both inflammatory pathways utilize a common signaling intermediate. Hence, at one stroke, it became clear that both vertebrate and invertebrate immunity use homologous proteins for the detection of invasive pathogens.

Depending upon the structural definition applied, Toll-like proteins may be found among vertebrates, invertebrates, and even among plants, where, as in animals, they appear to defend the host against invading microbes.

In *Drosophila*, seven Toll homologs have been identified. One of these, 18-wheeler, offers protection against bacterial infection. In mammals, 10 members of the Toll superfAMILY with leucine-rich ectodomains (toll-like receptors [Tlr] 1–10) have been cloned to date. In addition, both chains of the IL-1 and IL-18 receptors, SIGIRR and MyD88, have cytoplasmic Toll-like domains. Although the binding specificities of IL-1R and IL-18R are clearly defined, that of SIGIRR and most of the other Tlrs is not. It may be assumed that they, and perhaps other Tlrs yet unknown, are involved in pathogen detection, development, or both. The phylogeny of the Toll-like receptors has been analyzed by parsimony methods, and some of the mammalian receptors appear to have arisen in the recent evolutionary past (Du et al., 2000).

**Incorrect Claims of TLR2 Involvement in LPS Signaling**

A parallel line of inquiry aimed at identifying the LPS signal transducer was undertaken by Genentech, Inc. (South San Francisco, CA) and by Tularik, Inc. (South San Francisco, CA). These companies seized upon the identification of numerous Toll-like human mRNAs, achieved by Rock et al. (1998), Chaudhary et al. (1998), and Medzhitov et al. (1997), who used expressed sequence tag probes with homology to Toll to extract full-length mammalian cDNAs. Medzhitov, in particular, had shown that ligation of Tlr4 (then called h-Toll) by construction of a CD4:Tlr4 chimera, caused NF-\(\kappa\)B translocation to the nucleus of mammalian cells (Medzhitov et al., 1997). Although in no instance was a ligand for any mammalian Tlr known, the fact that an inflammatory effect could be rendered via the Tlrs stimulated a systematic search for LPS signal transducing activity.

The search was predicated on the transfection of a non-macrophage cell line with CD14 (the physical LPS receptor) and each of a collection of Tlrs, together with an NF-\(\kappa\)B reporter construct. In these studies, it was found that Tlr2, but not Tlr4, could endow 293 cells with LPS-sensing activity.

It is, at this point, quite clear that the observation was a system artifact, brought about by several misguided assumptions. First, it should not have been taken for granted that 293 cells lack only CD14 and a transmembrane co-receptor subunit, and are otherwise competent to sense LPS. In fact, they appear to lack other proteins as well, and the total structure of the LPS signaling complex remains unknown to this day. Second, the equation of NF-\(\kappa\)B translocation with LPS signal transduction was clearly a mistake. Many stimuli provoke NF-\(\kappa\)B translocation. The effects of LPS are unique. Third, the unphysiologic expression of any biologically active receptor might lead to artificial signaling (and in the case of Tlr2, apparently did so).

In fact, as emphasized earlier in this review, genetic arguments hold that only a single pathway for LPS signal transduction exist in mice, and by implication, in all mammals. Since mutations of Tlr4 fully abolish signaling by pure preparations of LPS, there is no room for the contention that Tlr4 comprises an alternative pathway. This fact alone would not exclude an essential contributing role for Tlr2, but such a role was never the primary claim. Rather, it was asserted that Tlr2 could act as an independent signal transducer. In any case, it is clear that it is neither essential nor even necessary: a frameshift mutation of Tlr2 fails to impair LPS signal transduction in hamsters (Heine et al., 1999). Furthermore, Takeuchi and coworkers (1999) determined that a knockout mutation of Tlr2 in mice has no discernible effect on LPS signal transduction.

The concurrent claim that Tlr4 does not transduce the LPS signal is likewise a spurious one. In the correct cellular environment (i.e., a macrophage), Tlr4 (and not Tlr2), expressed at low copy number, signals the presence of LPS as measured by activation of the TNF gene; surely the most relevant endpoint for study. Moreover, expression of the dominant-negative Tlr4\(^Lps-d\) isoform of Tlr4 blocks such signaling almost completely.

For all of these reasons, the notion that Tlr2 is an LPS signaling protein may be discarded. Ironically, the claim was embraced by many workers in the field, and the method applied in the demonstration of LPS signaling via Tlr2 has already been used to support the claim that it transduces many other microbial signals as well (Aliprantis et al., 1999; Brightbill et al., 1999; Schwandner et al., 1999; Yoshimura et al., 1999). Such reports must be viewed with skepticism, although there is no doubt that Tlr2 detects peptidoglycan (Takeuchi et al., 1999) and certain bacterial lipopeptides (Takeuchi et al., 2000), based on work with mice bearing a knockout mutation of Tlr2.

**Direct Contact between Tlr4 and LPS Is Suggested by Genetic Experiments**

*Drosophila* Toll is stimulated by the ligand Spätzle, a protein generated from an inactive precursor via the action of three upstream proteases, named Gastrulation-defective, Snake, and Easter (Belvin and Anderson, 1996). It is believed that Spätzle may also be activated by a second proteolytic cascade, initiated by an unknown product of pathogenic fungi and that this cascade is blocked by an endogenous serine protease inhibitor (Spn43Ac) (Levashina et al., 1999). Hence, Spätzle is itself required for the response to fungi, but Gastrulation-defective, Snake, and Easter are not.

The situation in mammals is quite different. Close contact between...
LPS and Tlr4 has been established through genetic analysis. Although human cells respond to LPS or its toxic lipid A moiety, they fail to respond to non-acetylated congeners of lipid A (e.g., LA-14-PP), which actually antagonize LPS signaling (Dal Nogare and Yarbrough, 1990; Golenbock et al., 1991). On the other hand, mouse macrophages are responsive to both forms of LPS (Birkland et al., 1990). This phenotypic difference cannot be ascribed to CD14, since both forms of LPS are bound by both human and mouse CD14, and transfection studies do not support the notion that chemical specificity is exercised at this level (Delude et al., 1995). Hence a more distal component of the signaling pathway must be considered.

In fact, the species specificity for responses to distinct lipid A partial structures is determined at the level of Tlr4. Hence, if immobilized C3H/HeJ macrophages are transduced to express the human Tlr4 protein, they respond only to intact lipid A; if they are transduced to express the normal mouse protein, they respond to both forms of lipid A (Poltorak et al., 2000). Similarly, human mononuclear cells transduced to express hamster Tlr4 respond to both forms of lipid A (Lien et al., 2000).

Insofar as Tlr4 “reads” the LPS structure, discriminating between acylated and non-acylated forms, it must come into close contact with LPS. However, the duration of contact is not clear, and many details concerning the ultimate fate of the complex must be answered by direct physicochemical studies.

**Downstream Signaling Events**

Coimmunoprecipitation studies established that Tlr4, like the IL-1 receptor, engages MyD88, a cytoplasmic protein with a Toll-like domain as well as “CARD” or death domains. This protein, in turn, engages the serine kinase interleukin-1 receptor-associated kinase, which activates TRAF6, leading to phosphorylation of IκB and nuclear translocation of NF-κB. Interestingly, a knockout mutation of MyD88, which markedly attenuates the lethal effect of LPS and most responses to IL-1, was consistent with delayed activation of NF-κB and with initiation of the mitogen-activated protein kinase cascade (Kawai et al., 1999). Both of these events are thought to be important to LPS signal transduction, and it comes as a surprise to see that they may persist in the absence of toxicity. It may be concluded on this basis that other proteins besides MyD88 must participate in the proximal signal transduction pathway. Moreover, given the striking difference between cellular activation by LPS as compared with IL-1, certain proteins that serve the LPS receptor must be unique in their role as such.

**Genetic Variation at the Tlr4 Locus**

Among humans and among mice, a single allele of Tlr4 predominates, but many relatively uncommon variants are observable, specifying changes within the coding region. The ectodomain of Tlr4 is far more variable than the bulk of the cytoplasmic domain, whether within species or between them. This may be taken to suggest that the ectodomain, which appears to contact LPS, has varied to accommodate variable ligand structure.

The cytoplasmic domain is highly conserved over most of its length. However, the C-terminal end is hypervariable among species. This portion of the protein may dictate the “set point” for LPS response among members of a given species. The set point may be optimized with respect to responses that cytokines elicit, or otherwise matched to the specific immune response (Smirnova et al., 2000).

It may be offered as a prediction that individuals with specific TLR4 alleles will prove resistant to infection by specific Gram-negative organisms. It is apparent that at least one rather new Tlr4 allele has already risen to high frequency among Caucasians. This allele (TLR4B; Gb:177766), representing a double amino acid substitution, most probably originated as the result of second mutation, occurring on the background of a common African variant allele (Smirnova et al., 2001). The African progenitor allele is now exceedingly rare in Caucasian populations. Although a “hitch-hiker” effect (linkage disequilibrium relationship with a favorable but unrelated gene) and simple drift cannot be excluded, it is also possible that the relevant compound allele confers resistance to a specific Gram-negative organism, one which exerts a powerful selective pressure in the recent evolutionary past. It has very recently been shown that the allele has a discernible phenotype in that heterozygotes are hyporesponsive to *Escherichia coli* LPS preparations (Arbour et al., 2000). A complete understanding of the molecular specificity of Tlr4—and indeed, the other Tlrs—will likely require several years of intensive study.

**References**


