Glycan Sequence-Dependent Nod2 Activation Investigated by Using a Chemically Synthesized Bacterial Peptidoglycan Fragment Library

Ning Wang,[a] Cheng-yuan Huang,[a] Mizuho Hasegawa,[b] Naohiro Inohara,[b] Yukari Fujimoto,*[a] and Koichi Fukase*[a]

Nucleotide oligomerization domain-containing protein 2 (Nod2), an innate immune receptor, recognizes bacterial cell-wall peptidoglycan (PGN), the minimum ligand of which is muramyl dipeptide (MDP). Enzymatic digestion of PGN appears to be important for Nod2 recognition. PGN is degraded by muramidase or glucosamidase through a process that produces two types of glycan sequence; glycans containing GlcNAcβ(1→4)MurNAc or MurNAcβ(1→4)GlcNAc. In this report, a range of disaccharide or tetrasaccharide fragments of each sequence were chemically synthesized, and their activities in stimulating human Nod2 (hNod2) were investigated. The results reveal that hNod2 recognitions is dependent on the glycan sequence, as demonstrated by comparing the activities of glycans with the same peptide moieties. (MurNAcβ(1→4)GlcNAc)$_2$-containing structures exhibited stronger activity than those containing (GlcNAcβ(1→4)MurNAc)$_2$. The results suggest that differences in the enzymatic degradation process affect the host's immunomodulation process.

Introduction

Innate immune receptors recognize a variety of microbial components, known as microbe-associated molecular patterns (MAMPs) or pathogen-associated molecular patterns (PAMPs), to activate the immune system. A variety of innate immune receptors, also known as pattern-recognition receptors (PRRs), have been identified, including Toll-like receptors (TLRs), Nod-like receptors (NLRs), RIG-like receptor (RLR) and C-type lectin receptor (CLR). Nucleotide oligomerization domain-containing proteins 1 and 2 (Nod1, Nod2),[1,2] which are founding members of the NLR family, recognize peptidoglycan (PGN).[3–5] PGN contains long glycan chains consisting of alternating N-acetyl-glucosamine (GlcNAc) and N-muramic acid (MurNAc) linked by a β(1→4) bond. The glycans are connected to one another through a peptide linkage at the carboxylic acid of MurNAc, outside the plasma membrane. In most bacteria, L-Ala-γ-D-Glu or L-Ala-γ-D-GluNH$_2$ (L-Ala-γ-D-isoGln) is connected to MurNAc, and the third amino acid connecting to the Glu (or GluNH$_2$) is usually a diaminocarboxylic acid: L-Lys in many Gram-positive bacteria or meso-diaminopimelic acid in most Gram-negative bacteria or some Gram-positive bacteria. Our group has synthesized a variety of PGN fragment structures in an effort to search for structures recognized by the PGN receptors.[6] This work has revealed the minimum ligand recognition structures of Nod1 and Nod2.[1,3] Because Nod1 and Nod2 are cytosolic proteins and both proteins prefer smaller ligands rather than PGN itself,[1–4] enzymatic digestion of PGN, followed by ligand transport appears to be critical for Nod1 and Nod2 recognition. NOD2 mutations are associated with a susceptibility to inflammatory diseases, including early-onset sarcoidosis (EOS), Blau syndrome (BS), and Crohn's disease (CD).[7–9]

Our previous studies revealed that Nod2 recognizes a muramyl dipeptide (MDP, MurNAc-L-Ala-γ-D-isoGln) as the minimum ligand structure and also recognizes peptidoglycan fragments containing MDP. MDP showed the most potent Nod2-stimulating activity, and the activity decreased as the glycan and peptide chain lengths increased.[6,10–13] In these studies, we synthesized tetrasaccharide and octasaccharide fragments that contained the GlcNAc-MurNAc unit. In host organisms, sequences containing the GlcNAc-MurNAc unit are produced by lysozyme, also known as muramidase, which cleaves β(1→4) linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues in PGN.

Various PGN fragments are released from bacteria during bacterial cell lysis. These fragments are produced by bacterial muramidase and glucosamidase and, therefore, contain two glycan sequences: repeating units of GlcNAc-MurNAc or MurNAc-GlcNAc (Figure 1). The disaccharides MurNAc-GlcNAc and GlcNAc-MurNAc with dipeptides were previously synthesized,[14] but not clear differences in the immunostimulatory activities of the compounds were observed. Peptidoglycan fragments with more than two MurNAc-GlcNAc repeating units have not previously been synthesized; therefore, we investigated the biological activities of such synthetic fragments by testing their innate immunostimulatory activities through the intercellular receptor, Nod2. This study reports the first synthesis of tetrasaccharide fragments, particularly compounds contain-

[a] N. Wang, Dr. C.-y. Huang, Prof. Dr. Y. Fujimoto, Prof. Dr. K. Fukase
Department of Chemistry, Graduate School of Science, Osaka University
1-1 Machikaneyama, Toyonaka, Osaka 560-0043 (Japan)

[b] Dr. M. Hasegawa, Prof. Dr. N. Inohara
Department of Pathology, University of Michigan Medical School
1150 W. Medical Center Dr., Ann Arbor, Michigan 48109 (USA)

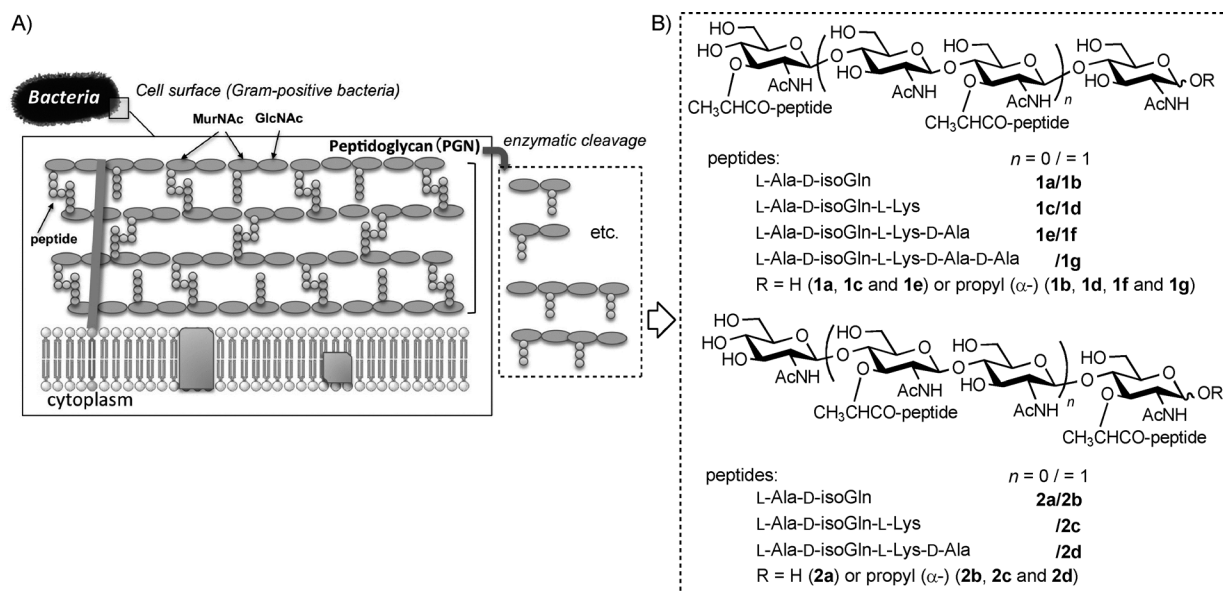Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201200655.

Figure 1. A) Cell surface of a Gram-positive bacterium and schematic diagram of the enzymatic cleavage of the cell-wall component, peptidoglycan. B) A PGN fragment library containing two types of glycan sequence.

ing MurNAc-GlcNAc repeating units. The activities of the tetrasaccharide-containing PGN fragments, as well as the di- and monosaccharide-containing fragments, were systematically investigated in human Nod2 (hNod2) for the first time.

Results and Discussion

Synthesis of peptidoglycan fragment library

Syntheses of PGN fragments have been reported previously, including by us.[5,6] In our studies, 2-N-Troc (Troc: 2,2,2-trichloroethoxycarbonyl) protection was used for both the glycosyl donor and acceptor for the glycosylation of glucosamine (GlcN) and muramic acid (MurN) in the synthesis of the complex PGN fragments.[10,11] A high β-selectivity was obtained with neighboring group participation of the N-Troc group. The glycosyl acceptor **4** showed a higher reactivity than the N-acetylated acceptors. In our latest study, we applied the same N-Troc protection strategy. We also examined the leaving groups in the donors: trichloroacetimidate[10,15] and N-phenyltrifluoroacetimidate.[16] N-phenyltrifluoroacetimidate showed good reactivity toward glycosylation at the 4-OH group in the MurNTroc residue, as shown below.

The repeating glycans in compounds **1a–1g** (Figure 1B) were constructed by preparing the key disaccharide glucosaminyl-β(1→4)-muramic acid intermediate **5**, as shown in Scheme 1, based on methods developed previously in our group.[10,11] β-Selective glycosylation with the

N-Troc-muramyl trichloroacetimidate donor **3** of the N-Troc-glucosaminyl acceptor **4** was carried out in the presence of catalytic amounts of trimethylsilyl trifluoromethanesulfonate (TMSOTf) as the activator to afford the disaccharide **5** in high yield (84%).

The disaccharide analogues were then synthesized from **5**, as illustrated in Scheme 2. Cleavage of the Troc groups of **5** with Zn/Cu in AcOH and subsequent acetylation with Ac_2O gave compound **6**. The allyl group was isomerized to a vinyl group by using H_2-activated [Ir(cod)(MePh$_2$P)$_2$]PF$_6$ (cod = 1,5-cyclooctadiene) to give compound **7**, and cleavage of the ethyl ester with LiOH gave **8**. The appropriate peptides were introduced to the liberated carboxylic acid **8** by using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride ("water-soluble carbodiimide", WSCD·HCl), 1-hydroxybenzotriazole (HOBt), and triethylamine (TEA). The vinyl group was then cleaved with iodine and H_2O, and hydrogenation with Pd(OH)$_2$ successfully gave the disaccharide fragments **1a**, **1c**, and **1e**.

Tetrasaccharide **12** was then synthesized by using **5** as a common synthetic intermediate for both the glycosyl donor
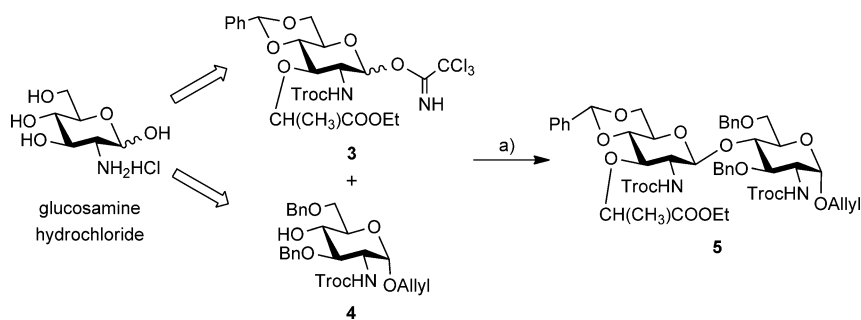
Scheme 1. Synthesis of the key disaccharide intermediate **5**. a) TMSOTf, CH$_2$Cl$_2$, −15°C, 4 Å molecular sieves, 20 min, 84%.
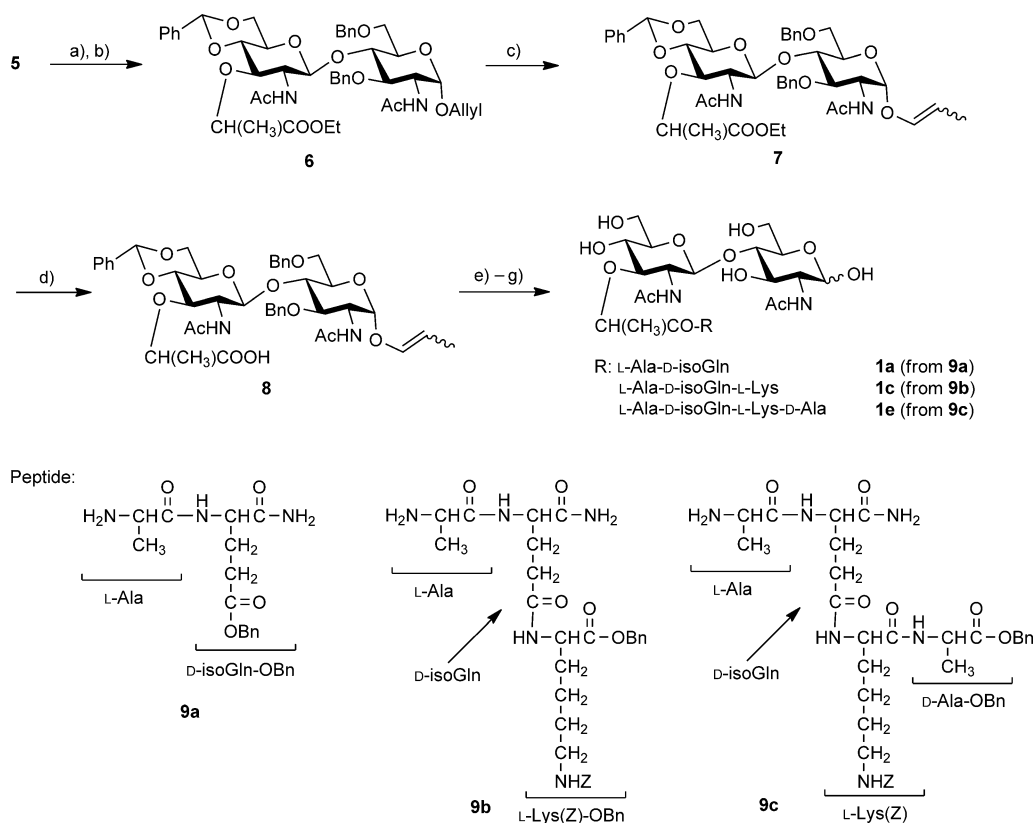
Scheme 2. Synthesis of the PGN disaccharide fragments **1a**, **1c**, and **1e**. a) Zn/Cu, AcOH, 30 min; b) Ac$_2$O/Py, 1 h, 74% (2 steps); c) [Ir(cod)H(MePh$_2$P)$_2$]PF$_6$; THF, 1.5 h, 68%; d) LiOH, dioxane/THF/H$_2$O (2:4:1), quant; e) **9a**, **9b** and **9c**, WSCD·HCl/HOBT/TEA DMF, 72% (with **9a**), 59% (with **9b**) or 48% (with **9c**); f) I$_2$/H$_2$O, 64, 54, 64%, respectively; g) H$_2$/Pd(OH)$_2$, AcOH, 91% (**1a**), 89% (**1c**), 90% (**1e**).

and acceptor (Scheme 3). The disaccharide donors **10a** and **10b** were prepared through cleavage of the allyl glycoside and subsequent conversion to their imidate forms (**10a**: trichloroacetimidate, **10b**: N-phenyltrifluoroacetimidate). Regioselective reductive ring opening of the 4',6'-O-benzylidene of **5** was carried out by using BH$_3$·Me$_3$N and BF$_3$·Et$_2$O in CH$_3$CN[17] to afford the disaccharide glycosyl acceptor **11** with a free 4-hydroxy group in 83% yield. For the preparation of the tetrasaccharide **12**, the trichloroacetimidate **10a** was first used as the glycosyl donor in a coupling reaction with the glycosyl acceptor **11**;[11] however, glycosylation between **10a** and **11** in the presence of TMSOTf as a Lewis acid gave the desired tetrasaccharide **12** only in 16% yield, with a 62% recovery of **11**. An increase in the reaction temperature or changes in the equivalents of reactants did not improve the yield. Possible reasons for the low yield include: 1) The 4-OH group of the disaccharide acceptor **11** has a low reactivity due to steric hindrance from the 3-O-lactyl moiety in the muramic acid residue; 2) trichloroacetimidate **10a** is highly reactive, but the activated cationic intermediate decomposes prior to the desired glycosylation because attack of **11** on the intermediate is sterically encumbered.

We then used the N-phenyltrifluoroacetimidate[16] **10b** as the glycosyl donor. Generally, glycosyl N-phenyltrifluoroacetimidates are also highly reactive but show better stabilities than the corresponding trichloroacetimidates. The formation of N-

glycosyl trifluoroacetamides as by-products is suppressed because the N-phenyl group of the eliminated N-phenyltrifluoroacetamide prevents the undesirable attack of the amide on the cationic intermediates. Glycosylation by N-phenyltrifluoroacetimidate **10b** was promoted by increasing the equivalent of acceptor **11** (donor/acceptor ratio 1:1.5). The glycosylation yield when using acceptor **11** improved dramatically to 61% of tetrasaccharide **12**.

Removal of the N-Troc groups from the amino groups at the 2-positions of **12**, N-acetylation of the liberated amino groups, and saponification of the ethyl esters produced the dicarboxylic acid **13** with a tetrasaccharide MGMG (MurNAc-GlcNAc)$_2$ sequence. The glycan backbone **13** was then coupled to di-, tri-, tetra-, and pentapeptides by using condensation reagents in DMF. Condensation of **13** with HCl·L-Ala-D-isoGln(OBn) and HCl·L-Ala-D-isoGln-L-Lys(Z)(OBn) was effected by using WSCD, HOBt, and triethylamine to give the protected tetrasaccharide containing two units of dipeptide **14** in 82% yield and the tripeptide **15** in 53% yield. Condensation of the longer peptide was achieved by using O-(7-azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HATU) as a coupling reagent along with triethylamine to obtain higher yields; this resulted in a tetrasaccharide with tetrapeptides, **16**, in 81% yield and a tetrasaccharide with pentapeptides, **17**, in 71% yield. All benzyl and benzylidene groups were then removed to obtain PGN fragments of tetrasaccharides having di-, tri-,
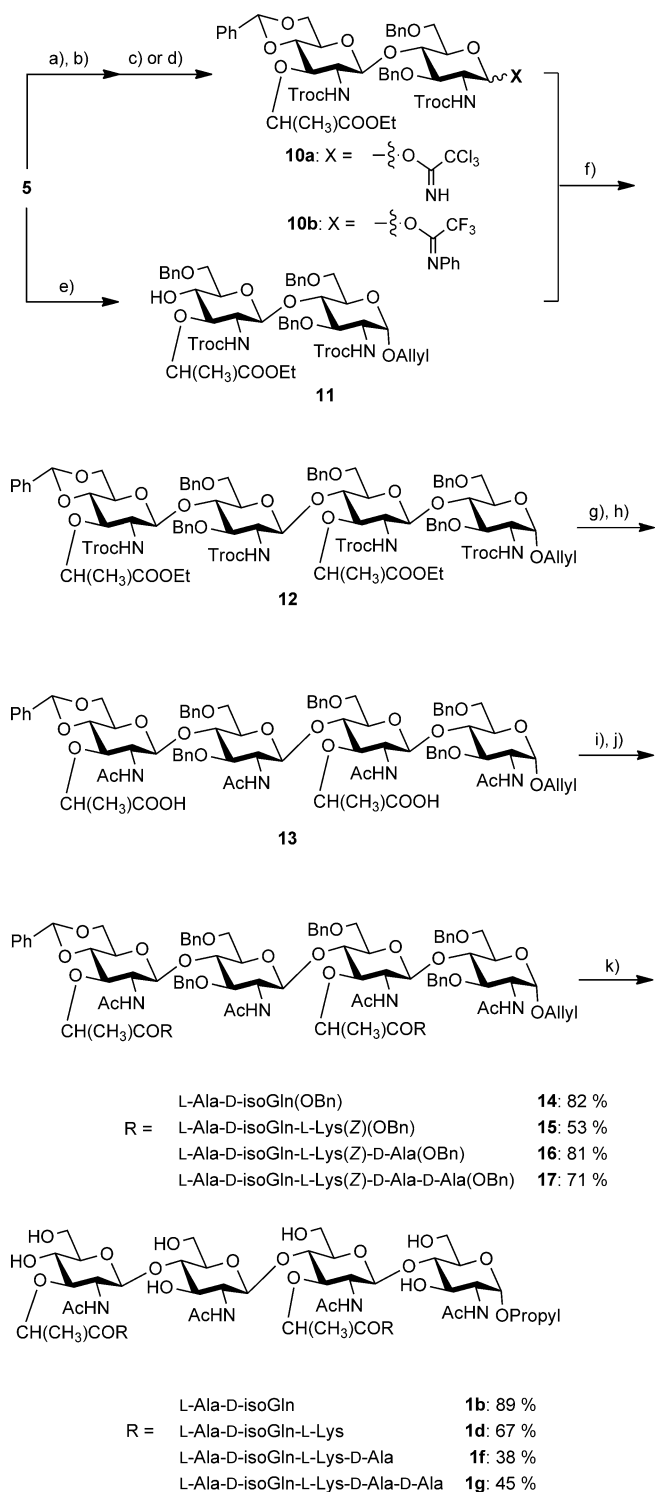
Scheme 3. Synthesis of PGN tetrasaccharide fragments. a) [Ir(cod)H-(MePh$_2$P)$_2$]PF$_6$, THF, 1.5 h; b) I$_2$/H$_2$O, 30 min, 81%; c) CCl$_3$CN/Cs$_2$CO$_3$, CH$_2$Cl$_2$, 30 min, quant. (**10a**); d) CF$_3$(NPh)Cl/Na$_2$CO$_3$, acetone, 3 days, 76% (**10b**); e) Me$_3$N·BH$_3$/BF$_3$·Et$_2$O, CH$_3$CN, 1 h, 83%; f) TMSOTf, CH$_2$Cl$_2$, −15 °C, 4 Å molecular sieves, 40 min, 16% (from **10a**) or 61% (from **10b**); g) Zn/Cu, AcOH, 3 h, Ac$_2$O/Py, 2 h, 62%; h) LiOH, quant.; Condensation: i) WSCD/HOBt/DMF/TEA, HCl·L-Ala-D-isoGln(OBn) (for **14**) or HCl·L-Ala-D-isoGln-L-Lys(Z)(OBn) (for **15**); j) HATU/DMF/TEA, HCl·L-Ala-D-isoGln-L-Lys(Z)-D-Ala(OBn) (for **16**) or HCl·L-Ala-D-isoGln-L-Lys(Z)-D-Ala-D-Ala(OBn) (for **17**); k) H$_2$/Pd(OH)$_2$, AcOH.

tetra-, or pentapeptides (**1b**, **1d**, **1f**, or **1g**, respectively). In order to compare the biological activities of the two kinds of glycan sequence, PGN fragments with the disaccharide GlcNAc-MurNAc units, such as a disaccharide with a dipeptide (**2a**) or a tetrasaccharide with di-, tri-, or tetrapeptide (**2b**, **2c**, or **2d**, respectively) as shown in Figure 1, were also synthesized according to the previously reported methods.[11]

Human Nod2 stimulation with PGN fragments in a glycan-sequence-dependent manner

The hNod2 stimulatory activity of each synthetic PGN fragment was then evaluated by using hNod2-transfected HEK293T cells, with NF-κB activation determined by using a luciferase reporter assay, as described previously.[18] As shown in Figure 2A, the synthesized tetrasaccharides containing PGN fragments (**1b**, **1d**, **1f**, **1g**, and **2b**–**2d**) were compared with the monosac-
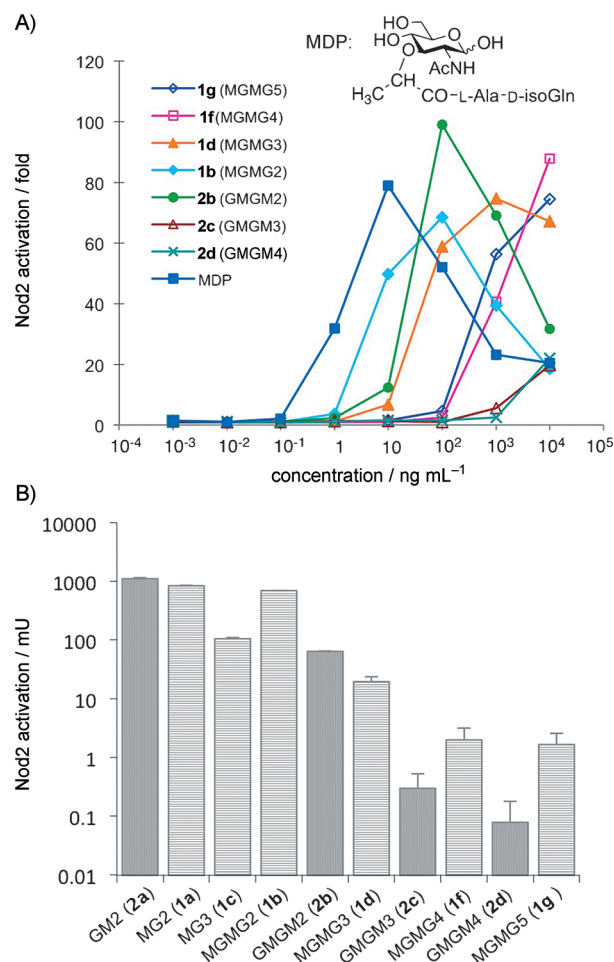

Figure 2. Stimulation of hNod2 by the synthesized PGN fragments. A) Dose-dependent hNod2 activation by the chemically synthesized tetrasaccharide containing PGN fragments, in comparison to a monosaccharide-dipeptide (MDP). B) hNod2 activation by the chemically synthesized disaccharide- and tetrasaccharide-containing PGN fragments at concentrations of 1 ng mL^{-1}. A Nod2 activation of 1000 mU is equal to the activation of MDP at 1 ng mL^{-1}. HEK293T cells were transfected with *hNOD2*, and the indicated amount of each compound was added to the cells. The ability of each compound to activate NF-κB was determined by using a luciferase reporter assay.[18]

charide dipeptide MDP in a dose-dependent manner. In both glycan sequences, the tetrasaccharide compounds with dipeptide (L-Ala-D-isoGln; **1b** and **2b**) exhibited stronger activities than the same glycan sequence groups with longer peptide chains. The dependence on peptide length was consistent with previous results.[11] Interestingly, differences in the glycan sequence caused significant differences in hNod2 activation, and the MurNAc-GlcNAc (MG)-containing sequence displayed more potent activity than the GlcNAc-MurNAc (GM) sequence. Among the dipeptide-containing fragments, MGMG2 (**1b**) showed stronger activity than GMGM2 (**2b**) at concentrations of 1 and 10 ng mL^{-1}. GMGM3 (**2c**) and GMGM4 (**2d**) exhibited only very weak activation even at higher concentrations (1000 and 10000 ng mL^{-1}), but MGMG3 (**1d**) and MGMG4 (**1f**) showed much stronger hNod2 activation.

Figure 2B, shows the hNod2 activities of the disaccharides (**1a**, **1c**, and **2a**), and also the tetrasaccharide fragments (**1b**, **1d**, **1f**, **1g**, and **2b**–**2d**) obtained at concentrations of 1 ng mL^{-1}. Among the PGN fragments, the disaccharides with dipeptides, **1a** and **2a**, showed potent activities comparable to the activation of the monosaccharide dipeptide, MDP. In these cases, it was not clear that the differences in glycan sequence affected hNod2 activation. Among the tetrasaccharide fragments, the activity clearly depended on the glycan sequence, and compounds with the MGMG sequence exhibited higher hNod2 activation than those containing the GMGM sequence, all other aspects of the peptide structures being the same. The hNod2 stimulatory abilities of the two types of glycan sequence differed significantly in the tetrasaccharide case: MGMG2 (**1b**) had an approximately tenfold higher activity than GMGM2 (**2b**); MGMG3 (**1d**) had an approximately 65-fold higher activity than GMGM3 (**2c**); and MGMG4 (**1f**) had an approximately 26-fold higher activity than GMGM4 (**2d**).

In this research, we have revealed major differences between the hNod2 innate immunostimulatory activities of peptidoglycan fragments having one of two types of glycan sequence. The MGMG (MurNAc-GlcNAc)$_2$ sequences (obtained through cleavage by *N*-acetylglucosaminidase) showed much stronger hNod2 activities than GMGM (GlcNAc-MurNAc)$_2$ sequences (cleaved by muramidase, including lysozyme), for a given peptide structure.

Conclusions

We have described the preparation of various PGN fragments containing alternating glycan disaccharide or tetrasaccharide sequences and a series of peptide chains, from di- to pentapeptides in good overall yields. The sequential glycosylation method provided an efficient approach to the divergent synthesis of PGN fragment structures with repeating glycan patterns. The chemically synthesized PGN fragments were tested for their hNod2 stimulatory activities. The results suggest that hNod2 activation by PGN fragments depends on the circumstances under which the PGN glycan is enzymatically cleaved, that is, by *N*-acetylglucosaminidase or muramidase. The results suggest that differences in the bacterial PGN degradation or construction enzymes (glycan cleaving enzymes or the peptidases) might affect immunomodulation in humans.

Experimental Section

Synthesis of the tetrasaccharide fragments: Compounds and spectroscopic data not described below are provided in the Supporting Information.

Glycosylation for the preparation of the tetrasaccharide intermediate 12 (from 10b): TMSOTf (5 μL, 0.03 mmol) was added to a mixture of the imidate **10b** (135 mg, 0.11 mmol), the acceptor **11** (200 mg, 0.18 mmol), and MS4A 4 Å molecular sieves in dry CH$_2$Cl$_2$ (20 mL) at −15 °C. After the mixture had been stirred at the same temperature for 20 min, the reaction was quenched with chilled sat. aq. NaHCO$_3$ (10 mL), and the mixture was extracted with CHCl$_3$ (50 mL). The organic layer was washed with sat. aq. NaHCO$_3$ (20 mL) and brine (20 mL), dried over Na$_2$SO$_4$, and concentrated in vacuo. The residue was purified by silica gel chromatography (30 g, toluene/EtOAc 7:1) to give **12** as a colorless solid (141 mg, 61%). ^{1}H NMR (CDCl$_3$, 500 MHz): δ = 7.45–7.18 (m, 30H), 5.85–5.82 (m, 2H), 5.41 (s, 1H), 5.24–5.12 (m, 3H), 4.88–4.59 (m, 14H), 4.50–4.10 (m, 13H), 4.02–3.87 (m, 7H), 3.77 (m, 1H), 3.67–3.49 (m, 8H), 3.40–3.38 (m, 3H), 3.25 (t, J = 10.0 Hz, 1H), 3.17 (br s, 1H), 3.06 (d, J = 10.0 Hz, 1H), 2.97–2.94 (m, 2H), 1.38 (d, J = 7.0 Hz, 3H), 1.29–1.24 ppm (m, 9H); ^{13}C NMR (CDCl$_3$, 125 MHz): δ = 175.6, 173.9, 155.9, 155.5, 154.3, 154.0, 139.5, 138.9, 138.1, 137.8, 137.6, 137.2, 133.5, 129.5, 129.1, 129.1, 128.6, 128.5, 128.4, 128.4, 128.3, 128.2, 128.1, 127.9, 127.8, 127.6, 127.4, 127.3, 127.1, 125.9, 117.9, 102.8, 101.8, 100.2, 96.6, 96.0, 95.6, 95.5, 82.4, 80.4, 78.2, 77.7, 75.0, 74.8, 74.6, 74.5, 74.3, 74.1, 74.0, 73.4, 73.9, 70.8, 68.4, 68.4, 68.0, 67.5, 65.7, 61.2, 61.1, 57.5, 57.3, 57.0, 54.7, 18.8, 18.4, 14.2, 14.1 ppm; HRMS (ESI-QTOF MS): calcd for C$_{91}$H$_{104}$Cl$_{12}$N$_4$O$_{29}$K: 2175.2686 [M+K]$^+$, found: 2175.2556.

Tetrasaccharide 13: Zn/Cu (prepared from 500 mg of Zn) was added to a solution of **12** (200 mg, 0.09 mmol) in AcOH (2 mL), and the mixture was stirred at room temperature for 30 min. The insoluble materials were filtered off, and the filtrate was concentrated in vacuo. The residual solvent was removed by coevaporation with toluene (10 mL). The residue was dissolved in pyridine (2 mL) and acetic anhydride (2 mL), and the solution was stirred at room temperature for 1 h. The reagents were removed by concentration with toluene (10 mL). The residue was purified by silica gel chromatography (50 g, CHCl$_3$/acetone 9:1) to give 2-*N*-acetyl tetrasaccharide as a white solid (93 mg, 62%). ^{1}H NMR (CDCl$_3$, 400 MHz): δ = 7.44–7.21 (m, 31H), 6.98 (m, 1H), 6.74 (d, J = 6.8 Hz, 1H), 6.07 (d, J = 8.4 Hz, 1H), 5.81 (m, 1H), 5.47 (s, 1H), 5.24–5.14 (m, 3H), 4.95 (d, J = 12.8 Hz, 1H), 4.84 (d, J = 3.6 Hz, 1H), 4.74–4.64 (m, 5H), 4.55–4.45 (m, 5H), 4.36–4.02 (m, 14H), 3.96–3.38 (m, 20H), 3.23 (t, J = 9.6 Hz, 1H), 3.13 (m, 1H), 3.05 (m, 1H), 1.98 (s, 3H), 1.94 (s, 3H), 1.88 (s, 3H), 1.78 (s, 3H), 1.37 (d, J = 6.8 Hz, 3H), 1.31–1.21 ppm (m, 9H); ^{13}C NMR (CDCl$_3$, 125 MHz): δ = 175.9, 174.7, 172.4, 172.3, 171.8, 169.8, 139.6, 138.8, 138.3, 138.1, 138.1, 138.0, 137.2, 133.7, 129.1, 128.5, 128.3, 128.3, 128.2, 128.2, 128.0, 127.9, 127.8, 127.7, 127.4, 127.3, 127.2, 125.9, 117.5, 102.4, 101.1, 100.8, 100.5, 96.3, 95.5, 82.6, 80.3, 76.4, 75.8, 75.4, 74.9, 74.6, 74.1, 73.8, 73.4, 73.4, 73.2, 72.8, 71.1, 68.5, 68.5, 68.2, 68.0, 65.9, 61.3, 61.0, 55.4, 55.0, 53.4, 52.1, 29.7, 29.3, 23.9, 23.6, 23.5, 23.4, 23.1, 18.8, 18.6, 14.2, 14.1 ppm; HRMS (ESI-QTOF MS): calcd for C$_{87}$H$_{108}$N$_4$O$_{25}$Na: 1631.7200 [M+Na]$^+$, found: 1631.7169.

LiOH (16 mg, 0.7 mmol) was added to a solution of 2-*N*-acetyl tetrasaccharide (93 mg, 0.06 mmol) in dioxane/THF/H$_2$O (2:4:1,

4.0 mL), and the mixture was stirred at room temperature for 1 h. The solution was neutralized with Dowex H$^+$ (Dowex 50W$\times$8, 200–400 mesh H form, Dow Chemicals), then applied to an HP-20 column (2 cm$\times$10 cm). Organic and inorganic salts were removed by elution with H$_2$O (160 mL), followed by elution with MeOH and concentration in vacuo to give a tetrasaccharide with a free lactic acid moiety, **13**, as a white solid (89 mg, quant). ^{1}H NMR (CD$_3$OD, 400 MHz): δ = 7.34–6.99 (m, 30 H), 5.83 (m, 1 H), 5.40 (s, 1 H), 5.20 (dd, J = 17.6, 1.6 Hz, 1 H), 5.09–5.00 (m, 2 H), 4.84 (d, J = 11.2 Hz, 1 H), 4.68 (d, J = 3.6 Hz, 1 H), 4.60–4.37 (m, 8 H), 4.25–4.21 (m, 3 H), 4.09–4.00 (m, 2 H), 3.95–3.31 (m, 22 H), 3.04–2.95 (m, 2 H), 1.89 (m, 6 H), 1.77 (s, 6 H), 1.31–1.19 ppm (m, 6 H); HRMS (ESI-QTOF MS): calcd for C$_{83}$H$_{100}$N$_4$O$_{25}$Na: 1575.6574 [M+Na]$^+$, found: 1575.6508.

Tetrasaccharide dipeptide backbone 14: WSCD·HCl (8 mg, 0.048 mmol) and triethylamine (15 μL, 0.096 mmol) were added at 0 °C to a solution of **13** (25 mg, 0.016 mmol), HCl·L-Ala-D-isoGln-OBn (22 mg, 0.064 mmol), and HOBt (7 mg, 0.048 mmol) in DMF (3 mL), and the mixture was stirred at RT overnight. The mixture was concentrated, and the residue was dissolved in CHCl$_3$. The solution was washed with citric acid (1 M, 20 mL), H$_2$O (20 mL), sat. aq. NaHCO$_3$ (20 mL), and brine (20 mL). The organic layer was dried over Na$_2$SO$_4$, and concentrated in vacuo. The residue was purified by silica gel chromatography (5 g, CHCl$_3$/MeOH 15:1) to give **14** as a white solid (28 mg, 82 %). ^{1}H NMR ([D$_7$]DMF, 400 MHz): δ = 8.12–8.04 (m, 3 H), 7.85 (d, J = 9.2 Hz, 1 H), 7.68–7.00 (m, 48 H), 5.90 (m, 1 H), 5.70 (s, 1 H), 5.14 (dd, J = 17.2, 1.9 Hz, 1 H), 5.08–5.17 (m, 6 H), 4.95 (d, J = 11.2 Hz, 1 H), 4.84 (d, J = 8.4 Hz, 1 H), 4.76 (d, J = 3.7 Hz, 1 H), 4.64–4.29 (m, 15 H), 4.20–3.57 (m, 24 H), 3.25–3.15 (m, 3 H), 2.49 (t, J = 8.0 Hz, 4 H), 2.31–2.20 (m, 2 H), 1.86–1.77 (m, 14 H), 1.29–1.20 ppm (m, 12 H); ^{13}C NMR ([D$_7$]DMF, 100 MHz): δ = 173.9, 173.9, 173.3, 173.2, 173.1, 173.0, 170.8, 170.8, 170.1, 140.7, 140.4, 139.8, 139.7, 139.6, 138.7, 137.3, 137.3, 135.3, 129.4, 129.1, 128.9, 128.7, 128.6, 138.5, 128.2, 128.2, 128.1, 128.0, 127.9, 127.8, 127.4, 126,7, 116,9, 101.8, 101.3, 97.5, 92.9, 81.6, 80.2, 80.0, 79.8, 79.6, 77.2, 75.3, 74.5, 73.0, 72.9, 71.0, 69.2, 68.8, 68.5, 66.7, 66.3, 53.3, 52.8, 50.0, 49.6, 36.0, 31.0, 28.0, 23.7, 23.6, 23.4, 22.8, 19.4, 19.2, 18.3, 17.7 ppm; HRMS (ESI-QTOF MS): calcd for C$_{113}$H$_{138}$N$_{10}$O$_{31}$Na$_2$: 1088.4662 [M+2 Na]$^{2+}$, found: 1088.4655.

Tetrasaccharide tetrapeptide backbone 16: HATU (11 mg, 0.019 mmol) and triethylamine (8 μL, 0.039 mmol) were added at 0 °C to a solution of **13** (15 mg, 0.006 mmol) and HCl·L-Ala-D-isoGln-L-Lys(Z)-D-Ala-OBn (19 mg, 0.019 mmol) in DMF (3 mL), and the mixture was stirred at RT overnight. The mixture was concentrated, and the residue was dissolved in CHCl$_3$. The solution was washed with citric acid (1 M, 20 mL), H$_2$O (20 mL), sat. aq. NaHCO$_3$ (20 mL), and brine (20 mL). The organic layer was dried over Na$_2$SO$_4$ and concentrated in vacuo. The residue was purified by silica gel chromatography (5 g, CHCl$_3$/MeOH 15:1) to give **16** as a white solid (23 mg, 81 %). ^{1}H NMR (CDCl$_3$/CD$_3$OD 4:1, 500 MHz): δ = 7.41–7.21 (m, 50 H), 5.85 (m, 1 H), 5.43 (s, 1 H), 5.25–5.06 (m, 8 H), 4.93–4.81 (m, 2 H), 4.68–4.25 (m, 16 H), 4.13–3.89 (m, 11 H), 3.47–3.36 (m, 11 H), 3.12 (m, 4 H), 2.35–2.18 (m, 6 H), 2.02–1.79 (m, 14 H), 1.48–1.10 ppm (m, 30 H); HRMS (ESI-QTOF MS): calcd for C$_{147}$H$_{184}$N$_{16}$O$_{39}$Na$_2$: 1421.6351 [M+2 Na]$^{2+}$, found: 1421.6357.

Deprotection of the tetrasaccharide dipeptide 1b: Palladium hydroxide (55 mg) in AcOH was added to a solution of **14** (28 mg, 0.013 mmol) in AcOH (1 mL), and the mixture was stirred under H$_2$ (2 MPa) for one day. The reaction was monitored by TLC, and the hydrogenolysis was continued until deprotection was complete. The Pd catalyst was filtered off by using celite, and the filtrate was concentrated. The residue was lyophilized from acetonitrile/H$_2$O to give **1b** (16.6 mg, 89 %) as a white solid. ^{1}H NMR (CD$_3$OD,

500 MHz): δ = 4.81–4.25 (m, 9 H), 3.87–3.29 (m, 26 H), 2.38 (br s, 4 H), 2.19 (br s, 2 H), 1.97 (m, 14 H), 1.60–1.58 (m, 2 H), 1.43–1.34 (m, 12 H), 0.95–0.91 ppm (dd, J = 12.0, 7.0 Hz, 3 H); ^{13}C NMR (CD$_3$OD, 125 MHz): δ = 176.3, 175.9, 175.6, 175.2, 175.2, 173.9, 173.9, 173.8, 173.5, 103.3, 103.1, 101.8, 98.2, 83.4, 81.9, 81.6, 80.0, 78.7, 78.4, 78.0, 77.4, 76.8, 76.0, 73.9, 72.2, 71.2, 70.9, 70.8, 62.5, 62.0, 61.6, 61.4, 57.0, 56.4, 55.9, 55.1, 53.9, 53.8, 50.7, 49.6, 31.6, 28.8, 28.4, 23.7, 23.4, 22.5, 19.5, 18.9, 17.7, 17.5, 11.0 ppm; HRMS (ESI-QTOF MS): calcd for C$_{57}$H$_{94}$N$_{10}$O$_{31}$Na: 1437.5984 [M+Na]$^+$, found: 1437.6008.

1c–1g were similarly synthesized. **1c:** ^{1}H NMR (D$_2$O, 500 MHz): δ = 5.11 (d, J = 2.6 Hz, 1 H), 4.49 (d, J = 8.5 Hz, 1 H), 4.27 (dd, J = 9.5, 4.5 Hz, 1 H), 4.22–4.13 (m, 2 H), 3.87–3.44 (m, 14 H), 3.32–3.26 (m, 3 H), 2.47–2.32 (m, 2 H), 2.15–1.88 (m, 8 H), 1.36–1.29 (m, 7 H), 1.11–1.00 ppm (m, 5 H); ^{13}C NMR (D$_2$O, 150 MHz): δ = 178.9, 175.9, 175.6, 175.1, 174.5, 174.2, 174.0, 101.3, 94.8, 90.4, 82.5, 79.7, 78.2, 75.6, 78.2, 75.6, 74.6, 72.5, 70.0, 69.3, 68.7, 61.3, 60.6, 60.1, 54.9, 53.6, 49.7, 39.3, 31.9, 31.1, 27.0, 26.3, 27.0, 26.3, 22.2, 22.1, 21.9, 18.8, 16.6 ppm; HRMS (ESI-QTOF): calcd for C$_{33}$H$_{57}$N$_7$O$_{17}$Na: 846.3709 [M+Na]$^+$, found: 846.3697.

1d: ^{1}H NMR (D$_2$O, 400 MHz): δ = 4.40–4.35 (m, 2 H), 4.26 (m, 1 H), 4.10–4.03 (m, 5 H), 3.79–3.20 (m, 29 H), 2.84 (t, J = 7.0 Hz, 4 H), 2.26–2.24 (m, 4 H), 2.07–2.03 (m, 2 H), 1.90–1.80 (m, 14 H), 1.66 (m, 2 H), 1.58–1.51 (m, 6 H), 1.43–1.39 (m, 2 H), 1.31–1.21 (m, 16 H), 0.74 ppm (t, J = 7.4 Hz, 3 H); HRMS (ESI-QTOF MS): calcd for C$_{69}$H$_{119}$N$_{14}$O$_{33}$K: 855.3851 [M+H+K]$^{2+}$, found: 855.3832.

1e: ^{1}H NMR (D$_2$O, 500 MHz): δ = 5.06 (d, J = 3.0 Hz, 1 H), 4.44 (d, J = 8.0 Hz, 1 H), 4.20–4.02 (m, 5 H), 3.82–3.38 (m, 12 H), 2.86 (t, J = 7.5 Hz, 2 H), 2.31–2.28 (m, 2 H), 2.04 (m, 1 H), 1.91 (s, 3 H), 1.86–1.82 (m, 4 H), 1.74–1.62 (m, 4 H), 1.32–1.30 (m, 5 H), 1.25 (d, J = 6.3 Hz, 3 H), 1.19 ppm (d, J = 7.3 Hz, 3 H); HRMS (ESI-QTOF): calcd for C$_{36}$H$_{63}$N$_8$O$_{18}$: 895.4260 [M+H]$^+$, found: 895.4213.

1f: ^{1}H NMR (D$_2$O, 400 MHz): δ = 4.39–4.37 (m, 3 H), 4.25–4.22 (m, 2 H), 4.17–4.10 (m, 8 H), 3.76–3.15 (m, 26 H), 2.85 (t, J = 7.2 Hz, 4 H), 2.28–2.26 (m, 4 H), 2.07–2.03 (m, 2 H), 1.90–1.80 (m, 14 H), 1.63–1.52 (m, 10 H), 1.44–1.42 (m, 2 H), 1.31–1.17 (m, 20 H), 0.74 (t, J = 7.4 Hz, 3 H); HRMS (ESI-QTOF MS): calcd for C$_{75}$H$_{129}$N$_{16}$O$_{35}$Na: 918.4352 [M+H+Na]$^{2+}$, found: 918.4310.

1g: ^{1}H NMR (D$_2$O, 400 MHz): δ = 4.43–4.38 (m, 2 H), 4.30–4.06 (m, 11 H), 3.88–3.25 (m, 28 H), 2.87 (t, J = 7.6 Hz, 4 H), 2.29–2.26 (m, 4 H), 2.09–2.01 (m, 2 H), 1.92–1.82 (m, 14 H), 1.67–1.53 (m, 8 H), 1.48–1.41 (m, 2 H), 1.33–1.24 (m, 28 H), 0.77 (t, J = 7.6 Hz, 3 H); HRMS (ESI-QTOF MS): calcd for C$_{81}$H$_{140}$N$_{18}$O$_{37}$: 978.4813 [M+2 H]$^{2+}$, found: 978.4783.

HEK293T bioassay for human Nod2 activity: Ligand-dependent NF-κB activation was determined by using 0.5×10^5 HEK293T cells transfected with expression plasmids of Nod2 (33 ng of pMX2-HANod2) in the presence of reporter plasmids, NF-κB-dependent pBxIV-luc, and control pEF1BOS-β-Gal, as described.[3] Briefly, HEK293T cells were transfected with expression plasmids by the calcium phosphate method. Eight hours post-transfection, the cells were treated with medium containing a synthesized ligand. 24 hours post-transfection, ligand-dependent NF-κB activation was determined by using a luciferase reporter assay.[18]

Acknowledgements

[1] N. Inohara, Y. Ogura, A. Fontalba, O. Gutierrez, F. Pons, J. Crespo, K. Fukase, S. Inamura, S. Kusumoto, M. Hashimoto, S. J. Foster, A. P. Moran, J. L. Fernandez-Luna, G. Nuñez, *J. Biol. Chem.* **2003**, *278*, 5509–5512.
[2] S. E. Girardin, I. G. Boneca, J. Viala, M. Chamaillard, A. Labigne, G. Thomas, D. J. Philpott, P. J. Sansonetti, *J. Biol. Chem.* **2003**, *278*, 8869–8872.
[3] M. Chamaillard, M. Hashimoto, Y. Horie, J. Masumoto, S. Qiu, L. Saab, Y. Ogura, A. Kawasaki, K. Fukase, S. Kusumoto, M. A. Valvano, S. J. Foster, T. W. Mak, G. Nuñez, N. Inohara, *Nat. Immunol.* **2003**, *4*, 702–707.
[4] S. E. Girardin, I. G. Boneca, L. A. Carneiro, A. Antignac, M. Jehanno, J. Viala, K. Tedin, M. K. Taha, A. Labigne, U. Zahringer, A. J. Coyle, P. S. DiStefano, J. Bertin, P. J. Sansonetti, D. J. Philpott, *Science* **2003**, *300*, 1584–1587.
[5] Y. Fujimoto, A. R. Pradipta, N. Inohara, K. Fukase, *Nat. Prod. Rep.* **2012**, *29*, 568–579.
[6] Y. Fujimoto, S. Inamura, A. Kawasaki, Z. Shiokawa, A. Shimoyama, T. Hashimoto, S. Kusumoto, K. Fukase, *J. Endotoxin Res.* **2007**, *13*, 189–196.
[7] Y. Ogura, D. K. Bonen, N. Inohara, D. L. Nicolae, F. F. Chen, R. Ramos, H. Britton, T. Moran, R. Karaliuskas, R. H. Duerr, J. P. Achkar, S. R. Brant, T. M. Bayless, B. S. Kirschner, S. B. Hanauer, G. Nuñez, J. H. Cho, *Nature* **2001**, *411*, 603–606.
[8] J. P. Hugot, M. Chamaillard, H. Zouali, S. Lesage, J. P. Cezard, J. Belaiche, S. Almer, C. Tysk, C. A. O'Morain, M. Gassull, V. Binder, Y. Finkel, A. Cortot, R. Modigliani, P. Laurent-Puig, C. Gower-Rousseau, J. Macry, J. F. Colombel, M. Sahbatou, G. Thomas, *Nature* **2001**, *411*, 599–603.
[9] C. Vignal, E. Singer, L. Peyrin-Biroulet, P. Desreumaux, M. Chamaillard, *Microbes Infect.* **2007**, *9*, 658–663.
[10] S. Inamura, K. Fukase, S. Kusumoto, *Tetrahedron Lett.* **2001**, *42*, 7613–7616.
[11] S. Inamura, Y. Fujimoto, A. Kawasaki, Z. Shiokawa, E. Woelk, H. Heine, B. Lindner, N. Inohara, S. Kusumoto, K. Fukase, *Org. Biomol. Chem.* **2006**, *4*, 232–242.
[12] Y. Fujimoto, Y. Konishi, O. Kubo, M. Hasegawa, N. Inohara, K. Fukase, *Tetrahedron Lett.* **2009**, *50*, 3631–3634.
[13] A. Kawasaki, Y. Karasudani, Y. Otsuka, M. Hasegawa, N. Inohara, Y. Fujimoto, K. Fukase, *Chem. Eur. J.* **2008**, *14*, 10318–10330.
[14] S. Kusumoto, K. Yamamoto, M. Imoto, M. Inage, M. Tsujimoto, S. Kotani, T. Shiba, *Bull. Chem. Soc. Jpn.* **1986**, *59*, 1411–1417.
[15] R. R. Schmidt, J. Michel, *Angew. Chem.* **1980**, *92*, 763–764; *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 731–732.
[16] B. Yu, H. Tao, *Tetrahedron Lett.* **2001**, *42*, 2405–2407.
[17] M. Oikawa, W. C. Liu, Y. Nakai, S. Koshida, K. Fukase, S. Kusumoto, *Synlett* **1996**, 1179–1180.
[18] M. Hasegawa, K. Yang, M. Hashimoto, J. H. Park, Y. G. Kim, Y. Fujimoto, G. Nuñez, K. Fukase, N. Inohara, *J. Biol. Chem.* **2006**, *281*, 29054–29063.