Akkermansia muciniphila phospholipid induces homeostatic immune responses

Multiple studies have established associations between human gut microbes and host physiology, but determining the molecular mechanisms underlying these associations has been challenging. Akkermansia muciniphila has been robustly associated with positive systemic effects on host metabolism, favourable outcomes to checkpoint blockade in cancer immunotherapy and homeostatic immunity. Here we report the identification of a lipid from A. muciniphila's cell membrane that recapitulates the immunomodulatory activity of A. muciniphila in cell-based assays. The isolated immunogen, a diacyl phosphatidylethanolamine with two branched chains (a15:0-i15:0 PE), was characterized through both spectroscopic analysis and chemical synthesis. The immunogenic activity of a15:0-i15:0 PE has a highly restricted structure–activity relationship, and its immune signalling requires an unexpected toll-like receptor TLR2–TLR1 heterodimer. Certain features of the phospholipid's activity are worth noting; it is significantly less potent than known natural and synthetic TLR2 agonists; it preferentially induces some inflammatory cytokines but not others; and, at low doses (1% of EC50) it resets activation thresholds and responses for immune signalling. Identifying both the molecule and an equipotent synthetic analogue, its non-canonical TLR2–TLR1 signalling pathway, its immunomodulatory selectivity and its low-dose immunoregulatory effects provide a molecular mechanism for a model of A. muciniphila's ability to set immunological tone and its varied roles in health and disease.
nuclear magnetic resonance (NMR) analysis identified a phosphatidylcholine (PC), the dominant membrane phospholipid in most bacteria. PEs have a glycerol core, a polar phosphoethanolamine head group at the sn-3 position, and two fatty-acid (FA) esters attached to the sn-1 and sn-2 positions (Fig. 1c). Additional NMR analysis revealed that both chains had methyl branches (Fig. 1c and Extended Data Fig. 1). One acyl chain had a terminal iso branch, and the other had a terminal anteiso branch, meaning that the methyl groups were on positions 12 and 13 of a 14-carbon FA chain (Fig. 1c). The order of the acyl groups was determined by selective hydrolysis to preferentially liberate the FA attached at the sn-2 position.

The active molecule’s chemical name is 12-methyltetradecanoyl-13-methyltetradecanoyl-sn-glycero-3-phosphoethanolamine, which is a15:0-i15:0 PE in standard lipid nomenclature. We did not find producers of a15:0-i15:0 PE in frequently encountered gut microbes nor in gut microbiomes with reported immunomodulatory effects. Membrane lipids reflect both evolutionary history and current environment. A. muciniphila, the only member of verrucomicrobia in the gut microbiota, is a phylogenetic outlier specialized for life in the mucin layer. Metabolomic analysis, phylogenetic placement and a distinctive microenvironment all support a singular association of A. muciniphila with a15:0-i15:0 PE.
The active fraction contained all the PEs produced by A. muciniphila: the later eluting fractions were triglycerides and the earlier eluting fractions were diacylglycerides with different head groups. The PE FAs were dominated (92%) by relatively short, branched-chain fatty acids (BCFAs): a15:0 (52%), i15:0 (24%) and i14:0 (16%) (Fig. 1d). Small amounts of a17:0 and i16:0 were also present. Bacteria make BCFAs to increase membrane fluidity, the same function unsaturated FAs have in animals. Anti-inflammatory activity has also been attributed to some BCFAs: a15:0 and i15:0 FAs are preferentially incorporated into membrane lipids, leading to increased fluidity17,18. BCFAs have been associated with human health, especially an anticorrelation with developing type 2 diabetes18,19.

Interestingly, BCFAs in human serum, independent of a connection with A. muciniphila or any other bacteria, have been strongly associated with human health, especially an anticorrelation with developing type 2 diabetes18,19.

The active compound (a15:0-i15:0 PE) was the major component (approximately 50%) of A. muciniphila’s lipid membrane and had a robust dose–response curve for induction of TNFα (Fig. 1e). In addition to dramatically upregulating TNFα release, it promoted the release of IL-6 (interleukin 6), but not IL-10 or IL-12p70 (Fig. 1f and Extended Data Fig. 2). Dendritic cells typically respond to bacterial metabolites through the pathogen-associated molecular pattern (PAMP) receptors, toll-like receptor 2 (TLR2) and toll-like receptor 4 (TLR4)20,21. Receptor specificity was established by generating mBMDCs from both tlr2−/− and tlr4−/− mice and using them along with wild-type cells in the cytokine induction assay22. PEs active in wild-type cell assays produced no TNFα induction in mBMDCs from tlr2−/− mice but showed robust TNFα induction in mBMDCs generated from tlr4−/− mice (Fig. 1g). Previous reports attributed A. muciniphila’s immunomodulatory activity to a membrane-associated protein (Amuc_1100) signalling through TLR222−24. Additional publications supporting the protein’s role in maintaining the intestinal mucosal barrier have also appeared23. Even though a15:0-i15:0 PE was the only active molecule detected in our study, other immunomodulatory contributors are a distinct possibility24,25.

With the identification of a15:0-i15:0 PE as an immunomodulatory molecule and TLR2 as its cognate receptor, we examined the pathway to determine the signal. It is well established that PE biosynthesis has three repetitive addition of two-carbon units25. After elongation, the BCFAs are deprotected, and the acyl groups were added in a stepwise fashion, with a commercially available protected chiral glycerol. With the future sn-1 position to install the hydroxyl groups at sn-2 and sn-3 were deprotected, and the acyl groups were added in a stepwise fashion, taking advantage of the greater reactivity at the sn-1 position to install this acyl chain first. The a15:0 carboxylic acid used in the synthesis had the stereochemistry appropriate for natural Ile. The synthetic a15:0-i15:0 PE had identical spectroscopic, chromatographic and biological properties to the natural molecule (Figs. 1e and 2d,e).

In addition to confirming the order of the acyl chains and two stereochemical issues, the synthetic scheme allowed a small library of natural and synthetic FAs and diacyl PEs to be assembled. The library was assayed to establish a preliminary SAR for the A. muciniphila lipids and their component parts. First, we established that FAs only activate TLR2 in a TLR2-dependent fashion as measured by ELISA. Pam3CSK4 and LPS were used as controls. Data are presented as mean values ± s.d. of technical replicates (n = 4). Unpaired t-test with two-tailed P value; ****P < 0.0001. With the identification of a15:0-i15:0 PE as an immunomodulatory molecule and TLR2 as its cognate receptor, we examined the pathway to determine the signal. It is well established that PE biosynthesis has three repetitive addition of two-carbon units25. After elongation, the BCFAs are deprotected, and the acyl groups were added in a stepwise fashion, with a commercially available protected chiral glycerol. With the future sn-1 position to install the hydroxyl groups at sn-2 and sn-3 were deprotected, and the acyl groups were added in a stepwise fashion, taking advantage of the greater reactivity at the sn-1 position to install this acyl chain first. The a15:0 carboxylic acid used in the synthesis had the stereochemistry appropriate for natural Ile. The synthetic a15:0-i15:0 PE had identical spectroscopic, chromatographic and biological properties to the natural molecule (Figs. 1e and 2d,e).

In addition to confirming the order of the acyl chains and two stereochemical issues, the synthetic scheme allowed a small library of natural and synthetic PEs and diacyl PEs to be assembled. The library was assayed to establish a preliminary SAR for the A. muciniphila lipids and their component parts. First, we established that FAs only activate TLR2 in the context of a diacyl PE, as none has any detectable activity on its own (Extended Data Fig. 5). This result is consistent with the tlr2−/− mBMDM analysis (Fig. 1g). The diacyl PE library members revealed a surprisingly strict set of structural requirements (Fig. 2e): (1) methyl branches are essential for TNFα induction, as the three PE analogues where both acyl groups have straight chains (n14:0, n15:0 and n16:0) had no detectable activity; (2) the two acyl chains must be different, as a15:0-a15:0 PE and i15:0-i15:0 PE had no detectable activity; and (3) positional order appears to play a minor role, as a15:0-i15:0 PE and i15:0-a15:0 PE were essentially equipotent. We did not detect i15:0-a15:0 PE in natural samples.
SAR studies on TLR2 ligands invariably focus on the head group that protrudes from the membrane-bound receptor. The conventional view of TLR2 signalling relates the lipid chains to providing hydrophobic anchors for a protruding head group that regulates receptor activation\(^\text{6,28,29}\). This view is supported by several structural studies on TLR2 receptors with bound ligand and SAR studies\(^\text{26,28,29}\). The extracellular part of TLR2 is a horseshoe-shaped, leucine-rich repeat with a long hydrophobic tunnel that binds two acyl chains (Fig. 3a,b). TLR2 typically requires formation of a heterodimer with either TLR1 or TLR6 for immune signalling\(^\text{5,26,29}\). CRISPR–Cas knockdowns of TLR6 and TLR1 showed that a TLR2–TLR1 heterodimer is required for TNFα induction, which is a surprising result for a diacyl lipid (Fig. 3a,c)\(^\text{5,26,29}\). The requirement for a non-canonical TLR2–TLR1 heterodimer indicates that the two acyl chains of a15:0-i15:0 PE occupy binding pockets in two different proteins, one in TLR2 and one in TLR1, forming an atypical signalling heterodimer with a buried head group (Fig. 3b and Supplementary Video 1). There are other TLR2 agonists that form TLR2–TLR1 heterodimers, and at least two of them (the synthetic molecules diprovocim and CU-T12-9) were developed as adjuvants for cancer immunotherapy, whereas another (polysaccharide A from Bacteroides fragilis\(^\text{24}\)) is produced by a member of the gut microbiome and associated with IL-10 production\(^\text{30–33}\). There is an important difference in potency between the synthetic agonists and a15:0-i15:0 PE: EC\text{50} values of pmol l\(^{-1}\) versus µmol l\(^{-1}\). A similar difference in immunogenicity has been noted in a study of immunomodulatory sphingolipids from B. fragilis\(^\text{24}\).

To complete this initial phase of our study, we sought to connect the active lipids from A. muciniphila to the selective cytokine responses of specific human immune cell lineages\(^\text{1}\). Human monocytes purified from peripheral blood were cultured and stimulated with natural and synthetic TLR2 agonists for 6 h, after which mRNA was extracted from peripheral blood were cultured and stimulated with natural and synthetic TLR2 agonists for 6 h, after which mRNA was extracted and sequenced (Extended Data Fig. 6). The a15:0-i15:0 PE induced pro-inflammatory cytokines such as TNFα and IL-6 comparably with lipopolysaccharide (LPS) and Pam3CSK4, albeit at higher doses, but
Next, we investigated the effects of a15:0-i15:0 PE on other immunogens by treating human monocyte-derived dendritic cells (which were differentiated from granulocyte-macrophage colony-stimulating factor (GM-CSF) and IL-4) with a15:0-i15:0 PE and the model agonists Pam3CSK4 (to stimulate TLR2–TLR1) and LPS (to stimulate TLR4). Variables were dose, duration (6 h or 21 h) and timing (cotreatment or sequential treatment). A particularly informative set of data came from the 21 h sequential treatment study in which a low dose of a15:0-i15:0 PE (0.15 μmol l⁻¹, approximately 1% of EC₅₀) was followed 18 h later by the addition of known agonists. This treatment regimen completely suppressed TNFα release (Fig. 3e and Extended Data Fig. 7). The suppressive effect was not seen with shorter periods between the lipid and agonist treatments (6 h and co-stimulation studies) nor with higher doses of lipid (Fig. 3f,g and Extended Data Fig. 7). These results support a model in which low doses of a15:0-i15:0 PE and delayed stimulation reset the cellular activation threshold and moderate other cellular immune responses, as indicated by the LPS response, which was reduced but not in a lipid-dependent fashion (Fig. 3e). Low dose and delayed stimulation reflect likely in vivo conditions. Larger doses and shorter times produce the expected dose-dependent response (Figs. 1e and 3f,g and Extended Data Fig. 7).

Conclusion

Since its discovery, multiple lines of investigation have indicated that A. muciniphila plays a considerable role in regulating human immune responses in a variety of contexts37,41. Our study indicates that A. muciniphila’s immunomodulatory activity can be replicated by a diacyl PE, a15:0-i15:0 PE, a lipid that is not noticeably different from other diacyl PEs forming the cell membranes of most bacteria found in the human gut42. Because of its generic structure, its remarkable activity would not have been easily identified by genomic or metabolomic analyses. It agonizes a non-canonical TLR2–TLR1 heterodimer to release a subset of inflammatory cytokines26–35. The potency of TLR2 heterodimers is conventionally thought to be governed by a peptide, peptide-like or (poly)saccharide moiety emerging from the dimer interface, and the absence of this chain in a15:0-i15:0 PE might be responsible for the molecule’s unusual immunomodulatory effects (Figs. 1e and 3b)26–30,32. Although there is still much to be learned about the pharmacology of a15:0-i15:0 PE, the existing data support a model in which repeated low-level stimulation of the TLR2–TLR1 signalling pathway resets the activation threshold so that weak signals are ignored and strong signals are moderated, thereby contributing to homeostatic immunity38,39. It is also important to note that the data underlying the model are from in vitro studies and in vivo studies will be needed to fully validate it. Overall, this study describes the molecular mechanism of a druggable pathway that recapitulates the dimer interface, and the absence of this chain in a15:0-i15:0 PE, a lipid that is not noticeably different from other diacyl PEs forming the cell membranes of most bacteria found in the human gut42. Because of its generic structure, its remarkable activity would not have been easily identified by genomic or metabolomic analyses. It agonizes a non-canonical TLR2–TLR1 heterodimer to release a subset of inflammatory cytokines26–35. The potency of TLR2 heterodimers is conventionally thought to be governed by a peptide, peptide-like or (poly)saccharide moiety emerging from the dimer interface, and the absence of this chain in a15:0-i15:0 PE might be responsible for the molecule’s unusual immunomodulatory effects (Figs. 1e and 3b)26–30,32. Although there is still much to be learned about the pharmacology of a15:0-i15:0 PE, the existing data support a model in which repeated low-level stimulation of the TLR2–TLR1 signalling pathway resets the activation threshold so that weak signals are ignored and strong signals are moderated, thereby contributing to homeostatic immunity38,39. It is also important to note that the data underlying the model are from in vitro studies and in vivo studies will be needed to fully validate it. Overall, this study describes the molecular mechanism of a druggable pathway that recapitulates in cellular assays the immunomodulatory effects associated with a prominent member of the gut microbiota.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-04985-7.
41. Picelli, S. Single cell methods, sequencing and proteomics. Methods Mol. Biol. 1979, 25–44 (2019).
42. Wingett, S. W. & Andrews, S. FastQ Screen: a tool for multi-genome mapping and quality control. F1000Research 7, 1338 (2018).
43. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32, 3047–3048 (2016).
44. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010).
45. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34, 525–527 (2016).
46. Long, F. et al. AceDRG: a stereochemical description generator for ligands. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2126–2132 (2004).
47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. 73, 112–122 (2017).
48. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. 27, 14–25 (2018).
49. Morin, A. et al. Collaboration gets the most out of software. eLife 2, e01456 (2013).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022
Bioassay-guided fractionation, purification and identification of \textit{a15:0-i15:0 PE}

The crude extract from the cell pellets (4 g) was dissolved in chloroform and fractionated by normal-phase chromatography using seven different solvent systems (A, 100% hexane; B, 100% chloroform; C, 100% ethyl acetate; D, 75% ethyl acetate/25% methanol; E, 90% acetone/10% methanol; F, 50% methanol/50% dichloromethane; and G, 100% methanol) with a silica column (Teledyne Isco, RediSep RF Gold Silica 12 g). The pro-inflammatory activity was highly detected in fractions F and G. The mixture of fractions F and G (120 mg and 210 mg, respectively, and 8.3% of total yield) was then subjected to reversed-phase semi-preparative high-performance liquid chromatography (HPLC) (Luna C8 (2), 250 × 10 mm, 5 µm) using the following gradient solvent system: 10% methanol/90% water isocratic for 10 min; gradient to 30% methanol/70% water for 10 min; then 30% methanol/70% water to 90% methanol/10% water for 20 min, 90% methanol isocratic for 10 min, gradient to 100% methanol for 25 min; flow rate, 2 ml/min). Fractions were collected every 1 min between 5 min and 75 min, generating 70 fractions. Fractions able to stimulate pro-inflammatory cytokine production from mBMDCs were combined and identified as bacterial PE with BCFAs (22 mg, yield = 0.35%). An essentially pure compound, later identified as \textit{a15:0-i15:0 PE}, was acquired at a retention time of 63 min (14 mg, yield = 0.35%).

The crude extract from the supernatants (15 g) was dissolved in methanol and filtered through a syringe filter (polytetrafluoroethylene (PTFE), 0.2 µm). The filtered extract was directly injected onto a reversed-phase preparative HPLC column (Luna C8 (2), 250 × 21.2 mm, 5 µm) with a gradient mobile solution (30% methanol/70% water/100% methanol for 30 min, 100% methanol isocratic for 30 min, 100% methanol for 30 min, 100% methanol isocratic for 30 min; flow rate, 10 ml/min). Fractions were collected every 2 min from 5 min to 55 min, generating 25 fractions. Fractions able to stimulate pro-inflammatory cytokine production from mBMDCs were combined and identified as bacterial PE with BCFAs (22 mg, yield = 0.35%). An essentially pure compound, later identified as \textit{a15:0-i15:0 PE}, was acquired at a retention time of 63 min (14 mg, yield = 0.35%).

NMR spectroscopy

All 1H NMR spectra were acquired at 500 MHz at 30 °C, and chemical shifts are represented on a δ (beta) scale. Residual proton in the NMR solvent (CDCl₃, δ 7.26) was used to reference chemical shifts. Data are represented as follows: assignment, chemical shift, integration, multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad) and coupling constant in hertz. All 13C NMR spectra were obtained at 125 MHz at 30 °C and chemical shifts are represented on a δ scale. The carbon resonances of the NMR solvent (CDCl₃, δ 77.17) were used to reference chemical shifts. Full assignment of protons and carbons were completed on the basis of the following 2D NMR spectroscopy experiments: gradient 1H–1H correlation spectroscopy, gradient 1H–13C heteronuclear single quantum coherence, gradient 1H–13C heteronuclear multiple bond connectivity. Mova v.14.2.0 was used to analyse NMR data of natural and synthetic compounds.

High-resolution mass spectrometry for \textit{a15:0-i15:0 PE} and other family members

High-resolution mass spectrometry data were collected using Agilent MassHunter Work Station LC/MS Data Acquisition 10.1 and Agilent LC-QTOF Mass Spectrometer 6530 equipped with a 1290 uHPLC system and electrospray ionization detector scanning from m/z 50 to 3,200. Then 5 µl aliquots of \textit{a15:0-i15:0 PE} and its family members were injected into a reversed-phase analytical column (Luna C18 (2), 100 × 2.1 mm, 5 µm) using a gradient solvent system with 0.1% formic acid (10% methanol/water to 90% methanol/water for 10 min, 90% methanol/water isocratic for 10 min, then gradient to 100% for 10 min; flow rate, 0.3 ml). Agilent MassHunter Qualitative Analysis B.07.00 software was used to analyse the data.

FA methyl esterification and GC–MS analysis of \textit{A. muciniphila PE}

A 0.1 mg sample of both \textit{a15:0-i15:0 PE} and complete \textit{A. muciniphila PE} were dissolved in 200 µl of methanol, and 1.4 mg of sodium methoxide was added to prepare a 0.5 mol l⁻¹ sodium methoxide solution. The reaction mixture was stirred at room temperature for 3 h then quenched by addition of 1N HCl. The methanolysis products were dried under vacuum and extracted with ethyl acetate and water (300 µl, v/v = 2:1). The water layers were removed, and each of the ethyl acetate layers containing FA methyl esters (FAME) were injected into a gas chromatograph (GC, Agilent MassHunter GC/MS Acquisition B.07.05.2479) combined with a HP-5 Ultra Inert column (0.25 mm × 30 m). The temperature of the injector and the detector in the GC was maintained at 150 °C. During analysis, the temperature of the GC column was controlled (150 °C for 3 min, 150–250 °C at 6 °C min⁻¹ and 250 °C for 3 min). The FAME derivatives of \textit{a15:0-i15:0 PE} were composed of i15:0 and a15:0 (1:1 ratio) having retention times at 10.2 min and 9.7 min, respectively. The gas chromatography–mass spectrometry (GC–MS) analysis of FAME derivatives of AmPE displayed i14:0 (15.7%), n14:0 (2.7%), a15:0 (51.7%), sn-1:1 (23.6%), a16:0 (0.6%), i16:0 (1.8%), a17:0 (1.7%) and a18:0 (2.2%), having retention times at 8.0, 8.6, 9.7, 10.2, 11.3, 11.9, 13.4 and 15.0 min, respectively (Fig. 1d). Agilent MassHunter Qualitative Analysis B.07.00 software was used to analyse GC–MS data.

O-deacylation for determination of \textit{a15:0} connected to sn-1

A 5 mg sample of \textit{a15:0-i15:0 PE} was prepared and lyophilized for 24 h. A 1 mg ml⁻¹ of NaOMe solution was prepared, and the mixture was dissolved in 500 µl of NaOMe solution at room temperature. The solution was stirred under argon for 30 min. After 30 min, the reaction was quenched by addition of 1N HCl and dried under vacuum. The O-deacylated product, \textit{a15:0 PE}, was purified by reversed-phase HPLC (Luna C8 (2): 250 × 10 mm, 5 µm) with an isocratic solvent system (45% acetonitrile/water over 30 min, ultraviolet 210 nm detection, flow rate 2 ml min⁻¹). The O-deacylated product (1.8 mg) was eluted at 12.5 min, and its structure was determined by one-dimensional and/or
2D NMR spectroscopy (Extended Data Table 2) and by low-resolution electrospray ionization mass spectrometry (ESI-MS) ([M+H]+ m/z at 440; molecular formula, C32H42NO12P).

Amino-acid feeding experiment
A volume of 5 ml of A. muciniphila BAA-835 grown in BHI was inoculated into three 11 bottles of M9 medium supplemented with 1.5 g of mucin from porcine stomach (Sigma-Aldrich) and either 1 mmol L⁻¹ of l-leucine, l-isoleucine or L-leucine/L-isoleucine mixture (1:1 ratio) or nothing as a control. The cultures were grown anaerobically at 37°C for 12 days. The cell pellets from these cultures were centrifuged and extracted with 40 ml of chloroform and methanol (1:1). The extract was dried under vacuum and dissolved in dimethyl sulfoxide at a 10 mg mL⁻¹ concentration and tested for activity in the mBMDC cytokine assay. Statistical significances were determined using an unpaired two-tailed Student’s t-test.

a15:0-i15:0 PE biosynthetic gene identification and analysis
Sequence comparison and analysis of the a15:0-i15:0 PE biosynthetic pathway to the previously reported BCF4 biosynthetic pathway and de novo biosynthetic pathway of leucine, isoleucine and valine were performed using blastp (NCBI RefSeq database, updated 8 September 2019), Kyoto Encyclopaedia of Genes and Genomes and Geneious v.11.1.4 for pairwise sequence alignments that were previously reported. The accession number for the genes used in this analysis is CP001071.1.

Total synthesis for small library of PEs
The total synthesis of a15:0-i15:0 PE, i15:0-a15:0 PE, a15:0-a15:0 PE, i15:0-i15:0 PE and n15:0-n15:0 PE was performed by previously reported methods38–40.

Animal and human-cell studies
Mouse experimental procedures complied with all relevant ethical regulations and were conducted according to protocol 2003N000158 approved by the Institutional Animal Care and Use Committee at Massachusetts General Hospital. Appropriate sample sizes were estimated based on the effect size and variance of cytokine measurements in myeloid cells stimulated with canonical TLR ligands. In all mouse experiments, animals were allocated to experimental groups based on the effect size and variance of cytokine measurements in myeloid cells stimulated with canonical TLR ligands. In all mouse experiments, animals were allocated to experimental groups based on genotype and/or age and sex matched. Male or female wild-type, TLR2⁻/⁻ or TLR4⁻/⁻ C57BL/6 mice at least 3–4 weeks old and preferably 7–12 weeks of age were used. Mice were housed with a 12 h light or dark cycle at an ambient temperature of between 18°C and 24°C and a relative humidity of between 30% and 70%.

Human monocytes were isolated from buffy coats collected from healthy donors at the Blood Donor Center at Massachusetts General Hospital. Monocytes were enriched for myeloid cells stimulated with canonical TLR ligands. In all mouse experiments, animals were allocated to experimental groups based on genotype and/or age and sex matched. Male or female wild-type, TLR2⁻/⁻ or TLR4⁻/⁻ C57BL/6 mice at least 3–4 weeks old and preferably 7–12 weeks of age were used. Mice were housed with a 12 h light or dark cycle at an ambient temperature of between 18°C and 24°C and a relative humidity of between 30% and 70%.

Peripheral blood mononuclear cell cytokine assay
Peripheral blood mononuclear cells (PBMCs) were enriched for monocytes using the RosetteSep Human Monocyte Enrichment Cocktail (STEMCELL Technologies, catalogue no. 15028). In brief, buffy coats were incubated with monocyte enrichment cocktail for 20 min at room temperature while rocking. They were then diluted with IX phosphate-buffered saline (PBS) and layered over the Ficol-Paque PLUS medium (GE Healthcare, catalogue no. 17-1440-02) and centrifuged for 20 min at 1,200 g. Enriched monocytes were collected and cultured with chromatographic fractions or purified compounds at 50 µg ml⁻¹ in DMEM media containing 10% FBS and 1% penicillin-streptomycin. LPS and Pam3CSK4 at a final concentration of 100 ng ml⁻¹ were used as controls. After overnight incubation, supernatant was collected and analysed for IL-6, IL-10, IL-12/IL-23p40 and TNFα cytokines using Human Flex Set Kits (BD CBA, catalogue nos. 558276, 558274, 560154 and 560112).

RNA sequencing
Monocytes were isolated from PBMCs as described previously41. Bulk RNA sequencing libraries were prepared using SmartSeq2. Libraries were sequenced on a NextSeq (Illumina), FastQC v.0.11.5 and MultiQC v.1.8 were used to confirm the quality of the sequenced libraries42.43. Next, kallisto v.0.46.1 was used with a GRCh38 reference to generate the counts of reads mapped to each gene44.45. The matrix of counts was used for the calculation of counts per million (CPM) values, and the generated CPM matrix was treated with log(CPM + 1) to obtain a log expression matrix. A gene with a CPM value greater than 1 was considered as expressed. Samples obtained after the above steps were then used to detect differentially expressed genes via EdgeR v.3.35.1 (ref. 44). The lists of differentially expressed genes were generated from likelihood ratio tests based on the generative linear model framework, following the prerequisite gene filtering, normalization and dispersion estimation steps of the software.

CRISPR targeting
PBMCs were isolated from buffy coats using Sepmate tubes (STEMCELL Technologies) and ammonium–chloride–potassium lysis buffer following the manufacturer’s protocol. Human monocytes were harvested from PBMCs by negative selection using RosetteSep human Monocyte Enriched Cocktail (STEMCELL Technologies) according to the manufacturer’s protocol. Alt-R sgRNAs were purchased from IDT and reconstituted to 100 µmol l⁻¹ with Nuclease-Free Duplex Buffer (IDT). In a sterile polymerase chain reaction strip, the sgRNAs were mixed with Cas9 (IDT, Alt-R S.p. Cas9 Nuclease V3) at a molar ratio of 2:1 (2 µg sgRNA at 100 µmol l⁻¹ + 2 µg Cas9 at 5 mg ml⁻¹) for each reaction and incubated at room temperature for over 20 min. Monocytes were washed twice with 5 ml of PBS and counted. Then 2 × 10⁶ cells per reaction were resuspended in 16 µl of P3 primary nucleofection solution.
The 16 µl of cells in P3 buffer was added to each Cas9–ribonuclease complex. The cell–ribonuclease protein mix was then immediately loaded into the supplied nucleofector cassette strip (Lonza) and nucleofected using 4D-Nucleofector with CM-137 programme. Then 180 µl of prewarmed medium was immediately added into each cassette well. A volume of 1 × 10^5 cells was seeded into a 96-well plate with medium (RPMI-1640 with 10% FBS, 2 mmol l^-1 Glutamax, 55 µmol l^-1 beta-mercaptoethanol, 100 µl ml^-1 penicillin, 100 µg ml^-1 streptomycin, GM-CSF 800 U ml^-1 and IL-4 500 U ml^-1). The medium was changed every 2–3 days. At day 5, MDDCs were stimulated with 10 µg ml^-1 of Akkermansia lipids, 100 ng ml^-1 of Pam3CSK4, 100 ng ml^-1 of FSL-1 or 100 ng ml^-1 of LPS for 18 h or as indicated. Cell supernatants were collected for human TNFα measurements by ELISA (Invitrogen) following the manufacturer’s protocol. SoftMax Pro v.6.2.1 (SpectraMax, Molecular Devices) was used to analyse ELISA plates. The sgRNA sequences used were as follows:

- Human TLR1: GGTCTTAGGAGAGACTTATG
- Human TLR2: GACCGCAATGGTATCTGCAA
- Human TLR6: ATTCATTTCCGTCGGAGAAC

TLR2–TLR1–a15:0-i15:0 PE complex modelling

Modelling of the a15:0-i15:0 PE ligand complex was based on the crystal structure of the TLR2–TLR1–Pam3CSK4 complex from the Protein Data Bank (PDB ID 2z7x)38. The Pam3CSK4 ligand was removed from the crystal structure coordinates, and an a15:0-i15:0 PE ligand was prepared using Lidia and AceDRG in Coot v.0.9 (refs. 46,47). The a15:0-i15:0 PE ligand placement in the ligand-binding pockets of TLR2 and TLR1 was guided by the electron density belonging to the acyl chains of the Pam3CSK4 ligand in the crystal structure. Structural figures and videos were generated using ChimeraX v.1.0 (ref. 48). Structural biology software was compiled and configured by SBGrid consortium49.

**Reporting summary**

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

**Data availability**

RNA sequencing data generated during this study are available in the NCBI Gene Expression Omnibus (GEO, GSE199367). NMR data generated during this study are available in Extended Data Tables 1 and 2. Complex modelling was based on the crystal structure from the Protein Data Bank (PDB ID 2z7x). The complete A. muciniphila BAA-835 genome was obtained from GenBank (CP001071.1).

**Acknowledgements**

We thank J. Moore of the Systems Bio FACS core at Harvard Medical School for assistance in setting up experiments. We thank A. Omar, J. Sulca, B. Li and E. Creasey at Massachusetts General Hospital for their assistance with mouse breeding, husbandry and tissue harvesting. We thank E. Helfrich for biosynthetic pathways discussions. This work was supported by the National Institutes of Health (grant no. R01AT009708 to J.C. and R.J.X., P30DK043351, R01DK127771, R01AI172947, and RC2DK114784 to R.J.X. and F32AT010415 to C.D.C.) and the Linde Family Programme in Cancer Chemical Biology (to S.J.B.).

**Author contributions**

M.B., C.D.C., X.L., J.K., H.V., D.B.G., S.J.B., R.J.X. and J.C. designed the experiments. M.B. and J.K. performed the isolation, purification and characterization in Fig. 1 and Extended Data Fig. 1. M.B. also analysed the genome for the biosynthetic pathways, performed the feeding experiment and conducted the mass spectrometry shown in Fig. 2 and Extended Data Fig. 3 and 4. C.D.C. performed the experiments in Figs. 1 and 2 and Extended Data Fig. 5. X.L. prepared the synthetic a15:0-i15:0 PE and analogues in Fig. 2. B.K.T. and A.S.B. helped perform the experiments shown in Fig. 1 and Extended Data Fig. 2. X.C. S.-M.P. and Z.L. performed the experiments shown in Extended Data Figs. 6 and 7. P.F. performed the structural ligand-binding rendering in Fig. 3. M.B., C.D.C. and J.C. wrote the manuscript. D.B.G and R.J.X. revised and edited the manuscript.

**Competing interests**

Four of the authors (M.B., C.D.C., R.J.X. and J.C.) have submitted a patent application for materials described in this manuscript. R.J.X. is co-founder of Inana Therapeutics and Celsius Therapeutics and is a consultant to Nestle; these organizations had no roles in this study. All other authors declare no competing interests.

**Additional information**

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-04985-7.

Correspondence and requests for materials should be addressed to Ramnik J. Xavier or Jon Clardy.

Peer review information Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Characterization of a15:0-i15:0 PE. Key COSY, HSQC, and HMBC correlations of a15:0-i15:0 PE. All experiments were repeated independently at least twice with similar results.
Extended Data Fig. 2 | Human-cell activation by PE of *A. muciniphila*.
a15:0-l5:0 PE induces TNFα, IL-6, IL-10, and IL-12/23p40 from human myeloid cells as measured by flow cytometry. Data are presented as mean values ± s.d. of technical replicates (n = 3). Experiments were repeated independently at least twice with similar results.
Extended Data Fig. 3 | Biosynthetic pathway of a15:0-i15:0 PE in *A. muciniphila*. Proposed biosynthesis of a15:0-i15:0 PE in *A. muciniphila* BAA-835.
Extended Data Fig. 4 | Biosynthetic pathways of branched-chain amino acids in *A. muciniphila*. *De novo* biosynthesis of l-leucine, l-isoleucine, and l-valine in *A. muciniphila* BAA-835.
Extended Data Fig. 5 | TNFα activation of straight chain fatty acids. TNFα production by mBMDCs treated with single chain fatty acids (from C17:0 to C13:0) as measured by ELISA. LPS was used as a control agonist. Data are representative of at least two independent experiments and show mean values ± s.d. of technical replicates (n = 4).
Extended Data Fig. 6 | Heat-map of human derived monocyte activation by natural and synthetic *A. muciniphila* lipids. Monocytes were purified from peripheral blood by negative selection with magnetic beads. Cells were immediately placed in culture and stimulated with the indicated stimuli for 6 h. After stimulation, cells were lysed for mRNA extraction and library preparation by SmartSeq2 for RNA sequencing. Monocytes were processed from two independent donors in technical triplicate. The final concentration of bacterial lipids was 50 µg/ml. Pam3CSK4, FSL-1, and LPS at a final concentration of 100 ng/ml were used as controls.
Extended Data Fig. 7 | Effects of treatment with synthetic *A. muciniphila* lipids in combination with Pam3CSK4 or LPS on human monocyte-derived dendritic cells. With long (18 h) delay times, low doses of synthetic a–c, a15:0-i15:0 PE or d–f, i15:0-a15:0 PE suppress immune responses to Pam3CSK4 and moderate immune responses to LPS. Both effects disappear with shorter delay times (3 h or none). LPS and Pam3CSK4 were used at final concentrations of 100 ng/mL. Data are presented as mean values ± s.d. of technical replicates (n = 4). Experiments were repeated independently at least twice with similar results.
### Extended Data Table 1 | NMR spectroscopic data for identification of the a15:0-i15:0 PE structure

| No. | $J_\text{in Hz}$ | $\delta_C$ | $\delta_H$ |
|-----|-----------------|------------|------------|
| 1a  | 4.37, dd (11.5, 2.5) | 62.8 | CH$_2$ |
| 1b  | 4.14, dd (11.5, 6.5) | 70.6 | CH |
| 2   | 5.21, m          | 64.1 | CH$_2$ |
| 3   | 4.07, br         | 62.4 | CH$_2$ |
| 4   | 3.15, br         | 40.6 | CH$_2$ |
| 1'  | 2.29, m          | 173.6 | C |
| 1'' | 1.58, m          | 34.6 | CH$_2$ |
| 4'  | 1.31-1.21, overlapped | 30.3-27.4 | CH$_2$ |
| 11'a| 1.26, m          | 36.7 | CH$_2$ |
| 11'b| 1.07, m          | 34.5 | CH |
| 12' | 1.27, m          | 29.5 | CH$_2$ |
| 13' | 1.13, m          | 11.6 | CH$_3$ |
| 14' | 0.83, t (6.5)    | 19.4 | CH$_2$ |
| 15' | 0.87, d (6.5)    | 173.3 | C |
| 1'' | 2.29, m          | 34.3 | CH$_2$ |
| 3'' | 1.58, m          | 25.1 | CH$_2$ |
| 4''-11'' | 1.31-1.21, overlapped | 30.3-27.4 | CH$_2$ |
| 12''| 1.16, m          | 28.2 | CH |
| 13''| 1.51, m          | 22.9 | CH$_3$ |
| 14''| 0.85, d (6.5)    | 22.8 | CH$_3$ |

$^a$ $^1$H: 500 MHz,

$^b$ $^{13}$C: 125 MHz

---

![Chemical Structure](image)

**a15:0-i15:0 PE**

$^1$H and $^{13}$C NMR data (500 MHz) of a15:0-i15:0 PE in CDCl$_3$. 
Extended Data Table 2 | NMR spectroscopic data for identification of the O-deacylated product structure

| No. | δH, mult (J in Hz)\(^a\) | δC\(^b\) |  |
|-----|----------------|--------|---|
| 1a  | 4.16, dd (11.5, 4.5) | 63.4 | CH₂ |
| 1b  | 4.10, dd (11.5, 6.5) |      |    |
| 2   | 3.96, m            | 70.0  | CH |
| 3   | 3.90, m            | 68.0  | CH₂ |
| 4   | 4.05, m            | 63.41 | CH₂ |
| 5   | 3.14, br           | 42.0  | CH₂ |
| 1'  |                  | 175.6 | C  |
| 2'  | 2.35, t (7.5)      | 35.1  | CH₂ |
| 3'  | 1.61, m            | 26.1  | CH₂ |
| 4'-10' | 1.31-1.21      | 33.2-28.3 | CH₂ |
| 11'a| 1.26, m            | 37.9  | CH₂ |
| 11'b| 1.10               |       |    |
| 12' | 1.27, m            | 35.8  | CH |
| 13' | 1.13               | 30.7  | CH₂ |
| 14' | 0.83, t (6.5)      | 11.8  | CH₃ |
| 15' | 0.87, d (6.5)      | 19.8  | CH₃ |

\(^a\) ¹H: 500 MHz, ¹³C: 125 MHz

\(^b\) ¹H NMR data (500 MHz) and ¹³C NMR data (500 MHz) of a15:0 PE in CD₃OD.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

- GenS 3.03 or SoftMax Pro 6.2.1 was used to analyze ELISA plates; Agilent Mass Hunter Work Station LC/MS Data Acquisition 10.1 and Agilent LC-QTOF Mass Spectrometer 6530 to collect HRMS data; Agilent Mass Hunter GC/MS Acquisition B.07.05.2479 to collect GC/MS data; and NovoExpress 1.4.1 to collect flow cytometry data

Data analysis

- Adobe Illustrator 2020 was used to assemble figures; GraphPad Prism 8 and Microsoft Excel 2016 to perform statistical analyses; Agilent Mass Hunter Qualitative Analysis B.07.00 to analyze HRMS and GC/MS data; FastQC v0.11.5 and MultiQC v1.8 to confirm quality of RNA sequencing libraries; kallisto v0.46.1 and EdgeR v3.35.1 to analyze RNA sequencing data; Geneious 11.1.4 to perform genome analyses; Mnova 14.2.0 to analyze NMR data of natural/synthetic compounds; FlowJo v10.7 to analyze flow cytometry data; Coot 0.9 to model ligand-receptor complex; and ChimeraX 1.0 to generate structural figures

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA sequencing data generated during this study are available in the NCBI Gene Expression Omnibus (GEO, GSE199367) at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE199367. NMR data generated during this study are available in Extended Data Tables 1 and 2. Complex modeling was based on the crystal...
Structure from Protein Data Bank (PDB ID: 2Z7X) at https://www.rcsb.org/structure/2Z7X. The complete A. muciniphila BAA-835 genome was obtained from GenBank (CP001071.1) at https://www.ncbi.nlm.nih.gov/nuccore/CP001071.1.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences      ☐ Behavioural & social sciences      ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Appropriate sample sizes were estimated based on the effect size and variance of cytokine measurements in myeloid cells stimulated with canonical TLR ligands.

Data exclusions

All relevant data were included in this study.

Replication

All experiments in this manuscript were performed at least twice and demonstrated the same or similar results as those published here.

Randomization

In all mouse experiments, animals were allocated into experimental groups based on genotype and/or age- and sex-matched. Human blood samples were obtained from de-identified, volunteer donors.

Blinding

For all biological experiments, investigators performing the experiment were blind to the identities of the samples being tested.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| Involved in the study |
|-----------------------|
| n/a                   |
| ☐ Antibodies          |
| ☑ Eukaryotic cell lines |
| ☑ Palaeontology and archaeology |
| ☑ Animals and other organisms |
| ☑ Human research participants |
| ☐ Clinical data       |
| ☑ Dual use research of concern |

Methods

| Involved in the study |
|-----------------------|
| n/a |
| ☑ ChIP-seq |
| ☐ Flow cytometry |
| ☑ MRI-based neuroimaging |

Antibodies

Antibodies were used in immunological assays with mBMDCs using Invitrogen's Mouse TNFalpha Uncoated ELISA kit, catalog # 88-7324.

Validation

All the commercial antibodies are validated by the manufacturers and came with quality assurance statements.

Animals and other organisms

Policy information about studies involving animals: ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Femurs and tibias were collected from male or female wild-type, TLR2-/-, or TLR4-/- C57BL/6 mice at least 3-4 weeks old and preferably 7-12 weeks of age. Mice were housed with a 12-hour light/dark cycle at an ambient temperature between 65-75°F and 30-70% relative humidity.

Wild animals

No wild animals were used in the study.

Field-collected samples

No field collected samples were used in the study.

Ethics oversight

Mouse experimental procedures complied with all relevant ethical regulations and were conducted according to protocol 2003N00158 approved by the Institutional Animal Care and Use Committee (IACUC) at Massachusetts General Hospital.
Human research participants

Policy information about studies involving human research participants

| Population characteristics | This is not applicable to our study. We obtained blood samples from de-identified, volunteer donors in order to isolate human monocytes for in vitro experiments. |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recruitment               | Participants were volunteer blood donors.                                                                                                                                                         |
| Ethics oversight          | Human monocytes were isolated from buffy coats collected from healthy donors at the Blood Donor Center at Massachusetts General Hospital in compliance with all relevant ethical regulations and according to protocol 2018P001504 approved by the Mass General Brigham Institutional Review Board (IRB). Donors provided informed written consent. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a ‘group’ is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

| Sample preparation          | CBA samples were prepared following the instruction manual of the BD Cytometric Bead Array (CBA) Mouse/Rat Soluble Protein Master Buffer Kit. |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------|
| Instrument                 | Agilent NovoCyte Flow Cytometer                                                                                                                                                                  |
| Software                   | NovoExpress 1.4.1, FlowJo v10.7                                                                                                                                                                   |
| Cell population abundance  | CBA populations were abundant for analysis following the instruction manual of the BD CBA Mouse/Rat Soluble Protein Master Buffer Kit.                                                        |
| Gating strategy            | CBA populations were gated following the instruction manual of the BD CBA Mouse/Rat Soluble Protein Master Buffer Kit.                                                                    |

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.